REFLEXIONES SOBRE EL SISTEMA EDUCATIVO ESPAÑOL

DANIEL SANTÍN
PAU BALART
ANTONIO CABRALES
JORGE CALERO
ÁNGEL DE LA FUENTE
J. ORIOL ESCARDÍBUL
FLORENTINO FELGUEROSO
GABRIELA SICILIA

FUNDACIÓN RAMÓN ARECES
FUNDACIÓN EUROPEA SOCIEDAD Y EDUCACIÓN

REFLEXIONES SOBRE EL SISTEMA EDUCATIVO ESPAÑOL

DIRECCIÓN DEL PROYECTO *REFLEXIONES SOBRE EL SISTEMA EDUCATIVO ESPAÑOL*
Fundación Europea Sociedad y Educación

COORDINACIÓN EDITORIAL
Mercedes de Esteban Villar
Instituto de Estudios Educativos y Sociales. Fundación Europea Sociedad y Educación

AUTORES
Daniel Santín. Dirección científica
Pau Balart
Antonio Cabrales
Jorge Calero
Ángel de la Fuente
J. Oriol Escardíbul
Florentino Felgueroso
Gabriela Sicilia

REFLEXIONES SOBRE EL SISTEMA EDUCATIVO ESPAÑOL

DANIEL SANTÍN
PAU BALART
ANTONIO CABRALES
JORGE CALERO
ÁNGEL DE LA FUENTE
J. ORIOL ESCARDÍBUL
FLORENTINO FELGUEROSO
GABRIELA SICILIA

FUNDACIÓN RAMÓN ARECES
FUNDACIÓN EUROPEA SOCIEDAD Y EDUCACIÓN

El contenido expuesto en este libro es responsabilidad exclusiva de sus autores.

Reservados todos los derechos.

Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética, o cualquier almacenamiento de información y sistema de recuperación, sin permiso escrito de la Editorial Centro de Estudios Ramón Areces y de la Fundación Europea Sociedad y Educación.

EDICIÓN 2015

© Editorial Centro de Estudios Ramón Areces S.A.
Tomás Breton, 21 - 28045 Madrid
T 915 398 659
F 914 681 952
cerasa@cerasa.es
www.cerasa.es

© Fundación Ramón Areces
Vitruvio, 5 - 28006 Madrid
www.fundacionareces.es

© Fundación Europea Sociedad y Educación
José Abascal, 57 - 28003 Madrid
www.sociedadyeducacion.org

© Autores

Diseño:
KEN / www.ken.es

Depósito legal: M-17250-2015

Impreso por:
ANEBRI, S.A.
Antonio González Porras, 35-37
28019 Madrid
Impreso en España / Printed in Spain

ÍNDICE

PRÓLOGO **9**

Fundación Ramón Areces
Fundación Europea Sociedad y Educación

PRIMERA PARTE **13**

CAPITAL HUMANO Y PRODUCTIVIDAD 15
Ángel de la Fuente

SEGUNDA PARTE. REFLEXIONES SOBRE EDUCACIÓN Y CAPITAL HUMANO **43**

El impacto de la educación infantil en los resultados de primaria: evidencia para España a partir de un experimento natural 45
Daniel Santín y Gabriela Sicilia

La maratón de Pisa: la perseverancia como factor del éxito en una prueba de competencias 75
Pau Balart y Antonio Cabrales

Recursos escolares y resultados de la educación 111
Jorge Calero y J. Oriol Escardíbul

TERCERA PARTE. REFLEXIONES SOBRE EL DISEÑO Y LA EVALUACIÓN DE LAS POLÍTICAS EDUCATIVAS **161**

Igualdad de oportunidades: inclusión y exclusión educativas en España 163
Jorge Calero y J. Oriol Escardíbul

La evaluación docente basada en el resultado como vía de mejora del sistema educativo 213
Pau Balart y Antonio Cabrales

Claves para mejorar la educación y formación de adultos en España en la post-crisis 247
Florentino Felgueroso

Evaluar para mejorar: hacia el seguimiento y la evaluación sistemática de las políticas educativas 283
Daniel Santín y Gabriela Sicilia

PRÓLOGO

Las fundaciones *Ramón Areces* y *Europea Sociedad y Educación* han coordinado durante dos años el trabajo de un equipo de investigación con el objetivo de profundizar en las relaciones entre rendimiento educativo y crecimiento económico, desde dos perspectivas fundamentales: por una parte, aquella que vincula el desarrollo y la competitividad con medidas que inciden directamente en la mejora del capital humano a través de la acción educativa; por otra, la que depende de los efectos de las políticas educativas, incluyendo la identificación, selección, implantación y evaluación de las más exitosas, a partir de evidencias procedentes de estudios comparados.

Con la creación de este grupo de trabajo, ambas entidades pretendieron impulsar una línea de investigación puntera sobre la relación causal entre educación y desarrollo económico, abriendo un proceso de reflexión en torno a un abanico de dimensiones y variables que orientaran algunas líneas de mejora sobre las que operar. Poner en valor los resultados de la literatura científica acerca de los cada vez más estrechos vínculos entre economía y educación; inspirar ideas para consolidar el proceso de recuperación económica mediante actuaciones específicas en el seno de los sistemas educativos y presentar nuevas propuestas que, desde la óptica de la economía de la educación, contribuyeran a superar con garantías de éxito las diferentes situaciones existentes que intervienen directa o indirectamente en el crecimiento y competitividad de las sociedades avanzadas, han sido algunos de los objetivos de esta obra. Para ello, se partió de la formulación de una serie de cuestiones básicas, inherentes a cualquier política de mejora: ¿qué tipo de educación y qué tipo de competencias predicen un crecimiento sostenible y una más equitativa distribución de oportunidades para el capital humano de un país?, ¿qué tipo de reformas educativas resultan exitosas y responden mejor a los desafíos educativos?, ¿qué factores de contexto deben ser tenidos en cuenta para implantar políticas educativas?

Un grupo de economistas especializados en economía de la educación y sistemas educativos abordan en esta obra distintos aspectos clave en las políticas educativas y sus efectos en el logro escolar, contando con información procedente de fuentes, datos, evidencias y conclusiones aportados por la literatura y la investigación científica, así como de indicadores y estudios empíricos nacionales e internacionales, en muchos casos pendientes de explotación por parte de la investigación económica y educativa. El objetivo último de esta publicación es ofrecer una guía para el lector no especialista, para responsables de la toma de decisiones y para los profesionales de la educación, sobre aquellos aspectos que facilitan la puesta en

marcha de políticas educativas fundamentadas en la investigación y las evidencias empíricas. Ambas fundaciones agradecen especialmente la dedicación de los autores de los trabajos así como la de los editores, y de los que de un modo u otro han participado en este proyecto.

Esta publicación se estructura en dos grandes apartados precedidos de una primera sección introductoria de la que es autor Ángel de la Fuente, donde se nos recuerda lo que la investigación ya ha demostrado en materia de educación y crecimiento económico. El primero de ellos, *Reflexiones sobre educación y capital humano*, atiende aspectos relacionados con factores que afectan directa o indirectamente a los usuarios del sistema educativo español: la escolarización en edades tempranas (Santín y Sicilia), la consideración del impacto de las cada vez más decisivas variables no-cognitivas en el rendimiento (Balart y Cabrales), y los recursos escolares, con destino individual y colectivo, y su impacto en los resultados (Calero y Escardíbul). En resumen, la mejora de la dotación del capital humano español, resultante de su paso por el sistema educativo, tanto en términos cuantitativos como cualitativos, se presenta como un desafío permanente para la economía española.

El segundo bloque *Reflexiones sobre el diseño y la evaluación de las políticas educativas* se centra en los fundamentos que deberían estar presentes en cualquier política educativa orientada a la mejora continua de las competencias y conocimientos de los estudiantes, así como el marco institucional, los recursos, procesos y el diseño de la evaluación que permitan su puesta en marcha y la obtención de resultados (Balart y Cabrales). A este fin responden también las conclusiones de los trabajos sobre inclusión y exclusión en España (Calero y Escardíbul), sobre evaluación y mejora (Santín y Sicilia), y sobre las competencias de nuestra población adulta y su relación con la reciente reforma de la formación para el empleo (Felgueroso).

El libro se completa con un apéndice que trata de mostrar cuál es la situación actual de la educación española, presentando un resumen breve de los principales gráficos de evolución o comparación y de datos y cifras del contexto educativo español. Para ello se han seleccionado algunos indicadores clave relacionados con los recursos empleados, la población que recibe el servicio en sus distintos niveles y los resultados educativos en términos de rendimiento académico, fracaso escolar, abandono educativo y formación de la población adulta.

La mejora de la productividad debe ser una de las prioridades más importantes de la política económica de España. Existe una creciente evidencia empírica de que la educación constituye un elemento significativo tanto en aquélla como en el empleo, habida cuenta del aumento de una demanda de producción basada en el intercambio del bien "conocimiento" y en la cualificación permanente del capital humano. Esta obra colectiva se planteó, además, como una respuesta desde la

investigación económico-educativa a las recomendaciones de la Unión Europea que ponen el acento en la puesta en marcha de iniciativas, basadas en la reflexión intelectual y académica, que permitan orientar la toma de decisiones de los responsables de la educación, a partir del conocimiento de la realidad nacional y de las expectativas de los usuarios de los sistemas de educación y formación. Esperamos haber cumplido con ambos objetivos.

FUNDACIÓN RAMÓN ARECES
FUNDACIÓN EUROPEA SOCIEDAD Y EDUCACIÓN

PRIMERA PARTE

CAPITAL HUMANO Y PRODUCTIVIDAD

CAPITAL HUMANO Y PRODUCTIVIDAD*

ÁNGEL DE LA FUENTE**
FEDEA e Instituto de Análisis Económico (CSIC)

RESUMEN

Este trabajo contiene un breve panorama de la literatura empírica sobre capital humano y productividad y resume los resultados de mi propio trabajo sobre el tema. En conjunto, la evidencia disponible sugiere que la inversión en educación tiene un efecto positivo, significativo y cuantitativamente importante sobre el crecimiento de la productividad. De acuerdo con mis estimaciones, además, la rentabilidad social de la inversión en capital humano es mayor que la del capital físico en la mayor parte de los países de la UE y en muchas regiones españolas.

INTRODUCCIÓN

Una de las características distintivas de las "nuevas" teorías del crecimiento económico ha sido el enriquecimiento del concepto relevante de capital. Mientras que los modelos neoclásicos tradicionales se centraban casi exclusivamente en la acumulación de capital físico (estructuras y maquinaria), las contribuciones más recientes han atribuido una importancia creciente a la acumulación de capital humano y conocimientos productivos, así como a la interacción entre estos dos factores intangibles.

* *Clasificación* JEL: O40, I20, O30, C19.

** El presente trabajo, basado en de la Fuente (2011), ha sido preparado para el proyecto "Reflexiones sobre el sistema educativo español", promovido por las fundaciones Ramón Areces y Europea Sociedad y Educación.

La evidencia empírica, sin embargo, no siempre ha sido consistente con los nuevos modelos teóricos. En el caso del capital humano, en particular, algunos estudios han obtenido resultados bastante desalentadores. Con frecuencia, las variables educativas no son significativas o incluso entran con el signo "equivocado" en ecuaciones de crecimiento, especialmente cuando éstas se estiman utilizando especificaciones en diferencias o técnicas de panel. La acumulación de resultados negativos en la literatura durante la segunda mitad de los noventa generó un creciente escepticismo sobre el papel de la educación en el proceso de crecimiento e incluso llevó a algunos autores (véase en particular Pritchett, 2001) a considerar seriamente las razones por las que la inversión educativa podría no contribuir al crecimiento de la productividad.

Muchos investigadores, sin embargo, mantuvimos una visión más optimista y argumentamos que los resultados negativos de ciertos estudios podían atribuirse a problemas técnicos que tienen mucho que ver con la dificultad de medir correctamente el capital humano. El presente artículo contiene un repaso rápido de varias ramas de la literatura que proporcionan evidencia consistente con esta hipótesis y un resumen algo más detallado de mi trabajo sobre el tema. El artículo está organizado como sigue. La sección 2 esboza el marco teórico que ha guiado la mayor parte de los estudios sobre la contribución de la educación al crecimiento económico, repasa las principales especificaciones empíricas que se han utilizado en la literatura y discute brevemente algunos de sus resultados clave. La sección 3 destaca algunos de los problemas de las series de escolarización más utilizados en la literatura empírica más temprana, discute sus implicaciones para los intentos de estimar la contribución de la educación al crecimiento de la productividad e introduce un conveniente indicador de calidad de los datos que puede utilizarse para cuantificar el contenido informativo de distintas series de escolarización y para estimar el tamaño del sesgo generado por la existencia de errores de medición. La sección 4 resume los principales resultados de una serie de trabajos que he escrito en su mayor parte en colaboración con Rafael Doménech. En ellos construimos nuevas series de escolarización para una muestra de más de veinte países de la OCDE y para las regiones españolas, elaboramos indicadores del contenido informativo de estas y otras series de escolarización y estimamos diversas especificaciones de crecimiento con ambas muestras. Utilizando estos resultados, hemos construido una serie de meta-estimaciones de los coeficientes del capital humano en una función de producción agregada del tipo Cobb-Douglas que corrigen el sesgo a la baja generado por el error de medición. Con esta corrección, encontramos que la contribución de la inversión en capital humano al crecimiento de la productividad es positiva, de un tamaño muy apreciable, e implica una rentabilidad social muy respetable que, para la mayor parte de los territorios en nuestras

dos muestras, se compara muy favorablemente con la ligada a la inversión en capital físico, lo que sugiere que existen importantes externalidades positivas.

CAPITAL HUMANO Y CRECIMIENTO ECONÓMICO: UN REPASO RÁPIDO DE LA LITERATURA

Los modelos teóricos que analizan el papel del componente educativo del capital humano en el proceso de desarrollo económico se construyen a partir de la hipótesis de que los conocimientos y competencias de la población trabajadora inciden positivamente sobre su productividad y sobre su capacidad para desarrollar y adoptar nuevas tecnologías[1]. Para explorar sus implicaciones y abrir el camino a su contrastación empírica, esta hipótesis se formaliza generalmente en una de dos formas no excluyentes. La más sencilla consiste en introducir el *stock* de capital humano (al que llamaré H en este trabajo) como un argumento adicional en una función de producción agregada por lo demás estándar que relaciona el producto nacional o regional con los *stocks* de factores productivos (generalmente el empleo y el capital físico) y con un índice de eficiencia técnica o productividad total de los factores (PTF). La segunda posibilidad consiste en incluir H en el modelo como uno de los determinantes de la tasa de progreso técnico (esto es, de la tasa de crecimiento de la PTF). Esto requiere la especificación de una función de progreso técnico que puede incluir variables adicionales tales como la inversión en I+D y la brecha existente entre cada país y la frontera tecnológica mundial.

En lo que sigue, me referiré a la primera de estas conexiones entre capital humano y productividad como *efectos de nivel* (puesto que el *stock* de capital humano tiene un efecto directo sobre el nivel de *output*) y a la segunda de ellas como *efectos de tasa* (porque H afecta ahora a la tasa de crecimiento del *output* a través de la PTF). El Recuadro 1 desarrolla un simple modelo de crecimiento con capital humano que formaliza la discusión precedente e incorpora ambos efectos.

Algunos modelos teóricos recientes sugieren que la acumulación de capital humano puede generar externalidades importantes que podrían justificar la intervención pública correctiva. El problema surge porque algunos de los beneficios de una fuerza laboral más educada tienden a "escaparse" y generan ganancias de *output* que no pueden ser apropiadas en la forma de mayores ingresos por aquellos que financian las inversiones relevantes, dando lugar a una "brecha" entre los rendimientos privados y sociales de la educación. Lucas (1988), por ejemplo, sugiere que el *stock* medio de capital humano a nivel agregado aumenta la productividad

1. Véase entre muchos otros Nelson y Phelps (1966), Lucas (1988) y Jones (1996).

de cada empresa manteniendo constante el *stock* propio de capital humano. También se suele suponer que los efectos de tasa del capital humano sobre el ritmo del progreso técnico incluyen un componente importante de externalidad porque es difícil apropiarse privadamente de todo el valor económico de las ideas. Azariadis y Drazen (1990), e implícitamente Lucas (1988) también, destacan que las cohortes más jóvenes se benefician del conocimiento y las destrezas acumuladas por sus mayores, generando externalidades intergeneracionales potencialmente importantes que operan tanto en casa como en la escuela. La literatura también sugiere que el capital humano puede generar externalidades "cívicas" más difusas dado que un aumento del nivel educativo de la población puede ayudar a reducir las tasas de criminalidad o contribuir al desarrollo de instituciones más efectivas.

RECUADRO 1. UN MODELO DESCRIPTIVO DE CAPITAL HUMANO Y CRECIMIENTO

En este recuadro se desarrolla un sencillo modelo de crecimiento con capital humano que tiene dos componentes: una función de producción agregada y una función de progreso técnico. Supondré que la función de producción es de tipo Cobb-Douglas:

$$(R.1)\ Y_{it} = A_{it}\, K_{it}^{\alpha k} H_{it}^{\alpha h} L_{it}^{\alpha l}$$

donde Y_{it} denota el *output* del territorio *i* en el momento *t*, L_{it} el nivel de empleo, K_{it} el *stock* de capital físico, H_{it} el *stock* de capital humano por trabajador, medido generalmente por el nivel educativo medio, y A_{it} un índice un índice de eficiencia técnica o productividad total de los factores (PTF) que resume el estado actual de la tecnología y recoge factores omitidos tales como la localización geográfica, el clima o la calidad de las instituciones. Los coeficientes α_i (con $i = k, h, l$) miden la elasticidad del *output* con respecto a los *stocks* de los distintos factores. Un aumento del 1% en el *stock* de capital humano por trabajador, por ejemplo, aumentaría el *output* en un α_h%, manteniendo constantes los *stocks* de los demás factores productivos y el nivel de eficiencia técnica.

Bajo la hipótesis habitual de que (R.1) presenta rendimientos constantes a escala en capital físico y trabajo manteniendo constante el nivel medio de formación (esto es, que $\alpha_k + \alpha_l = 1$), podemos definir una función de producción per cápita que relaciona la productividad media del trabajo con el nivel medio de formación y el *stock* de capital por trabajador. Sea $Q = Y/L$ el *output* por trabajador y $Z = K/L$ el *stock* de capital por trabajador. Dividiendo ambos lados de (R.1) por el empleo total, *L*, tenemos:

$$(R.2)\ Q = AZ^{\alpha k} H^{\alpha h}$$

La función de progreso técnico describe los determinantes del crecimiento de la productividad total de los factores. Supondré que el nivel de PTF del país *i* puede escribirse en la forma

$$(R.3) A_{it} = B_t X_{it}$$

[CONTINÚA EN PÁGINA SIGUIENTE]

RECUADRO 1. UN MODELO DESCRIPTIVO DE CAPITAL HUMANO Y CRECIMIENTO

donde B_t denota la frontera tecnológica mundial (esto es, el nivel máximo alcanzable de eficiencia en la producción dado el nivel actual de conocimiento científico y tecnológico) y $X_{it} = A_{it}/B_t$ es (un indicador inverso de) la "brecha tecnológica" entre el país *i* y la frontera. Supondremos que B_t crece a una tasa constante, *g*, y que la tasa de crecimiento de X_{it} viene dada por

$$(R.4)\ \Delta x_{it} = \gamma_{io} - \lambda x_{it} + \gamma H_{it}$$

donde x_{it} es el logaritmo de X_{it} y γ_{io} es un efecto fijo de país que ayuda a controlar por variables omitidas, tales como la inversión en I+D. Obsérvese que esta especificación incorpora un efecto de difusión tecnológica o *catch-up*. Si $\lambda > 0$, los países que están más cerca de la frontera tecnológica experimentarán tasas más reducidas de crecimiento de la PTF, lo que hará que los niveles de PTF tiendan a estabilizarse en relación a la frontera con el paso del tiempo y que sus valores de estado estacionario sean determinados parcialmente por el nivel educativo.

DE LA TEORÍA A LOS DATOS: ENFOQUES ALTERNATIVOS PARA EL ANÁLISIS EMPÍRICO

Los estudios empíricos sobre la contribución del capital humano a la productividad (o más generalmente sobre los determinantes del crecimiento económico) han seguido uno de dos enfoques alternativos. El primero de ellos se basa en la especificación y estimación de una ecuación *ad hoc* que relaciona el crecimiento del *output* total o per cápita con un conjunto de variables potencialmente relevantes que se identifican en base a consideraciones teóricas informales. En el segundo enfoque, por contra, lo que se intenta estimar es una relación estructural entre el nivel de *output* o su tasa de crecimiento y un conjunto de variables explicativas que se deriva de un modelo teórico explícito, construido a partir de una función de producción y, posiblemente, de una relación de progreso técnico del tipo descrito en el Recuadro 1.

Este marco básico para el análisis "estructural" de los determinantes del crecimiento puede dar lugar un gran número de especificaciones empíricas tal como se discute en el Recuadro 2. La función de producción puede estimarse directamente con las variables relevantes expresadas en niveles o en tasas de crecimiento siempre que dispongamos de datos sobre los *stocks* de los distintos factores productivos. Alternativamente, sus parámetros pueden recuperarse a partir de otras especificaciones (ecuaciones de convergencia o de estado estacionario) que están diseñadas para permitir la estimación cuando sólo se dispone de datos sobre flujos inversores en vez de sobre *stocks* de factores, o cuando estos últimos no se consideran fiables. Estas especificaciones se obtienen a partir de la función de producción, reemplazando los *stocks* de factores o sus tasas de crecimiento por aproximaciones convenientes construidas a partir de las tasas de inversión utilizando la metodolo-

RECUADRO 2. ALGUNAS ESPECIFICACIONES EMPÍRICAS HABITUALES

A la hora de estimar, resulta generalmente conveniente trabajar con la función de producción escrita en términos de logaritmos o de tasas de crecimiento. Utilizando minúsculas para indicar logaritmos y la combinación de minúsculas con el símbolo "Δ" para indicar tasas de crecimiento, la función de producción dada por la ecuación (R.1) en el Recuadro 1 genera las siguientes dos especificaciones:

$$(R.5)\ y_{it} = a_{it} + \alpha_k k_{it} + \alpha_h h_{it} + \alpha_l l_{it} + \varepsilon_{it}$$

$$(R.6)\ \Delta y_{it} = \Delta a_{it} + \alpha_k \Delta k_{it} + \alpha_h \Delta h_{it} + \alpha_l \Delta l_{it} + \Delta\varepsilon_{it}$$

donde ε_{it} y $\Delta\varepsilon_{it}$ son perturbaciones aleatorias.

Una dificultad que surge en este punto es que estas dos ecuaciones contienen términos que no son directamente observables (en particular, el nivel de PTF, a_{it}, o su tasa de crecimiento, Δa_{it}). Para proceder con la estimación, resulta necesario hacer algún supuesto adicional sobre el comportamiento de estos términos. Diferentes supuestos generarán especificaciones econométricas distintas. La posibilidad más sencilla consiste en suponer que la tasa de progreso técnico es constante en el tiempo y en el espacio, de forma que $\Delta a_{it} = g$ para todo i y t. En este caso, g puede estimarse como la constante de la regresión en la ecuación (R.6) y a_{it} se reemplaza en la ecuación (R.5) por $a_{io} + gt$, donde a_{io} y g dan lugar a un conjunto de constantes específicas para cada país y a una tendencia común, respectivamente. Un enfoque alternativo y más sofisticado consiste en especificar Δa_{it} como una función de otras variables en la ecuación (R.6). Una posible especificación es la función de progreso técnico descrita en las ecuaciones (R.3) y (R.4) del Recuadro 1.

Cuando no existen datos de *stocks* de factores o de sus tasas de crecimiento (o cuando la información existente no se considera fiable), las tasas observadas de inversión pueden utilizarse para construir aproximaciones a las variables que entran en las ecuaciones (R.5) y (R.6). Típicamente, estas aproximaciones se construyen utilizando un modelo generalizado de Solow siguiendo el procedimiento propuesto por Mankiw, Romer y Weil (1992). En este modelo, los valores de equilibrio a largo plazo de las intensidades factoriales son funciones sencillas de las tasas de inversión y el comportamiento de estos ratios fuera del estado estacionario puede aproximarse en función de las tasas de inversión y del nivel inicial de *output* por trabajador. Si estamos dispuestos a suponer que la mayor parte de los países están razonablemente cerca de sus equilibrios de largo plazo, la ecuación (R.5) puede substituirse por otra que relaciona el *output* por trabajador con las tasas de inversión en capital físico y humano. En caso contrario, la variable a explicar será la tasa de crecimiento del *output* y la ecuación también incluirá el *output* por trabajador en el período inicial para recoger la dinámica de transición a lo largo de la trayectoria de ajuste hacia el estado estacionario. Dos especificaciones estándar de las *ecuaciones* resultantes *de estado estacionario* o de *convergencia* (que no permiten la existencia de efectos de tasa) serían las siguientes:

$$(R.7)\ q_{it} = a_{io} + gt + \frac{\alpha_k}{1-\alpha_k-\alpha_h}\ ln\frac{s_{kit}}{\delta+g+n_{it}} + \frac{\alpha_h}{1-\alpha_k-\alpha_h}\ ln\frac{s_{hit}}{\delta+g+n_{it}}$$

y

$$(R.8)\ \Delta q_{it} = g + \beta(a_{io} + gt) + \beta\left(\frac{\alpha_k}{1-\alpha_k-\alpha_h}\ ln\frac{s_{kit}}{\delta+g+n_{it}} + \frac{\alpha_h}{1-\alpha_k-\alpha_h}\ ln\frac{ln\ s_{hit}}{\delta+g+n_{it}}\right) - \beta\, q_{it}$$

donde q es el logaritmo del *output* por trabajador, s_k y s_h denotan las tasas de inversión en capital físico y humano, medidas como fracción del PIB, n la tasa de crecimiento del empleo o de la población y δ la tasa de depreciación (que se supone igual para ambos tipos de capital). El parámetro β mide la velocidad de convergencia hacia el equilibrio a largo plazo y resulta ser una función del grado de rendimientos a escala en ambos tipos de capital considerados conjuntamente y de la longitud del período que media entre dos observaciones.

gía desarrollada por Mankiw, Romer y Weil (1992) dentro del marco de un modelo generalizado de Solow con varios tipos de capital.

UN REPASO RÁPIDO DE LA EVIDENCIA EMPÍRICA

Numerosos estudios empíricos han analizado la relación existente entre la inversión en capital humano y el crecimiento económico utilizando las distintas especificaciones esbozadas más arriba[2]. Los primeros intentos en esta dirección produjeron resultados positivos que tendían a confirmar la visión tradicionalmente optimista que los economistas han tenido sobre los rendimientos macroeconómicos de la inversión en educación. Landau (1983), Baumol *et al.* (1989), Barro (1991) y Mankiw, Romer and Weil (1992), entre muchos otros, encuentran que una variedad de indicadores educativos tienen el esperado efecto positivo sobre el crecimiento del *output*.

Durante la segunda mitad de los noventa, sin embargo, una nueva ronda de trabajos empíricos produjo resultados muy desalentadores sobre la contribución de la inversión educativa al crecimiento de la productividad agregada. A diferencia de los estudios anteriores, la mayor parte de estos trabajos emplearon datos de panel con frecuencia quinquenal y utilizaron técnicas de panel o especificaciones en diferencias para controlar por posibles diferencias no observadas entre países. En este marco, las variables educativas pierden con frecuencia la significatividad e incluso entran con el signo "equivocado" en ecuaciones de crecimiento (véanse por ejemplo, Benhabib y Spiegel (1994), Islam (1995), Caselli, Esquivel y Lefort (1996) y Pritchett (2001)).

Mientras que algunos investigadores han aceptado estos resultados sin demasiadas reservas, muchos otros se han mostrado muy escépticos (véase por ejemplo Barro (1997)) y han tendido a atribuir tales resultados a distintos problemas econométricos y de especificación así como a la mala calidad de los datos internacionales de educación. La existencia de errores de medición, en particular, ha sido considerada con frecuencia un problema potencialmente importante por dos razones. La primera es que es bien sabido que las series de años medios de escolarización utilizadas en la literatura contienen una gran cantidad de "ruido" y la segunda que, en cualquier caso, el número de años de escolarización sólo puede ser considerado un indicador muy imperfecto de las competencias y conocimientos de la fuerza laboral. El primero de estos problemas, además, es especialmente grave cuando se utilizan técnicas estadísticas en las que la estimación se basa en la varia-

2. Para una panorámica más detallada de la literatura relevante, véase la sección 3 del Anexo a de la Fuente y Ciccone (2003).

ción temporal de los datos porque los errores de medición generados por cambios en los criterios de clasificación y de recogida de datos generan una fuerte volatilidad espuria en las series de escolarización que hace difícil identificar su contribución al crecimiento de la productividad.

Aunque es demasiado pronto para considerar zanjado el asunto, mi lectura de la evidencia acumulada sobre el tema es optimista. Tenemos buenas razones para pensar que los resultados negativos de la literatura precedente pueden efectivamente atribuirse a las deficiencias de los datos educativos utilizados en estos estudios. Los trabajos que utilizan nuevas series "mejoradas" de escolarización y aquellos que emplean técnicas estadísticas para corregir los sesgos derivados de los errores de medición sugieren que aumentos en el nivel educativo tienen un efecto sustancial sobre el crecimiento de la productividad. Los resultados son aún más claros cuando se utilizan indicadores directos de los niveles de conocimientos y competencias de la población como *proxies* para el *stock* de capital humano, lo que sugiere que mejoras en la calidad de la educación pueden tener un efecto aún mayor sobre el producto agregado que aumentos en su cantidad.

La oleada de resultados negativos sobre la aportación de la educación al crecimiento que encontramos en la segunda mitad de los noventa está claramente asociada con la introducción de técnicas de panel. Mientras que en estudios anteriores sobre el tema la estimación se realizaba con datos de corte transversal (esto es, con una única observación por país que describía su comportamiento medio durante un período de varias décadas), en este grupo de estudios se utilizan varias observaciones por país tomadas sobre períodos más cortos y se emplean especificaciones en diferencias o técnicas de panel que básicamente eliminan la variación de corte transversal que existe en los datos antes de proceder a la estimación.

Aunque estas técnicas de estimación tienen la importante ventaja de que permiten controlar por diferencias no observables entre países, también presentan algunos inconvenientes, entre los que cabe destacar su mayor sensibilidad a la presencia de errores de medición en las variables. El problema surge, en particular, porque tales errores son generalmente mayores en la dimensión temporal que en la transversal pues tienden a cancelarse cuando se promedia sobre períodos largos. Esto sugiere, como ya he observado, que una posible explicación de los resultados adversos obtenidos en estudios de panel tiene que ver con la mala calidad de los datos de capital humano que se han utilizado en buena parte de la literatura de crecimiento. Como veremos enseguida, la mayor parte de las series de escolarización que se utilizaron en los primeros estudios sobre el tema contenían grandes cantidades de ruido que tienen su origen en diversas inconsistencias de los datos primarios utilizados para construirlas. La existencia de este ruido genera un sesgo a la

baja en la estimación de los coeficientes que miden el impacto del capital humano (esto es, una tendencia a subestimar sus valores) porque introduce volatilidad espuria en el *stock* de capital humano que, como cabría esperar, no se ve acompañada por cambios proporcionales en el nivel de productividad.

Algunos estudios ofrecen evidencia consistente con esta hipótesis. Comenzando con Krueger and Lindhal (K&L, 2001), algunos autores han construido indicadores estadísticos del contenido informativo de distintas series de escolarización (*ratios de fiabilidad*) que pueden utilizarse para calcular el tamaño esperado del sesgo de atenuación. Su conclusión es que el valor de este ratio es lo suficientemente bajo como para explicar la falta de significatividad de las variables educativas en estudios previos. Otros autores, incluyendo a Cohen y Soto (2007), de la Fuente y Doménech (D&D, 2001a y b, 2006 y 2014a), Lutz *et al.* (2007) y Barro y Lee (2001, 2013), han intentado mejorar la ratio señal/ruido de las series de escolarización mediante la explotación de nuevas fuentes de información y la introducción de diferentes correcciones. Todos ellos encuentran que los resultados sobre la contribución de la educación al crecimiento económico mejoran considerablemente cuando se utilizan las series revisadas. Volveré sobre estas cuestiones con mucho mayor detalle en las dos próximas secciones.

Otra línea de trabajo interesante es la que ha hecho uso de indicadores directos de conocimientos y destrezas que podrían ser mejores medidas del *stock* de capital humano que los años medios de escolarización. Aunque tales datos son todavía escasos, algunos estudios recientes sugieren que ésta puede ser una línea de investigación muy productiva. Hanushek y varios coautores[3] construyen indicadores de la calidad de la fuerza laboral utilizando los resultados medios en pruebas estandarizadas internacionales de desempeño en matemáticas, ciencia y lectura, mientras que Coulombe *et al.* (2004) utilizan datos tomados de IALS, un estudio internacional sobre el nivel de destrezas de la población adulta que ha sido realizado por la OCDE y Statistics Canada. En ambos casos, los resultados de regresiones de crecimiento apuntan a efectos sobre el *output* mayores incluso que los que se obtienen con los datos revisados de escolarización. Aunque no están del todo libres de problemas, estas estimaciones sugieren que la calidad de la educación es al menos tan importante como su cantidad y que la rentabilidad de cualquier política que mejore el rendimiento académico de los estudiantes es seguramente enorme, pues no sólo serán elevados sus beneficios sino que además los costes relevantes serán mucho más bajos que los de aumentar la estancia media de los jóvenes en el sistema educativo, lo tiene un elevado coste de oportunidad ligado al sacrificio de tiempo de trabajo.

3. Véanse entre otros Hanushek y Kimko (2000) y Hanushek y Woessmann (2008, 2010 y 2012).

BASES DE DATOS INTERNACIONALES DE ESCOLARIZACIÓN: PROBLEMAS E IMPLICACIONES

La mayor parte de los gobiernos recogen información sobre distintos indicadores educativos a través de los censos de población, las encuestas de población activa y otras encuestas, registros y estudios especializados. Diversas organizaciones internacionales recopilan estos datos y publican estadísticas comparativas que proporcionan información fácilmente accesible y (en principio) homogénea para un gran número de países. La fuente regular más importante de estadísticas educativas internacionales es el *Anuario Estadístico* de la UNESCO. Esta publicación ofrece series temporales razonablemente completas de tasas de escolarización para distintos niveles educativos en la mayor parte de los países del mundo y contiene también algunos datos sobre el nivel medio de instrucción de la población adulta, el gasto público en educación, el número de estudiantes por profesor y otras variables de interés[4].

Las series de tasas de escolarización que publica la UNESCO se han utilizado en un gran número de estudios empíricos sobre la relación entre educación y crecimiento económico. En muchos casos, esta elección se debe más a la facilidad de acceso a estos datos y a su amplia cobertura internacional que a su idoneidad teórica para el problema de interés. Las tasas de escolarización son probablemente un indicador aceptable, aunque imperfecto, del flujo de inversión educativa, pero no son necesariamente una buena medida del *stock* existente de capital humano puesto que el nivel medio de formación de la población adulta (que es en principio la variable de mayor interés desde un punto de vista teórico) responde a los flujos de inversión sólo de una manera gradual y con un retardo considerable.

En un intento de evitar estos problemas, distintos investigadores han construido bases de datos que intentan medir directamente el *stock* educativo de la población o de la fuerza laboral de un número elevado de países durante un período de varias décadas. Estas bases de datos generalmente combinan la información censal existente sobre el nivel de instrucción de la población con las tasas de escolarización de la UNESCO para aproximar la evolución de la composición de la población por nivel de estudios y construir series de años medios de instrucción. Los más conocidos de los intentos tempranos en esta línea son el trabajo de Kyriacou (1991), las primeras versiones de la base de datos de Barro y Lee (1993, 1996 y 2000) y las series construidas por diversos investigadores del Banco Mundial

4. Otras fuentes útiles son el *Anuario Demográfico* de la ONU, que ofrece datos de formación desagregados por grupos de edad y el informe anual de la OCDE sobre la educación en sus países miembros (*Education at a Glance*), que incluye gran cantidad de información sobre los *inputs* y los *outputs* del sistema educativo.

(Lau, Jamison y Louat (1991), Lau, Bhalla y Louat (1991) y Nehru, Swanson y Dubey (NSD, 1995).

En de la Fuente y Doménech (D&D, 2006) se repasa brevemente la metodología utilizada en estos estudios y se comparan las distintas series entre sí, con especial atención al caso de la OCDE, donde la calidad de la información de base disponible debería ser en principio mejor que en los países en vías de desarrollo. El análisis de las distintas series revela discrepancias muy significativas entre ellas en términos de las posiciones relativas de muchos países así como niveles o perfiles temporales poco plausibles para algunos de ellos. Aunque los distintos estudios generalmente coinciden cualitativamente a la hora de hacer comparaciones entre grandes regiones (p. ej. la OCDE versus países en vías de desarrollo en distintas áreas geográficas), las discrepancias son muy importantes cuando nos centramos en el grupo de países industriales, que es el que aquí nos concierne más directamente. Otro motivo para la preocupación es que las bases de datos existentes a menudo presentan saltos extremadamente grandes en niveles medios de formación sobre períodos muy cortos, especialmente en los niveles secundario y terciario.

En buena parte, estos problemas son consecuencia de las deficiencias de los datos primarios de base. Como observan Behraman y Rosenzweig (1994), existen buenas razones para dudar de la fiabilidad y consistencia de la información que proporciona la UNESCO tanto sobre tasas de escolarización como sobre niveles medios educativos. Nuestro análisis de las diferentes bases de datos educativas confirma este diagnóstico y sugiere que la mayoría de los problemas detectados en esas series tienen su origen en los datos primarios, que no parecen ser consistentes ni en el tiempo ni en el espacio, en su tratamiento de la formación profesional y técnica y en otros tipos de estudios, y reflejan unas veces el número de individuos que han comenzado un cierto nivel educativo y otras el número que lo ha completado.

EL SESGO DE ATENUACIÓN Y UN INDICADOR DE LA CALIDAD DE LOS DATOS

La mala calidad de los datos internacionales de escolarización es un problema serio porque tiende a oscurecer la relación existente entre las variables de interés y genera una tendencia a subestimar el impacto del capital humano sobre la productividad. Para entender el origen del *sesgo de atenuación* generado por la existencia de errores de medición, supongamos que el nivel de productividad, Q, es una función lineal del *stock* de capital humano, H, de forma que

$$(1)\ Q = bH + u$$

donde u es una perturbación aleatoria. Dada esta relación, variaciones en el *stock* de capital humano, H, inducirán cambios en el nivel de productividad, Q, y el examen de la magnitud relativa de las variaciones de ambas variables nos permitirá estimar el valor del coeficiente b. Ahora bien, si H está medido con error, de forma que lo que observamos en realidad no es H sino una *proxy* ruidosa

$$(2)\ P = H + \varepsilon,$$

donde ε es un error de medición aleatorio, entonces parte de la aparente variación del *stock* de capital humano (en el tiempo o entre países) se deberá al error de medición (esto es, será ruido en vez de auténtica señal). Puesto que tales variaciones lógicamente no inducen respuesta alguna en Q, esta variable parecerá menos sensible al *stock* de capital humano de lo que es en realidad, lo que sesgará a la baja el valor estimado de b.

En resumen, el sesgo de atenuación surge porque los errores de medición generan "ruido" que tiende a oscurecer la relación entre las variables de interés. Se puede demostrar que el tamaño del sesgo estará inversamente relacionado con el contenido informativo de la serie, medido por su *ratio de fiabilidad*, r. Este indicador se define como el cociente entre la señal y la suma de señal y ruido en los datos, esto es, como

$$(3)\ r \equiv \frac{var\,H}{var\,P} = \frac{var\,H}{var\,H + var\,\varepsilon}$$

donde *var* H mide la *señal* contenida en la serie (esto es, la variación real del *stock* de capital humano) y *var* ε el *ruido* que la distorsiona[5]. Este ratio resulta de gran utilidad, en primer lugar, porque nos proporciona un indicador del contenido informativo de cada una de las series, y en segundo porque el error en la estimación será inversamente proporcional a su valor. La ratio de fiabilidad, por tanto, puede utilizarse para corregir el sesgo de atenuación, obteniendo así estimaciones consistentes (no sesgadas en muestras grandes) del parámetro de interés[6].

Puesto que H y ε no se pueden observar separadamente, los ratios de fiabilidad no se pueden calcular directamente. Si pueden, sin embargo, estimarse utilizando la metodología desarrollada por Krueger y Lindhal (2001) siempre que dispongamos de varias *proxies* ruidosas para la misma variable de interés. El Recuadro 3 describe esta metodología y la extensión de la misma propuesta por de la Fuente y Doménech (2006).

5. Obsérvese que el denominador del último término de (3) supone implícitamente que el error de medición, ε, no está correlacionado con H.

RECUADRO 3. LA ESTIMACIÓN DE LAS RATIOS DE FIABILIDAD

Sean $P_1 = H + \varepsilon_1$ y $P_2 = H + \varepsilon_2$ dos *proxies* ruidosas del *stock* de capital humano, H. Es fácil comprobar que, si los términos de error de las dos series, ε_1 y ε_2, no están correlacionados entre sí, la covarianza entre P_1 y P_2 puede utilizarse para estimar la varianza de H, que es la única magnitud desconocida en la expresión (3) del texto donde se define la ratio de fiabilidad. Por lo tanto, bajo este supuesto, $\hat{r}_1$ puede estimarse como

$$(R.9)\ \hat{r}1 = \frac{\text{cov } (P1, P2)}{\text{var } P1}$$

que resulta ser la fórmula del estimador de mínimos cuadrados ordinarios del coeficiente de pendiente de una regresión de P_2 sobre P_1. Por tanto, para estimar la fiabilidad de P_1 estimamos una regresión de la forma $P_2 = c + r_1 P_1$[6]. Obsérvese, sin embargo, que si los errores de medición de las dos series están positivamente correlacionados ($E\varepsilon_1\varepsilon_2 > 0$) como cabría esperar en muchos casos, $\hat{r}_1$ sobreestimará la ratio de fiabilidad, lo que hará que tendamos a infravalorar el sesgo generado por el error de medición.

En de la Fuente y Doménech (2006) se desarrolla una extensión de este procedimiento que permite construir un estimador de mínima varianza de la ratio de fiabilidad cuando existen más de dos estimaciones con error de la misma variable subyacente, manteniendo la hipótesis de que los errores de medición de las distintas series no están correlacionados entre sí. Al igual que en K&L, la ratio de fiabilidad r_k de una serie dada de años de escolarización (digamos S_k) se obtiene utilizando a S_k para tratar de explicar otras estimaciones o *proxies* alternativas de la misma variable subyacente (S_j con $j \neq k$). La principal diferencia con el método original de K&L estriba en que, en vez de estimarse una serie de regresiones independientes con las distintas series de escolarización tomadas de dos en dos, el estimador eficiente de la ratio de fiabilidad para la serie k puede obtenerse como el coeficiente de pendiente de un modelo SUR restringido de la forma

$$(R.10)\ P_k = c_k + r_j P_j + u_k \quad \text{for } k = 1..., K$$

donde restringimos r_j a tener el mismo valor para todas las series de referencia, k, utilizadas en el lado izquierdo del sistema (R.10), donde k varía sobre la última versión disponible de todas las bases de datos distintas de j.

La fiabilidad de la base de datos de Barro y Lee (2000), por ejemplo, se estima utilizando esta serie como variable explicativa en un conjunto de regresiones SUR en el que la variable de referencia (la variable a explicar o dependiente) son los años medios de escolarización estimados por Kyriacou (1991), NSD (1995), D&D (2006) y Cohen y Soto (2001). Las demás versiones de la base de datos de Barro y Lee, sin embargo, no se utilizan como referencia porque la correlación de los errores de medición dentro de una familia determinada de estimaciones es casi con seguridad positiva y muy elevada, lo que tendería a inflar artificialmente la estimación de la tasa de fiabilidad.

6. Intuitivamente, la regresión de P_2 sobre P_1 nos da una idea de lo bien que P_1 explica la variable sin errores, H. Esto es así porque el error de medición en la variable dependiente (P_2 en este caso) será absorbido por la perturbación de la regresión sin que esto genere un sesgo (al contrario de lo que sucede cuando lo que se mide con error es la variable independiente). Por tanto, es *casi como si* estuviésemos haciendo una regresión del verdadero H sobre P_1 .

ALGUNOS RESULTADOS PARA LOS PAÍSES DE LA OCDE Y PARA LAS REGIONES ESPAÑOLAS

La discusión precedente sugiere dos formas complementarias de enfrentarnos al problema causado por la mala calidad de los datos educativos. Uno consiste en tratar de mejorar la calidad de los datos mediante la utilización de mejores fuentes primarias y la introducción de correcciones que neutralicen los efectos de cambios en criterios de clasificación, y el otro en utilizar ratios de fiabilidad para corregir el sesgo de atenuación. En una serie de trabajos relacionados entre sí, Rafael Doménech y yo (D&D, 2000, 2001a and b, 2002, 2006, 2008 y 2014 a y b) hemos utilizado ambas estrategias utilizando datos para una muestra de 21 países de la OCDE y de las regiones españolas. En ambos casos, el primer paso ha consistido en construir una nueva serie de escolarización que intenta aumentar la ratio entre señal y ruido. En el caso de España (D&D, 2008 y 2014b), la tarea ha sido relativamente simple porque la información primaria necesaria está fácilmente disponible en los censos de población, que han sido elaborados utilizando criterios de clasificación relativamente claros y estables en el tiempo.

UNA NUEVA BASE DE DATOS

Las series de escolarización para la OCDE (D&D, 2000, 2001b 2006 y 2014a) exigieron bastante más trabajo. La primera tarea consistió en recopilar toda la información que pudimos encontrar sobre la distribución por nivel de estudios de la población adulta en los países de la OCDE. Con este fin, hemos utilizado tanto publicaciones de organizaciones internacionales como fuentes nacionales (censos y encuestas, anuarios estadísticos y datos no publicados proporcionados por gobiernos nacionales y por la OCDE). Seguidamente hemos tratado de reconstruir un perfil razonable de la evolución del nivel medio de formación en cada país empleando toda la información disponible y un poco de sentido común. Para aquellos países en los que se dispone de series razonablemente completas, nos hemos basado fundamentalmente en fuentes nacionales. Para el resto, partimos de la estimación aparentemente más fiable de las existentes entre 1990 y 1995 (que generalmente se toma de la OCDE) y procedemos hacia atrás intentando evitar saltos poco razonables en las series que sólo pueden tener su origen en cambios en los criterios de clasificación. La construcción de las series requirió en algunos casos juicios subjetivos de valor para elegir entre estimaciones alternativas cuando las distintas fuentes están en desacuerdo. También hemos reinterpretado algunos datos tomados de compilaciones internacionales como referidos a niveles educativos ligeramente distintos de los indicados en las propias fuentes.

Las observaciones no disponibles situadas entre datos censales se estiman por simple interpolación lineal. En el resto de los casos se recurre a extrapolaciones que utilizan, siempre que es posible, información sobre el nivel educativo de la población desagregada por grupos de edad en años cercanos al de interés.

INDICADORES DE CALIDAD PARA DISTINTAS SERIES DE ESCOLARIZACIÓN

En de la Fuente y Doménech (D&D, 2002, 2006 y 2014c) se utiliza la metodología descrita en el Recuadro 3 para estimar las ratios de fiabilidad de las series de años medios de escolarización más utilizadas en la literatura, restringiéndonos a la muestra de 21 países cubierta por las series descritas en el apartado anterior. Este indicador se construye para cada una de las distintas transformaciones de la variable de años medios de escolarización que se utilizan en las especificaciones más habituales de la función de producción. En particular, se estiman ratios de fiabilidad para los años medios de formación medidos en niveles (S_{it}) y en logaritmos (s_{it}), para los incrementos medios anuales tanto en niveles como en logaritmos medidos entre observaciones quinquenales sucesivas (ΔS_{it} y Δs_{it}) y para el nivel de escolarización medido en desviaciones sobre las medias nacionales ($s_{it} - s_i$). Obsérvese que Δs_{it} corresponde a la tasa anual de crecimiento y que $s_{it} - s_i$ es la transformación intra-grupos frecuentemente utilizada para eliminar efectos fijos.

Los resultados del ejercicio se recogen en el primer panel del Cuadro 1 con las distintas bases de datos ordenadas de acuerdo con el promedio de sus ratios de fiabilidad. El segundo panel muestra resultados actualizados de un ejercicio similar para series más recientes y más largas, incluyendo las últimas actualizaciones disponibles de las bases de datos de B&L, C&S y D&D. En ambos paneles, la última fila del cuadro muestra el valor medio de la ratio de fiabilidad para cada transformación de las series (calculada sobre las distintas bases de datos) y la última columna recoge la ratio promedio de fiabilidad de cada serie (calculada sobre las distintas transformaciones).

Nuestros resultados muestran que la importancia de los errores de medición varía significativamente de unas bases de datos a otras, aunque el *ranking* exacto de cada una de ellas depende de la transformación de los datos que se utilice. Dos de las series más usadas en la literatura temprana de crecimiento, las construidas por Kyriacou (1991) y Barro y Lee (varios años), tienen ratios de fiabilidad respetables cuando se utilizan en niveles pero, tal como observan Krueger y Lindhal (2001), contienen muy poca información cuando se diferencian o se introducen efectos fijos.

CUADRO 1. ESTIMACIONES *SUR* DE LA RATIO DE FIABILIDAD, MUESTRA DE 21 PAÍSES DE LA OCDE

A. RESULTADOS DE D&D (2006) PARA LAS PRIMERAS SERIES DE ESCOLARIZACIÓN

	S_{it}	s_{it}	ΔS_{it}	Δs_{it}	s_{it}-s_i	*promedio*
D&D (2002)	0,754	0,775	0,337	0,769	0,917	0,710
C&S (2001)	0,806	0,912	0,330	0,467	0,547	0,612
D&D (2000)	0,720	0,761	0,100	0,550	0,818	0,590
Kyr. (1991)	0,723	0,600	0,024	0,065	0,111	0,305
B&L (2000)	0,707	0,603	-0,018	0,045	0,178	0,303
B&L (1996)	0,559	0,516	-0,017	0,039	0,146	0,249
B&L (1993)	0,526	0,436	-0,019	0,029	0,121	0,219
NSD (1995)	0,278	0,330	-0,021	0,066	0,095	0,150
promedio	0,634	0,617	0,090	0,254	0,367	0,392

B. RESULTADOS DE D&D (2014c) PARA SERIES RECIENTES DE ESCOLARIZACIÓN

	S_{it}	s_{it}	ΔS_{it}	Δs_{it}	s_{it}-s_i	*promedio*
D&D (2014)	0,994	1,072	0,462	0,850	0,822	0,840
C&S (2007)	0,886	0,878	0,599	0,704	1,065	0,826
L&S+	0,534	0,606	0,288	0,431	0,444	0,461
B&L (2013)	0,628	0,604	0,025	0,084	0,226	0,313
promedio	0,760	0,790	0,344	0,517	0,639	0,610

Notas:
- La estimación se realiza con todas las series expresadas en desviaciones sobre su promedio muestral en cada período con el fin de eliminar posibles efectos fijos temporales.
- *Clave:* D&D = de la Fuente y Doménech; C&S = Cohen y Soto; Kyr = Kyriacou; B&L = Barro y Lee; NSD = Nehru *et al.*; L&S+ = Lutz, Samir.
- Para L&S+ y C&S (2007) trabajamos con versiones actualizadas de los datos suministradas en 2012 por K. C. Samir y por M. Soto. La última está disponible en http://soto.iae-csic.org/Data.htm

Fuente: de la Fuente y Doménech (2006 y 2014c).

En el trabajo original (D&D, 2006) nuestra estimación de la ratio de fiabilidad promedio para todas las series y transformaciones era de 0,392. Puesto que el intervalo de posibles valores de esta variable está entre cero y uno (con el primer valor indicando que la serie contiene sólo ruido y el segundo que no hay errores de medición)[7], este resultado sugiere que la estimación típica del coeficiente de capital

7. Esto es cierto siempre y cuando los errores de medición de las distintas series no estén correlacionados entre sí o con H. Como se observa en el Cuadro 1, algunas de nuestras estimaciones de las ratios de fiabilidad están fuera de este intervalo, lo que implica que alguna de estas hipótesis no se cumple. En D&D (2002) construimos

humano en una ecuación de crecimiento estimada con las primeras series disponibles contenía un importante sesgo a la baja, incluso sin tener en cuenta la pérdida adicional de señal que se produce cuando la ecuación incluye otros regresores además del de capital humano (véase D&D, 2002).

Los intentos de aumentar el contenido informativo de las series de escolarización parecen haber tenido al menos un éxito parcial. Las últimas revisiones de las series educativas han aumentado en torno a un 50% la ratio de fiabilidad media de las series disponibles, hasta un valor de 0,610. Aún así, dependiendo de la serie y de la especificación utilizada, el sesgo a la baja en la estimación del coeficiente del capital humano en la función de producción podría ser todavía muy importante.

CALIDAD DE LAS SERIES EDUCATIVAS Y ESTIMACIONES DEL IMPACTO DEL CAPITAL HUMANO EN LA OCDE

Como hemos visto en el capítulo anterior, el tamaño esperado del sesgo de atenuación es una función decreciente de la ratio de fiabilidad de la serie empleada en la estimación. Esto sugiere que el valor estimado del coeficiente del capital humano en una regresión de crecimiento debería aumentar con la calidad de los datos educativos. En de la Fuente y Doménech (D&D, 2002 y 2006) se muestra que éste es efectivamente el caso. Para comprobarlo, estimamos varias especificaciones de una función de producción agregada utilizando las distintas series de escolarización analizadas en el primer panel del Cuadro 1 como indicadores alternativos del *stock* de capital humano y comprobamos que tanto el tamaño como la significatividad del coeficiente de esta variable aumentan de la forma esperada con la ratio de fiabilidad. Seguidamente, explotamos la correlación observada entre estas dos variables para construir diversas "meta-estimaciones" del parámetro de interés que corrigen el sesgo generado por el error de medición[8].

Resultados con distintas series educativas

Las ecuaciones que estimamos se derivan de una función de producción agregada del tipo Cobb-Douglas con rendimientos constantes a escala que incluye, además del *stock* de capital físico y el empleo, el nivel medio de formación de la mano de obra como uno de sus *inputs*. Esta ecuación se estima en niveles (con las variables expresadas en logaritmos), en niveles con efectos fijos de país y en diferencias. En

estimaciones alternativas de los ratios de fiabilidad bajo condiciones más generales y encontramos que las correcciones necesarias no cambian cualitativamente los resultados.

8. Una meta-estimación es una estimación que no se obtiene directamente de los datos sino que se construye a partir de otras estimaciones primarias.

D&D (2002) también se estima una cuarta especificación en diferencias en la que se introducen efectos fijos y un proceso de difusión o *catch-up* tecnológico que permite que, siendo otras cosas iguales, la tasa de crecimiento de la productividad total de los factores (PTF) sea directamente proporcional a la distancia tecnológica entre cada país y los Estados Unidos. En este caso, los efectos fijos de país capturan diferencias permanentes en niveles relativos de PTF que presumiblemente reflejarán diferencias en niveles de inversión en I+D y en otras variables omitidas[9].

Las distintas especificaciones se estiman con datos quinquenales para el período 1960-1990 con la muestra habitual de países de la OCDE, incluyendo en todos los casos efectos fijos temporales (*dummies* para los distintos subperíodos muestrales). Los valores del coeficiente que mide la elasticidad del *output* con respecto al nivel medio de formación obtenidos con las distintas especificaciones y bases de datos se muestran en el Cuadro 2. La última columna del cuadro muestra los valores medios del coeficiente de interés y de su estadístico *t* para cada especificación y las dos últimas filas los promedios de las mismas variables para cada base de datos.

CUADRO 2. ESTIMACIONES ALTERNATIVAS DEL COEFICIENTE DEL CAPITAL HUMANO (B) EN DISTINTAS ESPECIFICACIONES DE LA FUNCIÓN DE PRODUCCIÓN CON DIFERENTES SERIES EDUCATIVAS

	NSD	*KYR*	*B&L93*	*B&L96*	*B&L00*	*C&S*	*D&D00*	*D&D02*	*MEDIA*
niveles	0,078	0,186	0,141	0,165	0,238	0,397	0,407	0,378	0,249
	(2,02)	(2,18)	(4,49)	(4,82)	(6,19)	(7,98)	(7,76)	(6,92)	(5,30)
efectos fijos	0,068	0,066	0,136	0,115	0,203	0,608	0,627	0,958	0,348
	(0,76)	(1,86)	(3,30)	(1,80)	(3,74)	(4,49)	(3,99)	(6,51)	(3,31)
diferencias	0,079	0,009	0,089	0,083	0,079	0,525	0,520	0,744	0,266
	(0,70)	(0,15)	(2,52)	(1,47)	(1,28)	(2,57)	(2,17)	(3,10)	(1,75)
catch-up	- 0,206	0,014	0,056	-0,007	-0,019	0,573	0,587	0,540	0,192
	(1,61)	(0,29)	(1,80)	(0,11)	(0,31)	(3,52)	(3,47)	(2,89)	(1,24)
promedio	0,005	0,069	0,106	0,089	0,125	0,526	0,535	0,655	
	(0,47)	(1,12)	(3,03)	(2,00)	(2,73)	(4,64)	(4,35)	(4,86)	

Clave: Véanse las notas al Cuadro 1.

9. Todas las especificaciones se obtienen a partir de la ecuación (R.2) del Recuadro 1, utilizando los años medios de formación (*S*) como *proxy* para el *stock* de capital humano (*H*). En la última de estas especificaciones, se incorpora también una función de progreso técnico similar a la ecuación (R.4) del mismo recuadro, excepto en que se omite el *stock* de capital humano. El modelo estimado, por tanto, no permite la existencia de efectos de tasa. Hemos intentado estimarlo incorporando estos efectos pero los resultados no son satisfactorios. Este problema es bastante habitual en la literatura. Véase de la Fuente y Ciccone (2003) sobre las razones por las que puede resultar difícil separar los efectos de tasa de los de nivel.

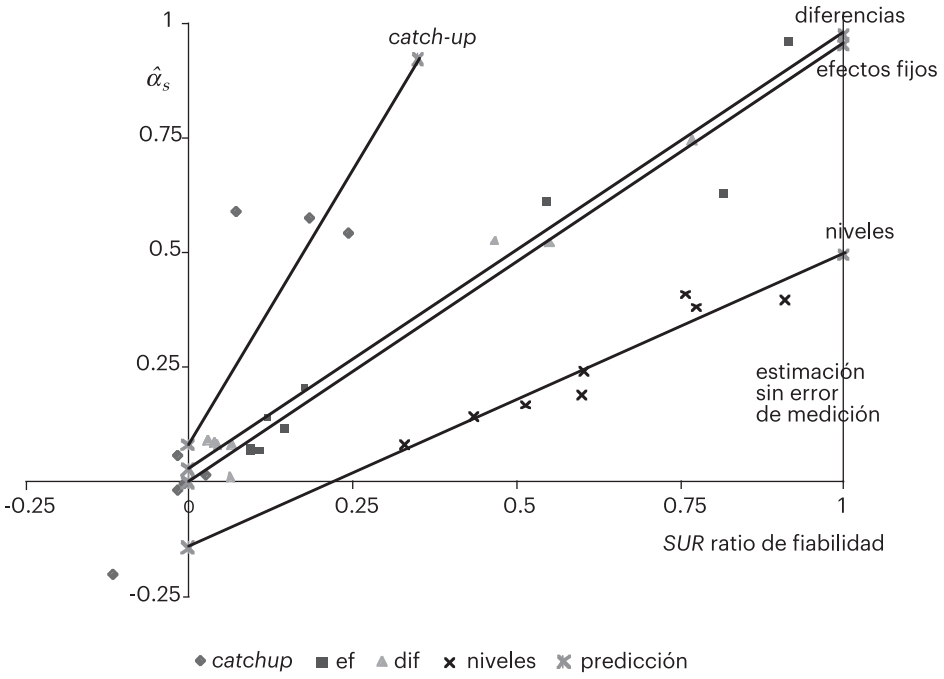
El patrón de resultados que emerge según vamos cambiando la fuente de los datos de escolarización es consistente con nuestra hipótesis sobre la importancia del sesgo derivado del error de medición. Para todas las bases de datos, el valor estimado del coeficiente de la variable educativa es positivo y significativo para la especificación en niveles sin efectos fijos de país (primera fila del cuadro) pero el tamaño del coeficiente estimado y su nivel de significatividad aumentan apreciablemente cuando utilizamos las series con mayores ratios de fiabilidad (que son las que corresponden a las últimas columnas del Cuadro 2). Las diferencias son aún más claras cuando la estimación se realiza con efectos fijos de país (segunda fila) o con los datos en tasas de crecimiento (tercera y cuarta fila). Los resultados obtenidos con las series de Kyriacou, Barrro y Lee y NSD expresadas en tasas de crecimiento son consistentes con los obtenidos por Kyriacou (1991), Benhabib y Spiegel (1994) y Pritchett (2001), quienes encuentran que el coeficiente del capital humano no es significativo (y a menudo tiene signo negativo) en funciones de producción estimadas en diferencias. Por otra parte, tanto nuestras series como las de Cohen y Soto generan valores positivos, relativamente grandes y precisos del coeficiente del capital humano en la mayor parte de las estimaciones.

Cómo corregir el sesgo derivado de los errores de medición

Los resultados resumidos en el Cuadro 2 indican con claridad que el error de medición genera un sesgo importante a la baja en la estimación del coeficiente del capital humano en la función de producción y muestran que la mejora en la calidad de los datos reduce este sesgo y conduce a resultados generalmente más favorables a la hipótesis de que la inversión educativa contribuye de forma sustancial al crecimiento de la productividad. El Gráfico 1 ilustra y refuerza estas conclusiones. En él he dibujado las distintas estimaciones del coeficiente de interés que aparecen en el Cuadro 2 como función de los ratios de fiabilidad (r) correspondientes (tomados del primerpanel del Cuadro 1), así como las rectas de regresión ajustadas para cada una de las especificaciones econométricas que hemos estimado en el apartado anterior. El gráfico muestra que existe una clara correlación positiva entre estas dos variables para cada especificación y sugiere que el valor correcto del coeficiente de capital humano es al menos 0,50 (que es la predicción para la ecuación en niveles cuando $r = 1$, esto es, cuando el capital humano se mide sin error).

Como sugiere el Gráfico 1, es posible extrapolar la relación observada en las distintas bases de datos entre la ratio de fiabilidad y el coeficiente estimado del capital humano con el fin de estimar el valor de este último que se obtendría en ausencia de errores de medición. De esta forma, se pueden construir meta-estimaciones de este parámetro que estarán libres del sesgo de atenuación, si bien esto ha

GRÁFICO 1. COEFICIENTE ESTIMADO DEL CAPITAL HUMANO *VS*. RATIO DE FIABILIDAD *SUR*



de hacerse con algo más de cuidado de lo que sugiere el gráfico (al menos cuando la ecuación de crecimiento incluye varios regresores).

En de la Fuente y Doménech (2002 y 2006) utilizamos un procedimiento de este tipo para obtener una serie de meta-estimaciones consistentes del parámetro de interés. Trabajando con las tres especificaciones lineales estimadas arriba (esto es, con todas ellas excepto el modelo de *catch-up* tecnológico) y con tres supuestos diferentes sobre la naturaleza de los errores de medición (y en particular sobre su posible correlación entre bases de datos de escolarización y con el resto de las variables explicativas del modelo de crecimiento), obtenemos diferentes estimaciones del valor del coeficiente del capital humano. De esta forma, generamos un intervalo bastante amplio de posibles valores para este parámetro. Bajo las hipótesis que consideramos más plausibles, nuestros resultados implican valores del mismo de entre 0,70 y 0,80.

Estos valores son significativamente más elevados que los obtenidos en la literatura anterior. Merece la pena resaltar que nuestra cota inferior para este parámetro es aproximadamente dos veces mayor que la estimación de 0,33 obtenida por Mankiw, Romer y Weil (1992). Esta última cifra se habría visto a comienzos de los noventa como un valor de consenso del parámetro de capital humano, pero

en años posteriores se ha considerado generalmente como demasiado optimista como consecuencia de los resultados negativos que hemos comentado en una sección anterior. Nuestros resultados, por contra, apuntan a una cifra mucho más elevada y sugieren, por tanto, que la inversión en capital humano es un importante factor de crecimiento cuyo efecto sobre la productividad ha sido infravalorado en estudios anteriores debido a la mala calidad de los datos de escolarización.

RESULTADOS REGIONALES PARA ESPAÑA

Nuestro análisis de los datos regionales españoles produce conclusiones cualitativamente similares sobre la contribución de la educación a la productividad. En de la Fuente y Doménech (2008) estimamos un modelo de *catch-up* con datos bienales para las regiones españolas que cubren el período 1965-1995. La especificación es idéntica a la estimada más arriba con datos de la OCDE excepto en que el capital físico se desagrega ahora en dos componentes, uno de los cuales recoge el *stock* de infraestructuras productivas (infraestructuras de transporte, suministro de agua y estructuras urbanas). Como *proxy* para el *stock* de capital humano utilizamos nuestras propias estimaciones de años medios de educación construidas a partir de datos censales y una estimación alternativa de la misma variable elaborada a partir de las series de escolarización construidas por Mas *et al.* (MPUSS, 2002). Esta serie se construye con datos primarios de la EPA y, presumiblemente como consecuencia de los reducidos tamaños muestrales existentes en algunas regiones, tiende a ser mucho más volátil que nuestra variable educativa.

Las estimaciones del coeficiente del capital humano obtenidas con las distintas series de nivel educativo se recogen en el Cuadro 3. Todas las ecuaciones contienen *dummies* temporales. Las ecuaciones [1] y [2] contienen el conjunto completo de *dummies* regionales mientras que en las ecuaciones [3] y [4] se retienen sólo aquellos efectos fijos regionales que resultaron significativos en la primera iteración.

La inspección del cuadro revela dos resultados interesantes en relación con el coeficiente del capital humano. En primer lugar, este parámetro pasa de ser prácticamente cero y no-significativo con los datos de MPUSS (2002) a tener un valor positivo elevado y significativo con nuestras series educativas. Este resultado es consistente con nuestras estimaciones del contenido informativo de ambas series, dado que la ratio de fiabilidad es de 0,900 para nuestra serie y de sólo 0,035 para la de MPUSS cuando las dos se miden en diferencias logarítmicas. En segundo lugar, nuestra estimación del coeficiente del capital humano en la función de producción con datos regionales españoles (0,835) es mayor que la obtenida para la muestra de países de la OCDE con una especificación similar (0,540 con el conjunto comple-

CUADRO 3. ESTIMACIONES DEL COEFICIENTE DEL CAPITAL HUMANO EN LA FUNCIÓN DE PRODUCCIÓN OBTENIDAS CON DISTINTAS SERIES Y ESPECIFICACIONES

	[1]	[2]	[3]	[4]
FUENTE DATOS:	***MPUSS***	***D&D***	***MPUSS***	***D&D***
Cap. humano	-0,013 (0,11)	0,835 (2,04)	-0,013 (0,11)	0,835 (4,13)
R^2 ajust.	0,749	0,753	0,757	0,763
Dummies region	todas	todas	signif.	signif.

Notas:
– Todas las ecuaciones incluyen *dummies* temporales.
– Estadístico *t* de White consistente a la heteroscedasticidad en paréntesis debajo de cada coeficiente.

to de *dummies* regionales y 0,394 cuando sólo se conservan las que resultan significativas). Una vez más, la explicación parece tener que ver al menos parcialmente con el contenido informativo de las series utilizadas en la estimación (la ratio de fiabilidad relevante para la estimación de la OCDE es de 0,246). De hecho, nuestra estimación del parámetro de interés con datos regionales españoles está claramente dentro del rango de meta-estimaciones obtenido para la OCDE en D&D (2006) tras corregir por el sesgo derivado de la existencia de errores de medición.

ALGUNAS IMPLICACIONES

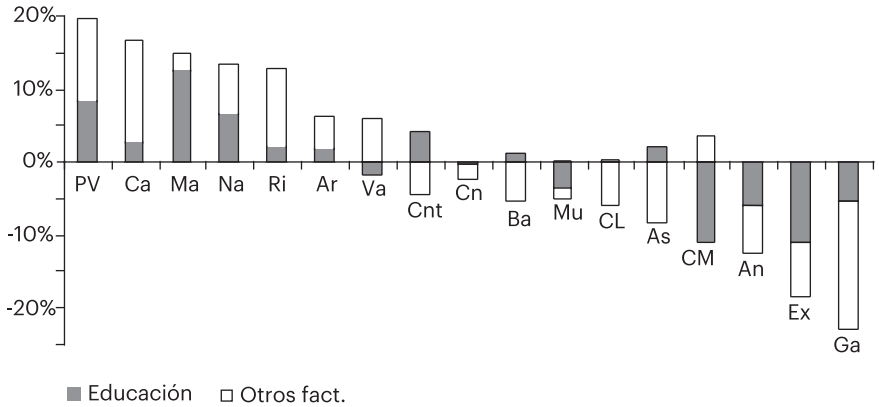
Los resultados resumidos en las secciones anteriores tienen algunas implicaciones importantes. Si el nivel medio de formación entra en la función de producción con un coeficiente cuyo valor está dentro del rango que hemos estimado, las diferencias en niveles educativos son una de las principales fuentes de diferenciales de productividad entre los países de la OCDE y las regiones españolas. Asimismo, la inversión en educación tiene una rentabilidad sustancial que, en la mayor parte de los territorios, mejora la disponible en otras formas alternativas de inversión.

Aunque no dispongo del espacio necesario para entrar en muchos detalles, no quisiera cerrar esta sección sin al menos una breve elaboración de estas dos afirmaciones. Utilizando las estimaciones recogidas en el Cuadro 3 y los datos subyacentes, en D&D (2008) hemos calculado la contribución de la educación a la productividad relativa de las regiones españolas, medida por el logaritmo del *output* real por puesto de trabajo expresado en desviaciones sobre la media muestral no ponderada de la misma variable. El Gráfico 2 muestra la descomposición de la productividad relativa de cada región en un componente educativo y un residuo que

recoge el efecto conjunto de todos los demás factores. Utilizando pesos de regresión para ponderar las distintas regiones, encontramos que la contribución media de la educación a la productividad relativa regional fue de un 40% en 1995 –esto es, que para la región española típica, el nivel educativo explica 4/10 del diferencial de productividad con la media muestral–[10]. Un cálculo similar para la OCDE atribuye a la educación un peso del 30%.

Nuestras estimaciones también implican que la rentabilidad social de la educación es muy respetable[11]. Combinando nuestros resultados sobre la contribución de la educación a la productividad con algunas estimaciones rápidas de su impacto sobre el empleo y con datos de gasto educativo, estimamos tasas de rentabilidad social en el año 1995 de entre el 10,1% y el 12,6% para España y de entre 8,3%

GRÁFICO 2. CONTRIBUCIÓN DE LA EDUCACIÓN A LA PRODUCTIVIDAD RELATIVA EN 1995



Clave: An = Andalucia; Ar = Aragón; As = Asturias; Ba = Baleares; Cn = Canarias; Cnt = Cantabria; CL = Castilla y León; CM = Castilla la Mancha; Cat = Cataluña; Va = Valencia; Ex = Extremadura; Ga = Galicia; Ma = Madrid; Mu = Murcia; Na = Navarra; PV = País Vasco; Ri = Rioja.

10. Definimos la productividad relativa de la región *i* ($qrel_i$) como la diferencia logarítmica entre el *output* por puesto de trabajo en la región y la media geométrica muestral de la misma variable. La contribución del capital humano a la productividad regional (cs_i) se obtiene multiplicando el coeficiente de este factor, α_s, por el nivel educativo relativo, también medido en diferencias logarítmicas con la media. Después de construir estas dos variables para cada región, estimamos una regresión de la forma
$cs_i = a * qrel_i + e_i$
donde e_i es una perturbación aleatoria. El coeficiente obtenido de esta forma, $a \cong cs_i / qrel_i$, mide la fracción del diferencial observado de productividad que, en promedio sobre toda la muestra, puede atribuirse al capital humano.
11. La rentabilidad social de la educación se define como la tasa de descuento que iguala el valor presente descontado de los incrementos de *output* generados por un incremento marginal del nivel educativo con el valor presente descontado de los costes directos y de oportunidad de la educación. Para más detalles sobre cómo estimar esta magnitud, véase de la Fuente (2003).

y 11,5% en la UE15[12]. En ambas muestras, las rentabilidades estimadas se comparan muy favorablemente con las disponibles en otras inversiones alternativas en la mayor parte de los territorios y tienden a ser más altas de lo que cabría esperar sobre la base de estimaciones microeconométricas de los rendimientos privados de la educación, lo que sugiere que puede haber importantes externalidades en juego. Ambos resultados sugieren que en la mayor parte de los territorios que hemos estudiado, una reasignación de recursos hacia la inversión en educación sería socialmente deseable.

CONCLUSIÓN

Los economistas académicos han sido tradicionalmente optimistas sobre la contribución de la educación al desarrollo económico y han asignado con frecuencia un papel clave a la acumulación de capital humano en los modelos formales, especialmente en la literatura sobre crecimiento endógeno desarrollada en las últimas décadas. Los resultados de los estudios empíricos sobre los determinantes del crecimiento económico han sido generalmente consistentes con esta visión de las cosas hasta mediados de los noventa. En la segunda mitad de los noventa, sin embargo, una nueva ronda de trabajos produjo resultados muy desalentadores sobre la contribución de la inversión educativa al crecimiento de la productividad agregada que generaron un animado debate entre "creyentes" y "escépticos" en relación con los efectos de la educación sobre el crecimiento de la productividad.

El presente trabajo contiene una panorámica selectiva y en algún sentido militante de la literatura relevante. Tras repasar brevemente el marco teórico, el artículo se centra en un problema (la mala calidad de los datos internacionales de escolarización) que puede explicar los desalentadores resultados de algunos estudios influyentes, en posibles formas de corregir el problema y en lo que sucede cuando así se hace. He sostenido que, como resultado de las deficiencias de los datos de base, las series de escolarización utilizadas en la literatura temprana sobre crecimiento y capital humano contenían una cantidad considerable de ruido que generaba un sesgo a la baja muy importante en la estimación del parámetro que mide la contribución del nivel educativo a la productividad. Esta conclusión se basa en la estimación de un indicador estadístico del contenido informativo de las series más utilizadas en la literatura y se ve reforzada por la constatación de que existe una clara tendencia a que los coeficientes estimados del capital humano aumenten en tamaño y mejoren en precisión según aumenta el contenido informativo de las

12. Para más detalles, véase de la Fuente y Doménech (2008) y de la Fuente (2003).

series de escolarización. Cuando esta relación se extrapola para obtener estimaciones del valor del coeficiente del capital humano que se obtendría en la ausencia de errores de medición, el ejercicio sugiere que la elasticidad del *output* con respecto al *stock* de capital humano no es, casi con total seguridad, inferior a 0,60 –esto es, alrededor de dos veces mayor que la más optimista de las estimaciones de referencia en la literatura temprana sobre capital humano y crecimiento–. Si esta conclusión es correcta, la inversión en la cantidad y calidad de la educación aparece como una de las mejores instrumentos de política de los que disponen los gobiernos para influir sobre el crecimiento de la productividad y, en última instancia, sobre el nivel de vida de sus ciudadanos.

REFERENCIAS

Azariadis, C. y A. Drazen (1990). "Threshold Externalities in Economic Development". *Quarterly Journal of Economics*, May, pp. 501-26.

Barro, R. (1991). "Economic Growth in a Cross Section of Countries". *Quarterly Journal of Economics* CVI (2), pp. 407-43.

Barro, R. (1997). *Determinants of economic growth: a cross-country empirical study.* Lionel Robbins Lectures, MIT Press, Cambridge, Mass.

Barro, R. y J-W. Lee (1993). "International Comparisons of Educational Attainment". *Journal of Monetary Economics* 32, pp. 363-94.

Barro, R. y J-W. Lee (1996). "International Measures of Schooling Years and Schooling Quality". *American Economic Review* 86(2), *Papers and Proceedings*, pp. 218-23.

Barro, R., y J-W. Lee (2001). "International data on educational attainment: Updates and implications". *Oxford Economic Papers*, 3, 541-563.

Barro, R. y J-W. Lee (2013). "A New Data Set of Educational Attainment in the World, 1950-2010". *Journal of Development Economics* 104, pp. 184-98.

Baumol, W., S. A. Batey Blackman y E. Wolf (1989). *Productivity and American Leadership: the Long View*. MIT Press.

Behrman, J. y M. Rosenzweig (1994). "Caveat emptor: cross-country data on education and the labor force". *Journal of Development Economics* 44, pp. 147-71.

Benhabib, J. y M. Spiegel (1994). "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data". *Journal of Monetary Economics* 34, pp. 143-73.

Caselli, F., G. Esquivel y F. Lefort (1996) "Reopening the convergence debate: a new look at cross-country growth empirics". *Journal of Economic Growth* 1(3), pp. 363-89.

Cohen, D. y M. Soto (2007). "Growth and human capital: good data, good results". *Journal of Economic Growth* 12(1), pp. 51-76.

Coulombe, S., J. F. Tremblay y S. Marchand (2004). "Literacy scores, human capital and growth across fourteen OECD countries". Statistics Canada.

de la Fuente, A. (2003). "Human capital in a global and knowledge-based economy, part II: assessment at the EU country level". European Commission, DG for Employment and Social Affairs, Brussels. http://digital.csic.es/bitstream/10261/1844/1/57603.pdf.

de la Fuente, A. (2011). "Human capital and productivity". *Nordic Economic Policy Review* 2, pp. 103-32.

de la Fuente, A. y A. Ciccone (2003). *Human capital and growth in a global and knowledge-based economy.* European Commission, DG for Employment and Social Affairs. Office for official publications of the European Communities, Luxembourg. http://ideas.repec.org/p/aub/autbar/576.03.html.

de la Fuente, A. y R. Doménech (D&D, 2000). "Human capital in growth regressions: how much difference does data quality make?" OECD Economics Department Working Paper no. 262, Paris.

de la Fuente, A. y R. Doménech (D&D, 2001a). "Schooling data, technological diffusion and the neoclassical model". *American Economic Review* 91(2), *Papers and Proceedings*, pp. 323-27.

de la Fuente, A. y R. Doménech (D&D, 2001b). "Educational attainment in the OECD, 1960-90". CEPR Discussion Paper no. 3390.

de la Fuente, A. y R. Doménech (D&D, 2002). "Human capital in growth regressions: how much difference does data quality make? An update and further results". CEPR Discussion Paper no. 3587. (http://ideas.repec.org/e/pde52.html#papers).

de la Fuente, A. y R. Doménech (D&D, 2006). "Human capital in growth regressions: how much difference does data quality make?" *Journal of the European Economic Association 4(1)*, pp. 1-36.

de la Fuente, A. y R. Doménech (2008). "Human capital, growth and inequality in the Spanish regions". En U. Stierle, M. Stierle, F. Jennings y A. Kuah, editores, *Regional Economic Policy in Europe. Challenges for Theory, Empirics and Normative Interventions*, pp. 15-44. Edward Elgar.

de la Fuente, A. y R. Doménech (2014a). "Educational attainment in the OECD, 1960-2010, version 3.1". FEDEA, Documento de Trabajo 2014-14, Madrid. http://documentos.fedea.net/pubs/dt/2014/dt-2014-14.pdf.

de la Fuente, A. y R. Doménech (2014b). "El nivel educativo de la población en España y sus regiones: 1960-2011". Mimeo, FEDEA.

de la Fuente, A. y R. Doménech (2014c). "Cross-country data on the quantity of schooling: a selective survey and some quality measures". Documento

de Trabajo 2014-15, Madrid. http://documentos.fedea.net/pubs/dt/2014/dt-2014-15.pdf.

Hanushek, E. y D. Kimko (2000). "Schooling, labor-force quality and the growth of nations". *American Economic Review* 90(5), pp. 1184-208.

Hanushek, E. y L. Woessmann (2008). "The role of cognitive skills in economic development". *Journal of Economic Literature* 46(3), pp. 607-68.

Hanushek, E. y L. Wössman (2010). "The economics of international differences in educational achievement". NBER *Working Paper* no. 15949, Cambridge, Mass.

Hanushek, E. y L. Wössman (2012). "Do better schools lead to more growth? Cognitive skills, economic outcomes and causation". *Journal of Economic Growth*. 17(4), pp. 267-321.

Islam, N. (1995). "Growth empirics: a panel data approach". *Quarterly Journal of Economics* 110, pp. 1127-70.

Krueger, A. y M. Lindahl (2001). "Education for growth: why and for whom?" *Journal of Economic Literature* XXXIX, pp. 1101-36.

Kyriacou, G. (1991). "Level and Growth Effects of Human Capital, A Cross-Country Study of the Convergence Hypothesis". Mimeo, NYU.

Landau, D. (1983). "Government Expenditure and Economic Growth: a Cross-Country Study". *Southern Economic Journal*, January, pp. 783-92.

Lau, L, D. Jamison y F. Louat (1991). "Education and Productivity in Developing Countries: an Aggregate Production Function Approach". *Report no.* WPS 612, World Bank.

Lau, L., S. Bhalla y F. Louat (1991). "Human and physical capital stock in developing countries: construction of data and trends". Draft mimeo, *World Development Report*, World Bank.

Lucas, R. (1988). "On the Mechanics of Economic Development". *Journal of Monetary Economics* 22, pp. 3-42.

Lutz, W., A. Goujon, K.C. Samir and W. Sanderson (2007). "Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970-2000". *Vienna Yearbook of Population Research* 2007, pp. 193-235. http://www.iiasa.ac.at/Research/POP/edu07/index.html.

Mankiw, G., D. Romer y D. Weil (1992). "A Contribution to the Empirics of Economic Growth". *Quartely Journal of Economics*, pp. 407-37.

Mas. M., F. Pérez, E. Uriel, L. Serrano y A. Soler (2002). "Capital humano, series 1964-2001," en *Capital humano y actividad económica* (CD Rom). Fundación Bancaja, Valencia.

Nehru, V., E. Swanson y A. Dubey (1995). "A New Database on Human Capital Stocks in Developing and Industrial Countries: Sources, Methodology and Results". *Journal of Development Economics*, 46, pp. 379-401.

Pritchett, L. (2001). "Where has all the education gone?" *World Bank Economic Review* 15(3), pp. 367-91.

Samir KC, B. Barakat, A. Goujon, V. Skirbekk, W. Sanderson and W. Lutz (2010). "Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050". *Demographic Research*: Volume 22, Article 15.

SEGUNDA PARTE

REFLEXIONES SOBRE EDUCACIÓN Y CAPITAL HUMANO

EL IMPACTO DE LA EDUCACIÓN INFANTIL EN LOS RESULTADOS DE PRIMARIA: EVIDENCIA PARA ESPAÑA A PARTIR DE UN EXPERIMENTO NATURAL

DANIEL SANTÍN
GABRIELA SICILIA
Universidad Complutense de Madrid

RESUMEN

En España la educación infantil es voluntaria entre los cero y los cinco años y gratuita a partir de los tres años. En este trabajo analizamos el impacto de los años de asistencia a la educación infantil sobre los resultados académicos en la Educación Primaria controlando por la posible presencia de endogeneidad. Para ello analizamos la Evaluación General de Diagnóstico 2009 aplicada a alumnos de cuarto curso de Educación Primaria. Esta base de datos tiene la peculiaridad de recoger para un grupo de escuelas información sobre dos grupos del mismo curso y de preguntar al Director del centro si la asignación de alumnos a grupos se realizó de manera aleatoria. Así, para el grupo de escuelas que aleatorizan, explotamos el hecho de que por azar en cada escuela siempre habrá un grupo que, en media, tiene más años de educación infantil que el otro. De esta forma podemos analizar esta diferencia como el tratamiento en un experimento natural, que nos permite identificar de forma exógena variaciones en los años de asistencia a educación infantil y explicar así las diferencias en rendimiento académico entre los dos grupos. Los resultados muestran que los años de educación infantil impactan de forma positiva y significativa sobre el resultado en las pruebas de matemáticas y lectura. Además, hay un efecto positivo y significativo en matemáticas por asistir a la educación infantil de primer ciclo mientras que existe un efecto negativo y significativo en ambas disciplinas por ser escolarizado después de los tres años.

INTRODUCCIÓN

Una de las primeras decisiones a las que deben enfrentarse las familias es la de la educación de sus hijos. Finalizado el permiso por maternidad-paternidad empiezan las primeras preguntas acerca de si escolarizar o no al niño/a en educación infantil, la puerta de entrada al sistema educativo. En caso afirmativo, las siguientes cuestiones serían a qué edad empezar con la escolarización y en qué tipo de centro. Distintos factores influyen en estas decisiones; la posibilidad de conciliar la vida laboral y el cuidado del niño, el coste del servicio, la restricción presupuestaria familiar, la disponibilidad de abuelos o de un tercero para cuidar al niño en el hogar, la importancia que los padres dan a esta etapa educativa, además de otros factores objetivos y afectivos. A su vez, las autoridades educativas deben decidir, también condicionados por sus propias restricciones presupuestarias y por otros compromisos políticos, acerca de la edad en la que la educación infantil debe ser gratuita con carácter universal, si debe tener un carácter educativo o meramente asistencial, además de regular los contenidos y la formación del profesorado necesaria para impartir este nivel con el grado de calidad requerido, entre otras cuestiones.

En España, la educación infantil consta de dos ciclos, ambos de carácter voluntario. El primer ciclo de educación infantil comprende las edades de cero, uno y dos años mientras que el segundo ciclo abarca las edades de tres, cuatro y cinco años. Los alumnos se agrupan por cursos en cohortes tomando como referencia el año natural. La provisión del primer ciclo es fundamentalmente privada, si bien existe una red de centros públicos con precios competitivos y con becas para facilitar el acceso a las familias con menos recursos pero con una oferta limitada que no cubre la demanda de plazas existente. Ello implica que actualmente en España la educación infantil se produce prácticamente al cincuenta por ciento en centros públicos y privados, si bien, tal y como muestra la Tabla 1, estos porcentajes varían según la Comunidad Autónoma (CC. AA. en adelante) de residencia. También suelen existir distintos cheques o bonos guardería proporcionados por las CC. AA. para facilitar el acceso al primer ciclo de esta etapa educativa en centros privados, donde para acceder a los mismos hay que cumplir unas condiciones determinadas fundamentalmente relacionadas con el nivel de renta familiar y por tanto el acceso a los mismos no es generalizado.

A pesar de la percepción extendida de que la educación infantil está ya totalmente generalizada desde edades tempranas, existen aún importantes diferencias por CC. AA. Tal y como muestra la Tabla 1, la tasa neta de escolaridad con un año supera el 30% en Andalucía, Aragón, Castilla la Mancha, Cataluña, Madrid y el País Vasco, mientras que no alcanza el 10% en Canarias, Cantabria, Extremadura y Ceuta.

A los dos años las diferencias se amplían aún más, siendo por ejemplo la tasa de escolarización del País Vasco del 91,3% mientras que en Extremadura solo alcanza el 5,5%.

TABLA 1. TASAS NETAS DE ESCOLARIDAD EN 0, 1 Y 2 AÑOS Y DISTRIBUCIÓN PORCENTUAL DEL ALUMNADO DE ESTAS EDADES SEGÚN TITULARIDAD DEL CENTRO. CURSO 2011-2012

	TASA NETA DE ESCOLARIDAD				% ALUMNADO DE PRIMER CICLO	
	0-2 AÑOS	0 AÑOS	1 AÑO	2 AÑOS	CENTROS PÚBLICOS	CENTROS PRIVADOS
ESPAÑA	30,7	9,7	31,8	49,8	51,7	48,3
Andalucía	33,7	8,2	35,6	55,9	41,0	59,0
Aragón	31,7	9,2	33,8	51,3	54,0	46,0
Asturias	16,9	8,1	17,6	24,4	83,7	16,3
Baleares	21,5	7,2	23,4	33,2	68,2	31,8
Canarias	7,1	2,1	6,2	12,3	61,3	38,7
Cantabria	24,0	3,2	9,0	57,9	74,7	25,3
Castilla y León	16,9	5,4	17,0	28,0	57,9	42,1
Cast. Mancha	33,9	10,3	35,8	54,4	64,8	35,2
Cataluña	36,4	11,5	38,5	58,8	61,7	38,3
C. Valenciana	25,3	7,9	27,4	39,5	38,9	61,1
Extremadura	3,3	1,3	3,1	5,5	0,0	100,0
Galicia	22,8	9,8	22,0	36,0	81,4	18,6
Madrid	43,3	15,4	47,3	67,3	44,5	55,5
Murcia	15,6	3,4	16,1	27,0	52,0	48,0
Navarra	10,5	4,0	11,0	16,3	78,2	21,8
País Vasco	51,9	20,3	44,0	91,3	54,5	45,5
La Rioja	15,7	5,4	16,8	24,6	20,3	79,7
Ceuta	3,8	0,8	3,5	7,0	0,0	100,0
Melilla	17,8	3,6	13,9	37,2	43,9	56,1

Fuente: MECD (2014).

El segundo ciclo de educación infantil es actualmente gratuito para las familias, lo que ha llevado a que esta etapa se haya extendido progresivamente en los últimos años hasta alcanzar a más del 90% de toda la población de tres años en todas las CC. AA. (Tabla 2).

La Tabla 2 también muestra cómo, en general, el número promedio de años de asistencia a educación infantil ha ido incrementándose en España hasta alcanzar en el curso 2011-2012 una media de 3,8 años. La tendencia en los últimos quince años ha sido creciente en prácticamente todas las CC. AA. pero todavía existen diferencias entre Cataluña, Madrid y el País Vasco, cuyos alumnos superan los cuatro años de escolarización infantil al ser comparadas con Baleares, Canarias, Extremadura, Murcia, Navarra, La Rioja, Ceuta y Melilla, donde sus alumnos no superan los 3,5 años.

TABLA 2. EVOLUCIÓN DE LA TASA NETA DE ESCOLARIDAD EN 3 AÑOS Y DEL NÚMERO MEDIO DE AÑOS DE ESCOLARIDAD EN EDUCACIÓN INFANTIL

	TASA NETA DE ESCOLARIDAD A LOS 3 AÑOS			NÚMERO MEDIO DE AÑOS DE ESCOLARIDAD EN EDUCACIÓN INFANTIL		
	2001-2002	2006-2007	2011-2012	2001-2002	2006-2007	2011-2012
ESPAÑA	92,1	95,8	95,2	3,2	3,5	3,8
Andalucía	79,1	96,4	97,0	2,8	3,0	3,9
Aragón	96,6	97,1	94,2	3,1	3,8	3,8
Asturias	94,4	97,8	96,3	3,0	3,2	3,5
Baleares	98,0	91,6	91,9	3,2	3,2	3,4
Canarias	94,4	96,0	94,9	2,9	2,9	3,1
Cantabria	94,1	96,0	94,5	3,0	3,4	3,6
Castilla y León	99,0	99,2	97,2	3,2	3,4	3,5
Cast. Mancha	97,1	98,8	96,0	3,1	3,1	3,9
Cataluña	100,0	95,7	95,4	3,9	3,8	4,0
C. Valenciana	89,3	93,7	90,3	3,1	3,2	3,6
Extremadura	92,8	97,3	98,2	3,0	3,1	3,1
Galicia	91,2	97,8	96,9	3,3	3,5	3,6
Madrid	95,0	93,0	94,7	3,5	3,8	4,2
Murcia	92,9	96,6	97,3	3,2	3,4	3,4

[CONTINÚA EN PÁGINA SIGUIENTE]

TABLA 2. EVOLUCIÓN DE LA TASA NETA DE ESCOLARIDAD EN 3 AÑOS Y DEL NÚMERO MEDIO DE AÑOS DE ESCOLARIDAD EN EDUCACIÓN INFANTIL

	TASA NETA DE ESCOLARIDAD A LOS 3 AÑOS			NÚMERO MEDIO DE AÑOS DE ESCOLARIDAD EN EDUCACIÓN INFANTIL		
	2001-2002	2006-2007	2011-2012	2001-2002	2006-2007	2011-2012
Navarra	96,9	98,2	95,4	3,0	3,7	3,2
País Vasco	100,0	98,6	97,3	3,7	4,4	4,5
La Rioja	97,7	96,4	94,9	3,1	3,1	3,3
Ceuta	81,8	96,9	90,7	2,9	3,1	2,9
Melilla	84,6	93,5	82,7	3,0	3,4	3,2

Fuente: MECD (2014).

Para poder apoyar en la toma de decisiones a padres y autoridades educativas acerca del número óptimo de años de educación infantil, la pregunta clave es: ¿hasta qué punto es importante la educación infantil para el devenir futuro de los niños? Desde un punto de vista teórico y empírico existe evidencia de que la intervención pública en educación infantil está justificada para lograr tres grandes objetivos sociales[1].

En primer lugar, el acceso a la educación infantil es un medio para facilitar la igualdad de oportunidades educativas, entendida como el aumento de la probabilidad de alcanzar un nivel de estudios determinado al comienzo del sistema educativo para alumnos que por sus circunstancias parten con desventaja. Así, de acuerdo con Currie (2001), Cawley *et al.* (2001), Berlinski *et al.* (2008), Heckman (2013) y Gertler *et al.* (2014), la educación infantil proporciona a los alumnos de nivel socioeconómico más desfavorecido una oportunidad, en forma de intervención temprana, para acercar sus resultados educativos futuros a los del alumnado con mejor entorno mediante la reducción del fracaso escolar y el aumento del rendimiento académico. Además, esta intervención tiene la ventaja de no generar el no deseado *trade-off* entre eficiencia y equidad ya que permite alcanzar ambos objetivos simultáneamente (Cunha y Heckman, 2007; Heckman y Masterov, 2007).

1. Existe una vasta literatura empírica que avala la importancia de la educación infantil sobre distintas dimensiones del bienestar futuro de los individuos y sobre la acumulación de capital humano que ofrece resultados robustos en diferentes países. No es el objetivo de esta investigación realizar una revisión exhaustiva de esta literatura, ya que ese trabajo excede los objetivos de este capítulo, sino tan solo presentar una muestra de trabajos recientes. El lector interesado puede acudir a las citas de estos trabajos y a la página web del Profesor James Heckman, Premio Nobel de Economía; http://www.heckmanequation.org/ para un estudio en profundidad de esta cuestión.

En segundo lugar, la educación infantil no solo mejora las capacidades cognitivas y no cognitivas de los alumnos relacionadas con su desempeño escolar, además, existe una amplia evidencia de que mejora la productividad de los individuos y su éxito en la etapa adulta (Heckman *et al.*, 2006; Borghans *et al.*, 2008). Si bien otras políticas de mejora educativa también generarían estos resultados, la intervención a edades tempranas tiene como característica que su rentabilidad es mucho más alta en términos económicos (Heckman *et al.*, 1999; Martin y Grubb, 2001).

En tercer lugar, la educación infantil es un mecanismo potente para generar beneficios en otros aspectos de la vida de los individuos (*i.e.* contribuye a mejorar las habilidades y el estado de salud de los individuos), así como también para producir ciertas externalidades positivas en la economía como la reducción de riesgos sociales asociados a la criminalidad o el embarazo de adolescentes (Nores y Barnett, 2010; Hines *et al.*, 2011; Campbell *et al.*, 2014).

A pesar de la amplia literatura internacional, en España existe menos evidencia empírica que permita establecer una relación causal entre educación infantil y resultados educativos. Ello es posiblemente debido a la dificultad de encontrar una estrategia correcta de identificación de la verdadera causalidad entre los años de educación infantil y el rendimiento educativo[2]. En efecto, si los años de asistencia a la etapa de educación infantil están correlacionados con variables no observables (por ejemplo la motivación de la familia, la ayuda que recibe en el hogar o el interés de los padres en que su hijo asista a una buena escuela y en que obtenga altas calificaciones), y a su vez estas variables influyeran positivamente en el resultado académico de los alumnos, entonces un modelo de regresión múltiple posiblemente sobreestimará el impacto de la asistencia a educación infantil, y por tanto sus resultados podrían estar sesgados y no ser consistentes. Para el caso de España cabe destacar como excepciones los trabajos de Hidalgo-Hidalgo y García-Pérez (2012) y de Felfe *et al.* (2013).

Hidalgo-Hidalgo y García-Pérez (2012) utilizan la información contenida en TIMSS-PIRLS para estudiar el impacto de la asistencia a educación infantil sobre los resultados en lengua, matemáticas y lectura. Para controlar los efectos potenciales de la endogeneidad, estos autores emplean como variable adicional la población potencial de alumnos por centro de educación infantil en cada región durante los años en los que cada alumno pudo asistir a infantil en función de su edad. Además, emplean un modelo de efectos fijos para controlar la posible heterogeneidad en-

2. Para una excelente discusión acerca de los problemas que origina la endogeneidad en la estimación de modelos econométricos en el ámbito educativo, así como para conocer los principales métodos experimentales y cuasi-experimentales de identificación de la causalidad puede acudirse a Schlotter *et al.* (2012) y Webbink (2005). Para una ilustración de los conceptos básicos de las metodologías de evaluación de impacto (grupo tratado, grupo control, sesgo de selección, etc.) puede acudirse al capítulo de evaluación educativa en el marco de este mismo proyecto.

tre grupos debido al número de años medio de educación infantil. Sus resultados concluyen que los alumnos que asistieron a educación infantil al menos tres años obtienen entre 13 y 16 puntos más en el resultado en lectura, es decir son un tres por ciento mejores en su rendimiento que los que asistieron menos años.

Felfe *et al.* (2013) estudian el impacto que tuvo la extensión gratuita de la educación infantil a los tres años a principios de los años noventa con la implantación de la LOGSE (Ley Orgánica General del Sistema Educativo) en el curso 1991-1992 sobre el rendimiento académico y la repetición de curso a los 15 años. Para ello explotan el experimento natural que se produjo como consecuencia de la distinta extensión de la reforma a lo largo de las CC. AA. utilizando los datos de las oleadas de PISA 2003, 2006 y 2009. Utilizando el método de diferencias en diferencias encuentran que la reforma supuso un incremento a los 15 años del resultado en lectura (matemáticas) de entre 0,10 y 0,13 (0,07) desviaciones estándar. Además encuentran que este efecto es más pronunciado en las chicas y en alumnos procedentes de niveles socioeconómicos bajos.

Vista la importancia de cuantificar el impacto de la educación infantil en España el objetivo de este capítulo es doble. En primer lugar se propone una metodología para identificar la causalidad entre los años de asistencia a educación infantil y los resultados en una etapa educativa posterior basada en la asignación aleatoria de alumnos dentro de cada escuela, y por tanto en la creación de un experimento natural. En segundo lugar se presenta una aplicación empírica para el caso de España utilizando datos de la Evaluación General de Diagnóstico realizó el Ministerio de Educación en 2009 donde se cuantifica el impacto de los años de asistencia a educación infantil sobre el rendimiento en matemáticas y lectura de los alumnos de cuarto curso de Educación Primaria.

Para cumplir con estos dos objetivos el trabajo queda organizado de la siguiente manera. El siguiente apartado discute la metodología y la estrategia de identificación de la causalidad del efecto de la educación infantil en el rendimiento académico de los alumnos. En el tercer apartado se presenta la base de datos y las variables clave en las estimaciones, además de comprobar empíricamente que realmente existe aleatoriedad en la distribución de alumnos entre los grupos evaluados. El cuarto apartado recoge los resultados de los modelos y por tanto la cuantificación del impacto de la asistencia a la educación infantil sobre los resultados educativos en matemáticas y lectura de alumnos que cursan cuarto de primaria. Para finalizar, el último apartado estará dedicado a concluir el capítulo y a presentar algunas sugerencias sobre el futuro de la educación infantil en España.

METODOLOGÍA Y ESTRATEGIA DE IDENTIFICACIÓN DE LA CAUSALIDAD

Proponemos un modelo de función de producción educativa a nivel de grupos educativos para estimar cuál es el efecto medio en el rendimiento académico del grupo cuando este tiene en media más años de educación infantil que otro grupo. Así, el resultado promedio de N alumnos n=1, 2,…N asignados al grupo j; j=1, 2,…J; en la escuela s: s = 1…S vendrá determinado por la siguiente función:

$$\bar{A}_{js} = \alpha_0 + \beta_p P_{js} + \beta_B \bar{B}_{js} + \beta_s S_s + \beta_k \bar{K}_{js} + \beta_g \bar{g}_{js} + \beta_h h_s + \varepsilon_{js} \qquad (1)$$

donde $\bar{A}_{js}$ denota el resultado medio en una prueba de rendimiento objetiva de los N alumnos del grupo j en la escuela s. Este rendimiento medio depende de un conjunto de variables observables y no observables. Así, P_{js} muestra las características observables asociadas al grupo que incluyen las características del profesor encargado del grupo, sus prácticas docentes y otras características relevantes del grupo como el tamaño de la clase; $\bar{B}_{js}$ recoge el nivel socioeconómico y las características observables medias de los alumnos del grupo; S_s es un factor que muestra las características observables del centro educativo, que lógicamente son las mismas para todos los grupos que pertenecen a la misma escuela y finalmente $\bar{K}_{js}$ es un factor asociado al número de años medio de educación infantil que asistieron los alumnos que pertenecen al grupo j.

Además de las variables observables anteriores, la Ecuación 1 recoge dos variables adicionales no observables; un factor $\bar{g}_{js}$ que recoge características no observables medias del grupo, por ejemplo habilidades cognitivas o no cognitivas de los alumnos o la motivación de sus padres por lograr una buena educación en sus hijos o características no observables de los profesores relacionadas fundamentalmente con su motivación y esfuerzo; y un término h_s que recoge características no observables relacionadas con la gestión eficiente y el liderazgo de la dirección del centro educativo además de un término de error aleatorio ε_{js}.

Asumimos que h_s y $\bar{g}_{js}$ son distintas a lo largo de las escuelas. En primer lugar, no todas las escuelas realizan una gestión eficiente de los recursos de los que disponen aunque parece razonable asumir que la gestión, buena o mala, de una misma escuela afecta por igual a todos los niños que en ella estudian, independientemente del grupo al que asisten. En segundo lugar, los padres más motivados por la educación de sus hijos tienden a recoger más información tanto a la hora de elegir las mejores escuelas como en las decisiones cotidianas que afectan a su educación. Igualmente, los maestros más efectivos y motivados suelen dedicar más tiempo a informarse acerca de la calidad de las escuelas a la hora de elegir el centro educativo

donde trabajar. El término $\bar{g}_{js}$ será en media similar o no en ambos grupos dentro del mismo centro dependiendo del criterio en el que se base la distribución de los alumnos en cada grupo.

Por tanto, un aspecto crucial es la regla que se utiliza para asignar los alumnos a los grupos dentro de cada escuela. En algunos centros los directores utilizarán algún criterio no aleatorio de agrupación de los alumnos; por ejemplo a partir de la habilidad puesta de manifiesto por ellos, su idioma materno o algún otro criterio no aleatorio. Lo mismo ocurriría cuando los profesores o los padres más motivados pueden presionar para elegir grupo o para cambiar a su hijo al que consideran mejor grupo. En ambos casos, estas variables no observables que también afectan al resultado académico de los alumnos, serían capturadas por el término de error de la Ecuación 1, el cual resultaría correlacionado con las restantes variables observables incluidas en el modelo como regresores. Esta correlación produciría un sesgo a la hora de estimar los coeficientes de la Ecuación 1, y por tanto la estimación del verdadero impacto de las variables observables sobre el resultado académico no sería correcta.

Así, si asumimos que solo existen dos grupos j = 1, 2 dentro de cada escuela[3] *s*, el rendimiento medio de cada grupo vendría dado por las ecuaciones (2) y (3):

$$\bar{A}_{1s}=\alpha_{01}+\beta_p P_{1s}+\beta_B \bar{B}_{1s}+\beta_s S_s+\beta_k \bar{K}_{1s}+\beta_g \bar{g}_{1s}+\beta_h h_s+\varepsilon_{1s} \quad (2)$$

$$\bar{A}_{2s}=\alpha_{02}+\beta_p P_{2s}+\beta_B \bar{B}_{2s}+\beta_s S_s+\beta_k \bar{K}_{2s}+\beta_g \bar{g}_{2s}+\beta_h h_s+\varepsilon_{2s} \quad (3)$$

Con lo cual, el modelo en diferencias entre los dos grupos sería:

$$(\bar{A}_{1s}-\bar{A}_{2s}) = (\alpha_{01}-\alpha_{02}) + \alpha_p (P_{1s}-P_{2s}) + \delta_b (\bar{B}_{1s}-\bar{B}_{2s}) + \Upsilon_k (\bar{K}_{1s}-\bar{K}_{2s}) + \eta_g (\bar{g}_{1s}-\bar{g}_{2s}) + \varepsilon_s \quad (4)$$

donde $\bar{A}_{1s}-\bar{A}_{2s}$ denota la diferencia en el resultado educativo entre los dos grupos de la misma escuela *s*. Si los alumnos no son asignados aleatoriamente a los grupos existirá un sesgo de selección producido, bien por la dirección del centro al diseñar la composición de los grupos y asignar los profesores a cada grupo o bien porque los padres más motivados elegirán el que consideran mejor grupo para sus hijos. Todo ello producirá que en general $(\alpha_{01}-\alpha_{02}) \neq 0$; $E\,(P_{1s}-P_{2s}) \neq 0$; $E\,(\bar{B}_{1s}-\bar{B}_{2s}) \neq 0$; $E\,(\bar{K}_{1s}-\bar{K}_{2s}) \neq 0$ y $E\,(\bar{g}_{1s}-\bar{g}_{2s}) \neq 0$. Como mencionamos anteriormente, en este caso, $\bar{g}_{js}$ no es observada por el investigador, lo que supone que entre los dos grupos existirá una diferencia significativa que implicará que la estimación de la ecuación (4) esté generalmente sesgada y sea inconsistente.

3. Aunque la extensión a más grupos es inmediata, la metodología se presenta considerando únicamente dos grupos por simplicidad. En caso de que existieran más grupos se tomaría uno de los grupos como referencia y se calcularían las diferencias del resto de grupos con respecto al grupo referenciado.

Sin embargo, en otras ocasiones las escuelas asignan los alumnos a los grupos de forma aleatoria. En estos casos, la diferencia media esperada para las variables observables y no observables asociadas a los alumnos entre ambos grupos será cero debido a la aleatorización. Si contamos con una muestra de M escuelas donde dentro de cada escuela los N alumnos se asignaron a los J grupos aleatoriamente, podemos explotar la variabilidad que surge entre las covariables al realizar la aleatorización escuela por escuela. En efecto, en este caso tendremos que las diferencias en resultados académicos entre grupos serán debidas al efecto del profesor, que siempre es distinto, pero también al efecto de otras diferencias en covariables donde "por azar" se hayan generado diferencias. En este caso, la suma de las diferencias medias esperadas en las variables de los dos grupos para todas las escuelas será igual a cero, pero dado que el número de alumnos dentro de cada escuela es limitado, se deberían encontrar diferencias entre los grupos en la variable de interés que pueden ser relevantes[4].

La idea de esta metodología es considerar que en cada escuela se produce un experimento natural donde podemos observar y explotar las pequeñas diferencias que se han generado por azar considerando en cada centro como grupo tratado (control) al grupo que en promedio cuenta con mayores (menores) valores de la variable que queramos estudiar[5], en nuestro caso los años de educación infantil. En resumen, dentro de cada escuela tendremos dos grupos. El grupo tratado (T) será el grupo donde por azar el número de años promedio de asistencia a educación infantil por parte de los alumnos es más elevado Alternativamente el grupo de control (C) está formado por el grupo de alumnos con menos años promedio de educación infantil, $K_{TS} > K_{CS}$. La diferencia en rendimiento entre los grupos T y C a lo largo de todas las escuelas $\bar{A}_{Ts}$- $\bar{A}_{Cs}$ que asignan a los alumnos aleatoriamente vendrá dada por la ecuación (5):

$$(\bar{A}_{Ts} - \bar{A}_{Cs}) = (\alpha_{01} - \alpha_{02}) + \alpha_p (P_{Ts} - P_{Cs}) + \delta_B (\bar{B}_{Ts} - \bar{B}_{Cs}) + \Upsilon_k (\bar{K}_{Ts} - \bar{K}_{Cs}) + \eta_g (\bar{g}_{Ts} - \bar{g}_{Cs}) + \varepsilon_s \qquad (5)$$

donde $\bar{A}_{Ts}$- $\bar{A}_{Cs}$ denota las diferencias entre los grupos tratado y control en la variable de interés relacionada con el resultado educativo. En este caso, tendremos que

4. Sería equivalente a lanzar al aire una moneda un número limitado de veces N (número de alumnos en cada grupo de cada escuela que en este ejemplo asumimos que es el mismo) en los dos grupos de cada escuela asignando el resultado a cada grupo. El proceso se realiza en los dos grupos N *x* 2 y se repite este procedimiento M veces (una por cada escuela). La esperanza matemática de la diferencia entre 'caras y cruces' al considerar las M *x* 2 *x* N tiradas es cero. Sin embargo, para cada escuela en presencia de muestras pequeñas de tamaño N, por azar, normalmente el resultado en uno de los grupos presentará una diferencia en la covariable de interés sobre el otro que podemos considerar como el tratamiento en este experimento natural. En nuestro caso, el resultado de tirar 25 veces la moneda en cada uno de los dos grupos de cada escuela normalmente producirá que uno de los grupos tenga más caras que el otro.
5. En este trabajo estudiamos el impacto que la educación infantil tiene en el rendimiento académico. No obstante, esta misma metodología podría ser empleada para analizar la influencia de otras variables.

$E(\alpha_{01}-\alpha_{02})=0, E(P_{Ts}-P_{Cs})=0, E(\bar{B}_{Ts}-\bar{B}_{Cs})=0$ y $E(\bar{g}_{Ts}-\bar{g}_{Cs})=0$ y mientras que $E(\bar{K}_{Ts}-\bar{K}_{Cs})\neq 0$ debido a que hemos organizado los datos como si se trataran de un experimento natural. Así definido, el modelo permite establecer a partir de la estimación del coeficiente γ_k qué parte de la diferencia de los resultados entre los dos grupos en las escuelas es debida al tratamiento y cuantificar su efecto.

ANÁLISIS EMPÍRICO EN ESPAÑA

LA EVALUACIÓN GENERAL DE DIAGNÓSTICO 2009. DATOS Y VARIABLES

Para llevar a cabo las estimaciones utilizaremos la información de la Evaluación General de Diagnóstico (EGD) del Ministerio de Educación, Cultura y Deporte de España aplicada durante 2009 a una muestra de alumnos de cuarto curso de primaria a lo largo de todo el territorio español. La muestra seleccionada contenida en esta base de datos es representativa a nivel nacional[6]. La EGD 2009 fue aplicada a 28.708 alumnos repartidos en 887 centros educativos, aunque efectivamente la base de datos incluye información para 27.125 alumnos en 882 centros.

La EGD está enfocada en medir los conocimientos, destrezas y actitudes adquiridos por los estudiantes en cuatro áreas del conocimiento que definen las competencias básicas: 'lingüística', 'matemática', 'social y ciudadana' y 'conocimiento y la interacción con el mundo físico'. De forma similar a otros estudios internacionales (PISA, TIMSS, PIRLS), complementariamente a las pruebas se provee un cuestionario a los alumnos y sus familias, a los directores de los centros y al profesorado que proporciona información adicional sobre factores contextuales, de recursos y procesos organizativos y de aprendizaje dentro del aula y la escuela. Este amplio volumen de información permite analizar los rendimientos académicos según diversos factores y circunstancias. A continuación se detallan las variables seleccionadas para nuestro análisis y una breve descripción estadística de las mismas.

La estimación del impacto de la asistencia a la educación infantil en los resultados académicos de los alumnos será medida a través de tres variables asociadas que hemos construido en relación con la asistencia a preescolar. Para construir estas variables se utilizaron las respuestas proporcionadas por las familias a la pregunta '¿A qué años empezó el colegio o guardería?[7]'. Los años de asistencia a

6. Para una descripción detallada del proceso de selección muestral en la EGD-2009 véase INEE (2010).
7. Las posibles respuestas eran 'antes de los dos años', 'a los dos años', 'a los tres años', 'a los cuatro años', 'a los cinco años' y 'a los seis o más años'. La pregunta se realizó a los padres y a los alumnos por lo que para su construcción se tomó la respuesta de los padres y solo en el caso en que el padre no respondiera se tomó la contestación proporcionada por el alumno.

educación infantil en España van desde cero, si no asistió nunca a infantil antes de ingresar a primaria a cinco, si asiste a infantil desde antes de los dos años. La primera es una variable continua que recoge los años de asistencia a educación infantil (*Años_preescolar*). La segunda variable que se analizará es el porcentaje de alumnos en el grupo que asistieron al primer ciclo de educación infantil la cual se computa agregando a los niños que empezaron su educación entre cero y dos años (*Asiste_2*). La última variable analizada es el porcentaje de alumnos del grupo que comenzaron a asistir a educación infantil antes de los 4 años, esto es, entre los cero y los tres años (*Asiste_3*).

Las variables de resultado académico seleccionadas sobre las que se mide el efecto de la asistencia a la educación infantil son el resultado en matemáticas (*mates*) y el resultado en lectura (*lectura*), con el fin de captar dos dimensiones cognitivas vehiculares y complementarias y sobre las cuales podría tener diferente impacto la asistencia a infantil. Ambas variables surgen de promediar a nivel de grupo el resultado de los alumnos en cada prueba.

Por último, se incluyen en las distintas estimaciones las siguientes variables de control relacionadas con las características del alumnado, del grupo y de los centros educativos:

- ISEC: índice de estatus social, económico y cultural medio del grupo. Este índice es elaborado por los analistas que diseñaron la base de datos y se construye mediante un análisis factorial a partir de la respuesta de las familias en las siguientes preguntas: nivel más alto de estudios de cualquiera de los padres, profesión con un nivel más elevado de cualquiera de los padres, número de libros en el hogar y el nivel de recursos domésticos[8]. El índice resultante es una variable continua y se ha expresado como un valor tipificado con media cero y desviación típica 1 a nivel nacional.
- *Inmigrantes*: porcentaje de alumnos inmigrantes dentro del grupo.
- *Repetidores*: porcentaje de alumnos repetidores dentro del grupo.
- *Trimestre 1*: porcentaje de alumnos dentro del grupo que nacieron en el primer trimestre del año.
- *Chicas*: porcentaje de niñas en el grupo.
- *Monoparentales*: porcentaje de niños dentro del grupo que pertenecen a familias monoparentales, definida esta última en base a la respuesta del niño sobre si vive o no con su madre y/o padre biológico. Aquellos niños que han

8. Para mayor detalle ver "Evaluación general de diagnóstico 2009. Educación Primaria. Cuarto curso. Informe de resultados" (2010). Instituto de Evaluación, Ministerio de Educación.

respondido vivir con uno sólo de sus padres biológicos se computan como niños que viven en familias monoparentales.

- *Tamaño*: cantidad de alumnos en el grupo.
- *Pública*: variable dicotómica que toma valor uno si el centro es de titularidad pública y cero si se trata de un centro privado concertado[9].

ESTRATEGIA PARA LA IDENTIFICACIÓN DEL IMPACTO DE LA EDUCACIÓN INFANTIL

Una primera estrategia para estimar cómo la asistencia a educación infantil influye en los resultados académicos puede ser realizada a nivel de alumno a partir de variables individuales (Hidalgo y García-Pérez, 2012). Sin embargo, como hemos visto, si los años de asistencia a la educación infantil (*K*) están correlacionados con un factor no observable como la motivación de la familia y el interés en que asista a una buena escuela y obtenga altas calificaciones (*g*) y este factor *g* también influye en el resultado académico, tendremos un problema de endogeneidad y, por tanto, la estimación del efecto de *K* posiblemente estará sesgado y no será consistente.

La Tabla 3 muestra la estimación de este modelo a nivel individual con la muestra de los 27.125 alumnos para los que inicialmente se disponía de información en la EGD[10]. Los resultados evidencian que incluso controlando por otras variables el efecto de los años de educación infantil es fuertemente significativo tanto para matemáticas como para lectura.

Para el caso de España existe un sesgo de selección potencial en el número de años de educación infantil al que acuden los alumnos que, en caso de ser estimado convencionalmente, haría que los resultados de la Tabla 3 fueran cuestionables. Las causas por las que la decisión de los padres acerca de la cantidad de años de educación infantil de sus hijos no es aleatoria provienen fundamentalmente de dos fuentes. En primer lugar, la educación infantil de primer ciclo en general no es gratuita por lo que la mayoría de familias deben asumir un esfuerzo económico que en función de sus circunstancias laborales y familiares podrían decidir no llevar a cabo. En segundo lugar, en la elección de escuela pública o concertada algunos centros concertados ofrecen plazas en la educación infantil de primer ciclo que posteriormente facilita la escolarización final en el centro[11].

---

9. No se incluyen en el análisis los centros privados independientes.
10. Las variables '*repetidor*', '*inmigrante*', '*Trimestre1*', '*monoparental*' y '*mujer*' son variables dicotómicas que reciben valor uno si el alumno posee esa característica y cero si no la posee.
11. Por ejemplo, los padres más motivados a la hora de elegir las mejores escuelas concertadas tratarán de enviar a sus hijos a la etapa privada de educación infantil de primer ciclo con el objetivo de ganar un punto de libre disposición que en muchos casos resulta decisivo para permanecer definitivamente en el centro seleccionado.

TABLA 3. ESTIMACIÓN A NIVEL DE ALUMNO DEL EFECTO DE LA ASISTENCIA A EDUCACIÓN INFANTIL SOBRE LOS RESULTADOS EN MATEMÁTICAS Y LECTURA EN EDUCACIÓN PRIMARIA

	MATES			LECTURA		
	COEF.	DESV. EST.	P-VALOR	COEF.	DESV. EST.	P-VALOR
CONSTANTE	418,770	2,906	0,000	398,462	2,942	0,000
Años_preescolar	3,588	0,506	0,000	4,232	0,512	0,000
ISEC	24,862	0,553	0,000	27,281	0,560	0,000
Repetidor	-41,922	1,747	0,000	-51,000	1,769	0,000
Inmigrantes	-7,801	1,781	0,000	-13,428	1,803	0,000
Trimestre1	16,228	1,183	0,000	14,124	1,198	0,000
Monoparental	-5,446	1,739	0,002	-5,975	1,761	0,001
Mujer	-12,382	1,011	0,000	11,224	1,024	0,000
Pública	-7,890	1,106	0,000	-8,115	1,120	0,000

Fuente: Elaboración propia a partir de la EGD 2009.

La segunda estrategia de análisis consiste en tratar de identificar una variación exógena en la variable que queremos analizar para determinar su verdadera influencia. En nuestro caso estimamos el impacto de los años de educación infantil con un modelo en diferencias entre grupos educativos tal y como fue detallado en el apartado de la metodología. Para seguir esta dirección, es necesario seleccionar sólo aquellos centros donde existen dos grupos de alumnos, que son 442 centros del total. Asimismo, resulta clave asegurar la aleatoriedad en la asignación de los alumnos a cada grupo dentro de la misma escuela de modo que se pueda evitar cualquier tipo de sesgo de selección por parte de la dirección del centro, profesores y padres. Para ello, la EGD 2009 dispone de siete preguntas en el cuestionario aplicado a los directores que indagan acerca de los criterios de agrupación del centro. La Tabla 4 recoge estos criterios y la valoración sobre si el mismo responde o no a un criterio de asignación aleatoria.

De los 442 centros que poseen dos grupos de cuarto curso de primaria, hemos descartado 220 centros donde los criterios de agrupación declarados por el director son no aleatorios. Es decir, se excluyen los centros donde la agrupación se realiza según "motivos lingüísticos", "rendimiento académico", "buscando homogeneidad según características del alumnado" u "otro criterio". En definitiva, se considera que la asignación fue aleatoria en 434 grupos con 9.438 alumnos correspondientes a 217 centros[12] de los cuales el 65,4% son públicos y el 34,6% restante son privados.

12. De los 222 centros iniciales que poseían dos grupos y agrupaban según algún criterio aleatorio se elimina-

TABLA 4. CRITERIOS DE AGRUPACIÓN DEL ALUMNADO EN LOS CENTROS DE EDUCACIÓN PRIMARIA

PREGUNTA EN EGD	CRITERIO DE AGRUPACIÓN LLEVADO A CABO	ALEATORIO
PD26A	Orden alfabético de apellidos u otro criterio aleatorio	SÍ
PD26B	Equilibrio entre niñas y niños	SÍ
PD26C	Motivos lingüísticos	NO
PD26D	Según rendimiento académico	NO
PD26E	Buscando homogeneidad según características del alumnado	NO
PD26F	Procurando la heterogeneidad entre el alumnado	SÍ
PD26G	Otro criterio	NO

Fuente: Elaboración propia a partir de la EGD 2009.

Las variables de la Ecuación 5 son computadas a nivel de grupo y posteriormente se calculan las diferencias entre los grupos Tratado (aquel con mayor valor de la variable de interés) y Control (aquel con menor valor de la variable de interés) dentro de cada centro para llevar a cabo las estimaciones (a excepción de la variable *Pública* que está definida a nivel de centro).

Tal y como fue mencionado anteriormente en el apartado metodológico, la información proveniente de la EGD 2009 permite tratar los datos como si procedieran de un cuasi-experimento o experimento natural. En este trabajo consideramos tres tratamientos; Años_preescolar, *Asiste_2 y Asiste_3* donde 'el grupo tratado' en cada centro es aquel con mayor valor de la variable[13]. Por tanto, un primer paso imprescindible antes de llevar a cabo las estimaciones de los modelos es verificar que realmente la única diferencia entre ambos grupos (tratado y control) radica en los años de educación infantil que se han producido por azar a la hora de asignar aleatoriamente alumnos a grupos dentro de cada escuela. Para ello, en la Tabla 5 se comparan las medias de ambos grupos para todas las variables relevantes cuando el tratamiento es 'Años_preescolar'.

De la Tabla 5 se desprende que efectivamente sólo se rechaza la hipótesis nula de igualdad de medias entre grupos para los años de asistencia a preescolar, mien-

ron cinco centros que presentaban valores extremadamente atípicos en alguna de las variables relevantes incluidas en el análisis.

13. El grupo tratado y el grupo de control en las tres variables coinciden en la mayoría de ocasiones aunque hay excepciones, lo que supone que en algunos colegios los grupos tratado y control intercambian su rol. Para simplificar la exposición de los resultados a lo largo del texto se muestran únicamente las estadísticas descriptivas y los test de diferencia de medias entre grupos del tratamiento *'años preescolar'*, sin embargo las regresiones son estimadas con las diferencias de los grupos de cada uno de los tres tratamientos. Los resultados y las conclusiones de los test de diferencias de media entre grupos aleatorizados y no aleatorizados con los tratamientos '*Asiste_2*' y '*Asiste_3*' no cambian y los mismos están disponibles en el Anexo de este trabajo.

tras que para las restantes variables se rechaza dicha hipótesis en todos los casos. Esto es, no existen diferencias significativas en las medias de las variables observables de ambos grupos salvo por el tratamiento. Por ello, es plausible suponer que tampoco existen diferencias significativas en las variables no observables. Resulta también muy relevante llevar a cabo este mismo análisis de diferencia de medias para los grupos de los centros excluidos del análisis por agrupar según criterios no aleatorios para dimensionar el impacto de estimar directamente el efecto de la asistencia a educación infantil con toda la muestra de colegios (Tabla 6).

TABLA 5. TEST DE IGUALDAD DE MEDIAS ENTRE LOS GRUPOS TRATADO Y DE CONTROL CUANDO LA ASIGNACIÓN DE ALUMNOS A LOS GRUPOS ES ALEATORIA

		N	MEDIA	DESVIACIÓN TÍPICA	T-TEST	P-VALOR
Años_preescolar	Tratado	217	3,879	0,348	6,283	0,000***
	Control	217	3,654	0,397		
Mates	Tratado	217	506,8	41,1	0,570	0,569
	Control	217	504,5	41,3		
Lectura	Tratado	217	508,8	42,9	1,158	0,247
	Control	217	503,8	45,6		
ISEC	Tratado	217	3,423	0,549	0,610	0,543
	Control	217	3,390	0,569		
Repetidor	Tratado	217	0,094	0,081	-1,044	0,297
	Control	217	0,103	0,096		
Inmigrantes	Tratado	217	0,089	0,120	-0,754	0,451
	Control	217	0,098	0,119		
TRI1	Tratado	217	0,244	0,101	0,855	0,393
	Control	217	0,236	0,097		
Chicas	Tratado	217	0,486	0,120	0,398	0,691
	Control	217	0,482	0,114		
Monoparental	Tratado	217	0,089	0,069	-0,980	0,327
	Control	217	0,096	0,075		
Tamaño	Tratado	217	21,687	3,465	-0,355	0,723
	Control	217	21,806	3,568		

Nota: El tratamiento es la variable *'años preescolar'*. *** La diferencia es significativa al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

TABLA 6. TEST DE IGUALDAD DE MEDIAS ENTRE LOS GRUPOS TRATADO Y DE CONTROL CUANDO LA ASIGNACIÓN DE ALUMNOS A LOS GRUPOS NO ES ALEATORIA

		N	MEDIA	DESVIACIÓN TÍPICA	T-TEST	P-VALOR
Años_preescolar	Tratado	220	3,930	0,355	7,704	0,000
	Control	220	3,658	0,386		
Mates	Tratado	220	509,3	41,6	1,769	0,078*
	Control	220	502,2	42,6		
Lectura	Tratado	220	506,4	41,5	1,469	0,143
	Control	220	500,5	43,3		
ISEC	Tratado	220	3,461	0,514	1,695	0,091*
	Control	220	3,373	0,573		
Repetidor	Tratado	220	0,089	0,080	-1,303	0,193
	Control	220	0,099	0,088		
Inmigrantes	Tratado	220	0,088	0,112	-2,647	0,008***
	Control	220	0,121	0,147		
TRI1	Tratado	220	0,235	0,096	-0,508	0,612
	Control	220	0,239	0,099		
Chicas	Tratado	220	0,475	0,114	-1,254	0,211
	Control	220	0,488	0,109		
Monoparental	Tratado	220	0,091	0,068	-1,136	0,257
	Control	220	0,099	0,077		
Tamaño	Tratado	220	21,800	3,829	0,112	0,911
	Control	220	21,759	3,821		

Nota: El tratamiento es la variable '*años preescolar*'. * La diferencia es significativa al 90% / *** La diferencia es significativa al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

Los resultados muestran cómo en este caso sí existen algunas diferencias significativas en otras covariables. En particular existen diferencias entre las medias del ISEC, matemáticas y el porcentaje de inmigrantes que son más favorables para el grupo tratado, y por tanto podrían existir también diferencias en las variables no observables, lo cual confirma la necesidad de excluir estos centros del análisis para garantizar la aleatoriedad del cuasi experimento y evitar el posible sesgo en los re-

sultados. Una vez justificada la selección final de escuelas, la Tabla 7 presenta los principales descriptivos de las variables en diferencias incluidas en los análisis.

La Tabla 7 muestra cómo la media de las diferencias de medias es prácticamente igual a cero en todas las variables menos en el tratamiento y en el resultado en matemáticas y lectura. En este caso, sí que podemos inferir que la aleatoriedad produce que esta diferencia de medias también sea cero en las variables no observables.

TABLA 7. ESTADÍSTICAS DESCRIPTIVAS DE LAS DIFERENCIAS MEDIAS DE LAS VARIABLES INCLUIDAS EN LOS ANÁLISIS ENTRE LOS GRUPOS TRATADO Y CONTROL

	N	MEDIA	DESVIACIÓN TÍPICA	MÍNIMO	MÁXIMO
Dif_Años_preescolar	217	0,23	0,18	0,00	0,77
Dif_mates	217	2,3	33,6	-71,0	106,8
Dif_lectura	217	4,9	32,2	-85,2	85,3
VARIABLES DE CONTROL					
Dif_ISEC	217	0,03	0,30	-0,59	1,03
Dif_repetidor	217	-0,01	0,10	-0,34	0,25
Dif_inmigrante	217	-0,01	0,07	-0,26	0,19
Dif_TRI1	217	0,01	0,13	-0,41	0,45
Dif_monoparental	217	-0,01	0,09	-0,27	0,24
Dif_chicas	217	0,00	0,11	-0,37	0,35
Dif_tamaño	217	-0,12	2,38	-6	8
Pública	217	0,65	0,48	0	1

Nota: El tratamiento es la variable 'años preescolar'.
Fuente: Estimaciones propias a partir de la EGD 2009.

RESULTADOS DE LAS ESTIMACIONES

Se estimaron dos modelos, uno para cada variable dependiente (matemáticas y lectura), para cada uno de los tres tratamientos asociados a la asistencia a la educación infantil. En primer lugar se analizó el efecto que la diferencia en años de asistencia a educación infantil media tiene sobre la diferencia obtenida en matemáticas y lectura respectivamente (Tabla 8).

Los resultados evidencian un efecto positivo y significativo de la mayor asistencia media a la educación infantil tanto sobre el resultado medio del grupo en

TABLA 8. ESTIMACIÓN DEL EFECTO DE LOS AÑOS DE ASISTENCIA A EDUCACIÓN INFANTIL SOBRE EL RESULTADO EN MATEMÁTICAS Y EN LECTURA

	DIF_MATES				DIF_LECTURA			
VARIABLES EXPLICATIVAS	**COEF.**	**DESV. EST.**	**COEF. TIPIF**	**P-VALOR**	**COEF.**	**DESV. EST.**	**COEF. TIPIF**	**P-VALOR**
Dif_Años_ preescolar	31,03	13,00	0,16	0,018**	22,68	12,54	0,13	0,072***
Dif_ISEC	5,36	7,66	0,05	0,485	13,20	7,39	0,12	0,075***
Dif_repetidor	-72,24	23,84	-0,20	0,003*	-61,55	23,00	-0,18	0,008*
Dif_inmigrantes	6,32	31,64	0,01	0,842	36,98	30,54	0,08	0,227
Dif_TRI1	22,60	16,91	0,09	0,183	20,72	16,32	0,09	0,206
Dif_ monoparental	-18,20	24,98	-0,05	0,467	23,90	24,11	0,07	0,323
Dif_chicas	-7,58	19,64	-0,03	0,700	12,44	18,96	0,04	0,512
Dif_tamaño	1,47	0,96	0,10	0,129	0,91	0,93	0,07	0,327
Pública	-9,05	4,75	-0,13	0,058***	0,96	4,58	0,01	0,834
Constante	0,33	4,44	---	0,941	-1,43	4,28	---	0,739

Nota: Estimación mediante MCO. *Coeficiente significativo al 90% / **Coeficiente significativo al 95% / ***Coeficiente significativo al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

matemáticas como en lectura, siendo el primer efecto más significativo (99%) que el segundo (90%). Un incremento de una desviación estándar en la diferencia de los años de asistencia a educación infantil (0,18 años, o lo que es lo mismo en torno a dos meses) produce un incremento de 0,16 (0,13) desviaciones estándar del resultado en matemáticas (lectura) del grupo, es decir en torno a 5,5 (4,2) puntos. O lo que es lo mismo, un año de asistencia a educación preescolar de diferencia entre los dos grupos produciría una mejora en el resultado medio en matemáticas (lectura) de aproximadamente 31 (23) puntos, es decir una ganancia de un 6% (4,4%) en una clase con un resultado medio de 500 puntos. Esta mejora es muy similar, aunque algo más reducida, a la obtenida por Berlinski *et al.* (2009) para el caso de Argentina, quienes cifraron la ganancia de un año de educación infantil en una mejora en los resultados educativos de tercer curso de primaria de un 8%.

La estimación de este modelo muestra también otros resultados interesantes para algunas de las variables de control presentes en el análisis. Así, la presencia de

más repetidores en un grupo con respecto al otro disminuye el resultado medio de la clase. Este resultado podría ser debido probablemente a dos efectos; por un lado el resultado más bajo de los alumnos repetidores y por otro lado su posible influencia en que el profesor vaya algo más lento en el desarrollo de contenidos y que por tanto la presencia de estos alumnos repetidores afecte al resto de compañeros de clase. Este resultado pone de manifiesto la conveniencia de repartir a los alumnos repetidores entre ambos grupos equiproporcionalmente y de realizar un trabajo específico con ellos. En la prueba de lectura un nivel socioeconómico de la clase ligeramente mayor también afecta (con una significación estadística del 90%) a la diferencia media en lectura. En cuanto a los resultados en matemáticas se observa que las diferencias entre dos grupos pertenecientes a la educación pública son menores que entre dos grupos que asisten a la educación privada. Este resultado podría indicar que el efecto de tener un buen o mal profesor es en media más acusado en la educación privada que en la pública, y que esta última logra compensar en cierto modo las diferencias iniciales en las características de los alumnos.

A continuación analizamos si se producen diferencias entre los dos grupos cuando consideramos como tratamiento el porcentaje de niños que asistieron a la educación infantil de primer ciclo (entre cero y dos años). La Tabla 9 muestra los resultados.

En este caso el tratamiento resulta ser solo significativo (95%) para el resultado en matemáticas. Así, un 10% de diferencia en alumnos que han asistido al primer ciclo de educación infantil supone un aumento en la diferencia media de la prueba de matemáticas de 5,5 puntos. Nuevamente la diferencia en el porcentaje de alumnos repetidores para los dos grupos penaliza significativamente el resultado de la media de la clase. En este caso además, la presencia de más alumnos nacidos en el primer trimestre del año se asocia significativamente (90%) con mejores resultados en las dos pruebas. Al igual que en los años de educación escolar, las diferencias en matemáticas (no así en lectura) entre los dos grupos como consecuencia del tratamiento disminuyen por el hecho de asistir a una escuela pública. Para el caso de lectura, las diferencias en nivel socioeconómico entre los dos grupos influyen positivamente en que la diferencia en el resultado medio de lectura entre los dos grupos sea significativa.

Finalmente, analizamos el porcentaje de alumnos que asistieron a la educación infantil antes de los cuatro años. Al ser la educación en España gratuita desde los tres años, la mayoría del alumnado empieza su escolarización a esta edad. Es por ello que lo que se evalúa a través de este tratamiento es cómo influye tener en el grupo alumnos que empezaron su escolarización más allá de los tres años. La Tabla 10 muestra los resultados.

TABLA 9. ESTIMACIÓN DEL EFECTO DE LA ASISTENCIA A EDUCACIÓN INFANTIL ANTES DE LOS 3 AÑOS SOBRE EL RESULTADO EN MATEMÁTICAS Y EN LECTURA

	DIF_MATES				DIF_LECTURA			
VARIABLES EXPLICATIVAS	**COEF.**	**DESV. EST.**	**COEF. TIPIF**	**P-VALOR**	**COEF.**	**DESV. EST.**	**COEF. TIPIF**	**P-VALOR**
Dif_Asiste_2	54,92	26,05	0,14	0,036**	-3,88	25,49	-0,01	0,879
Dif_ISEC	7,46	7,77	0,07	0,338	15,21	7,60	0,14	0,047**
Dif_repetidor	-81,61	23,77	-0,23	0,001***	-69,16	23,27	-0,20	0,003***
Dif_inmigrante	-16,53	31,75	-0,03	0,603	22,30	31,07	0,05	0,474
Dif_TRI1	32,98	16,72	0,13	0,050*	27,34	16,36	0,11	0,096*
Dif_ monoparental	-11,94	25,11	-0,03	0,635	22,40	24,57	0,06	0,363
Dif_chicas	-8,27	19,72	-0,03	0,675	13,35	19,30	0,05	0,490
Dif_tamaño	1,10	0,97	0,08	0,259	0,69	0,95	0,05	0,467
Pública	-9,86	4,77	-0,14	0,040**	-2,08	4,66	-0,03	0,656
Constante	-1,23	4,52	---	0,785	2,91	4,42	---	0,512

Nota: Estimación mediante MCO. *Coeficiente significativo al 90% / **Coeficiente significativo al 95% / ***Coeficiente significativo al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

Los resultados de la estimación evidencian que la presencia de más alumnos escolarizados antes de los cuatro años en el grupo supone un aumento significativo en la diferencia media de resultados entre ambos grupos tanto para matemáticas como para lectura. En este caso, diez puntos de diferencia en el porcentaje de alumnos que asistieron con tres años o antes suponen un aumento del resultado en matemáticas (lectura) de 9 (8,5) puntos. La variable relacionada con el porcentaje de repetidores sigue siendo significativa en ambos resultados, mientras que en este caso un mayor nivel socioeconómico únicamente produce un mejor resultado en lectura.

TABLA 10. ESTIMACIÓN DEL EFECTO DE LA ASISTENCIA A EDUCACIÓN INFANTIL ANTES DE LOS 4 AÑOS SOBRE EL RESULTADO EN MATEMÁTICAS Y EN LECTURA

	DIF_MATES				DIF_LECTURA			
VARIABLES EXPLICATIVAS	**COEF.**	**DESV. EST.**	**COEF. TIPIF**	**P-VALOR**	**COEF.**	**DESV. EST.**	**COEF. TIPIF**	**P-VALOR**
Dif_asiste_3	91,33	47,58	0,14	0,056***	85,18	45,08	0,14	0,060***
Dif_ISEC	6,74	7,69	0,06	0,382	14,30	7,29	0,13	0,051***
Dif_repetidor	-71,77	24,30	-0,20	0,004*	-55,25	23,03	-0,16	0,017**
Dif_inmigrante	10,69	32,38	0,02	0,742	44,32	30,68	0,10	0,150
Dif_TRI1	25,62	16,98	0,10	0,133	21,60	16,08	0,09	0,181
Dif_ monoparental	-17,12	25,21	-0,05	0,498	25,29	23,88	0,07	0,291
Dif_chicas	-9,53	19,95	-0,03	0,633	16,42	18,91	0,06	0,386
Dif_tamaño	1,37	0,97	0,10	0,157	0,71	0,91	0,05	0,439
Pública	-6,30	4,90	-0,09	0,200	1,52	4,64	0,02	0,743
Constante	-1,13	4,13	---	0,785	0,00	3,91	---	0,999

Nota: Estimación mediante MCO. *Coeficiente significativo al 90% / **Coeficiente significativo al 95% / ***Coeficiente significativo al 1%.
Fuente: Estimaciones propias a partir de la EGD 2009.

CONCLUSIONES

De todos los análisis efectuados se desprenden las siguientes conclusiones. En primer lugar, dado que la asistencia a educación infantil en España no es exógena, la estimación directa de su impacto sobre el rendimiento académico ofrece serias dudas acerca de qué parte del efecto positivo que se encuentra responde a la variable analizada y qué parte a las variables observables y no observables que influyen en que una familia decida enviar más o menos años a su hijo a este ciclo educativo que tiene en general carácter voluntario.

Para corregir este posible sesgo tenemos en cuenta el hecho de que en un conjunto de escuelas, aproximadamente la mitad de la muestra de la EGD 2009, la distribución de alumnos entre grupos se realiza en base a criterios aleatorios. Considerar esta aleatoriedad en la asignación es un paso fundamental para alcanzar una buena estrategia de identificación ya que garantiza que las diferencias medias

esperadas en variables observables y no observables de los alumnos entre los dos grupos a lo largo de todas las escuelas sean igual a cero. A partir de esta distribución aleatoria, explotamos el hecho de que se produce un experimento natural ya que generalmente habrá un grupo que en media tiene más años de educación infantil que el otro. De esta forma, podemos estimar la importancia de esta diferencia para explicar las diferencias de resultados académicos entre grupos.

Los resultados muestran que los años recibidos de educación infantil afectan significativamente y de forma positiva al rendimiento académico de los alumnos, lo cual podría sugerir que la educación infantil tiene efectivamente un carácter formativo y no meramente asistencial. El principal resultado de este trabajo muestra que un año de asistencia a educación preescolar de diferencia entre los dos grupos produciría una mejora en el resultado medio en matemáticas (lectura) de aproximadamente 31 (23) puntos, es decir una ganancia de un 6% (4,4%) en una clase con un resultado medio de 500 puntos. Asimismo, los resultados muestran que hay un efecto negativo en el grupo donde más alumnos comienzan la educación infantil después de los tres años. En cuanto a la educación infantil de primer ciclo (entre los cero y los dos años), parece existir un efecto positivo y significativo en cuarto de primaria pero solamente para el resultado en matemáticas. Dado que las estimaciones se han realizado a nivel de diferencias inter-grupos podemos señalar que además del posible efecto negativo para el propio alumno con una baja asistencia a educación infantil, se pueda producir una externalidad negativa sobre el resto de alumnos ('efecto compañeros'). Posiblemente el canal que explicaría este resultado sea que el profesor se adapte a las características de sus alumnos y compense a los alumnos con menos años de educación infantil con más tiempo y actividades de refuerzo en detrimento de los alumnos con más años de educación infantil. En definitiva, los resultados apuntan a que la composición socioeconómica y educativa del grupo afecta significativamente a los resultados incluso dentro de los grupos de una misma escuela.

A partir de estos resultados cabe plantearse que para mejorar el sistema educativo español a través de la educación infantil resultaría positivo abrir un debate que tenga en cuenta dos líneas de trabajo. En primer lugar, dado que la educación infantil importa, debería plantearse una ampliación de la gratuidad de la educación infantil a los dos años. En caso de que los recursos presupuestarios disponibles no sean suficientes, cabría entonces al menos aumentar las becas para escolarizar de forma temprana a aquellos alumnos socioeconómicamente más desfavorecidos, que son quienes tienden a no asistir en esta etapa de escolarización. En segundo lugar, la educación infantil tiene un carácter formativo y no meramente asistencial. Es por ello que debería también plantearse la necesidad de analizar mediante méto-

dos experimentales las distintas actividades pedagógicas y lúdicas que se realizan en el aula para promover una mejor calidad y mayor homogeneidad entre los centros que permita incrementar el retorno de esta etapa educativa.

REFERENCIAS

Berlinski, S., Galiani, S. y Gertler, P. (2009). "The effect of pre-primary education on primary school performance". *Journal of Public Economics*, 93 (1), 219-234.

Berlinski, S., Galiani, S. y Manacorda, M. (2008). "Giving children a better start: Preschool attendance and school-age profiles". *Journal of Public Economics*, 92 (5-6), 1416-1440.

Borghans, L., Duckworth, A. L., Heckman, J. J. y Ter Weel, B. (2008). "The economics and psychology of personality traits". *Journal of Human Resources*, 43 (4), 972-1059.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E. y Pan, Y. (2014). "Early childhood investments substantially boost adult health". *Science*, 343 (6178), 1478-1485.

Cawley, J., Heckman, J. J. y E. J. Vytlacil (2001). "Three observations on wages and measured cognitive ability". *Labour Economics* 8 (4), 419-442.

Cunha, F. y Heckman, J. J. (2007). "The Technology of Skill Formation". *American Economic Review*, 97 (2), 31-47.

Currie, J. (2001). "Early childhood education programs". *Journal of Economic perspectives*, 15 (2), 213-238.

Felfe, C., Nollenberger, N. y Rodríguez-Planas, N. (2013). "Can't buy mommy's love? Universal childcare and children's long-term cognitive development". *Journal of Population Economics*, 1-30.Felfe, C.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., ... y Grantham-McGregor, S. (2014). "Labor market returns to an early childhood stimulation intervention in Jamaica". *Science*, 344 (6187), 998-1001.

Heckman, J. J. (2008). "Schools, skills, and synapses". *Economic inquiry*, 46 (3), 289-324.

Heckman, J. J. (2013). *Giving kids a fair chance*. MIT Press.

Heckman, J. J., R. J. LaLonde y J. A. Smith (1999). "The economics and econometrics of active labor market programs". In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3A, Chapter 31, pp. 1865-2097. New York: North-Holland.

Heckman, J. J. y Masterov, D. V. (2007). "The productivity argument for investing in young children". *Applied Economic Perspectives and Policy*, 29 (3), 446-493.

Heckman, J. J., Stixrud, J. y Urzua, S. (2006). "The effects of cognitive and noncogni-

tive abilities on labor market outcomes and social behavior". *Journal of Labor Economics* 24(3), 411-482.

Hidalgo-Hidalgo, M. y García-Pérez, J. I. (2012). "Impacto de la asistencia a Educación Infantil sobre los resultados académicos del estudiante en Primaria". En PIRLS-TIMSS 2011. Volumen II, *Informe español*, 105-142.

Hines, P., McCartney, M., Mervis, J. y Wible, B. (2011). "Laying the foundation for lifetime learning". *Science*, 333 (6045), 951-951.

INEE (2010) *Evaluación General de Diagnóstico 2009. Eduación Primaria. Cuarto Curso. Informe de Resultados*. Ministerio de Educación, Madrid.

Martin, J. P. y D. Grubb (2001). "What works and for whom: A review of OECD countries' experiences with active labor market policies". *Swedish Economic Policy Review* 8(2), 9-56.

MECD (2014). *Las cifras de la educación en España*. Ministerio de Educación, Cultura y Deporte. Madrid.

Nollenberger, N. y Rodríguez-Planas, N. (2011). "Child care, maternal employment and persistence: a natural experiment from Spain". IZA *Discussion paper*, Nº 5888.

Nores, M. y Barnett, W. S. (2010). "Benefits of early childhood interventions across the world: (Under) Investing in the very young". *Economics of Education Review*, 29 (2), 271-282.

Rivkin, S. G., Hanushek, E. A. y Kain, J. F. (2005). "Teachers, schools, and academic achievement". *Econometrica*, 73(2), 417-458.

Schlotter, M., Schwerdt, G. y Woessmann, L. (2011). "Econometric methods for causal evaluation of education policies and practices: a non-technical guide". *Education Economics*, 19(2), 109-137.

Webbink, D. (2005). "Causal effects in education". *Journal of Economic Surveys*, 19(4), 535-560.

ANEXO

TABLA A1. TEST DE IGUALDAD DE MEDIAS ENTRE LOS GRUPOS TRATADOS Y DE CONTROL CUANDO LA ASIGNACIÓN DE ALUMNOS A LOS GRUPOS ES ALEATORIA Y EL TRATAMIENTO ES *ASISTE_2*

		N	MEDIA	DESVIACIÓN TÍPICA	EST. T	P-VALOR
Asiste_3	Tratado	217	0,630	0,164	6,832	0,000***
	Control	217	0,517	0,179		
Mates	Tratado	217	505,4	41,0	-0,146	0,884
	Control	217	505,9	41,5		
Lectura	Tratado	217	507,5	43,5	0,564	0,573
	Control	217	505,1	45,1		
ISEC	Tratado	217	3,435	0,551	1,072	0,284
	Control	217	3,378	0,565		
Repetidor	Tratado	217	0,095	0,081	-0,762	0,447
	Control	217	0,102	0,096		
Inmigrante	Tratado	217	0,088	0,118	-0,901	0,368
	Control	217	0,098	0,121		
TRI1	Tratado	217	0,245	0,098	1,048	0,295
	Control	217	0,235	0,100		
Chicas	Tratado	217	0,490	0,119	1,018	0,309
	Control	217	0,478	0,114		
Monoparental	Tratado	217	0,093	0,071	0,211	0,833
	Control	217	0,092	0,073		
Tamaño	Tratado	217	21,558	3,567	-1,121	0,263
	Control	217	21,935	3,457		

Nota: *Coeficiente significativo al 90% / **Coeficiente significativo al 95% / ***Coeficiente significativo al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

TABLA A2. TEST DE IGUALDAD DE MEDIAS ENTRE LOS GRUPOS TRATADOS Y DE CONTROL CUANDO LA ASIGNACIÓN DE ALUMNOS A LOS GRUPOS ES ALEATORIA Y EL TRATAMIENTO ES *ASISTE_3*

		N	MEDIA	DESVIACIÓN TÍPICA	EST. T	P-VALOR
Asiste_3	Tratado	217	0,963	0,057	7,964	0,000***
	Control	217	0,908	0,084	7,964	
Mates	Tratado	217	506,2	42,2	0,291	0,771
	Control	217	505,1	40,2	0,291	
Lectura	Tratado	217	509,5	42,5	1,495	0,136
	Control	217	503,1	45,8	1,495	
ISEC	Tratado	217	3,421	0,555	0,543	0,587
	Control	217	3,392	0,563	0,543	
Repetidor	Tratado	217	0,092	0,082	-1,436	0,152
	Control	217	0,104	0,095	-1,436	
Inmigrante	Tratado	217	0,090	0,120	-0,635	0,526
	Control	217	0,097	0,119	-0,635	
TRI1	Tratado	217	0,243	0,099	0,699	0,485
	Control	217	0,236	0,099	0,699	
Chicas	Tratado	217	0,480	0,120	-0,732	0,465
	Control	217	0,488	0,114	-0,732	
Monoparental	Tratado	217	0,090	0,070	-0,762	0,447
	Control	217	0,095	0,075	-0,762	
Tamaño	Tratado	217	21,765	3,475	0,109	0,913
	Control	217	21,728	3,558	0,109	

Nota: *Coeficiente significativo al 90% / **Coeficiente significativo al 95% / ***Coeficiente significativo al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

TABLA A3. TEST DE IGUALDAD DE MEDIAS ENTRE LOS GRUPOS TRATADOS Y DE CONTROL CUANDO LA ASIGNACIÓN DE ALUMNOS A LOS GRUPOS NO ES ALEATORIA Y EL TRATAMIENTO ES *ASISTE_2*

		N	MEDIA	DESVIACIÓN TÍPICA	EST. T	P-VALOR
Asiste_2	Tratado	220	0,654	0,176	7,355	0,000***
	Control		0,527	0,186		
Mates	Tratado	220	508,26	41,6	1,241	0,215
	Control		503,26	42,8		
Lectura	Tratado	220	506,59	41,6	1,545	0,123
	Control		500,34	43,2		
ISEC	Tratado	220	3,45	0,518	1,355	0,176
	Control		3,38	0,571		
Repetidor	Tratado	220	0,094	0,082	0,103	0,918
	Control		0,093	0,086		
Inmigrante	Tratado	220	0,096	0,118	-1,356	0,176
	Control		0,113	0,144		
TRI1	Tratado	220	0,236	0,096	-0,193	0,847
	Control		0,238	0,099		
Chicas	Tratado	220	0,483	0,114	0,304	0,762
	Control		0,480	0,11		
Monoparental	Tratado	220	0,092	0,069	-0,686	0,493
	Control		0,097	0,076		
Tamaño	Tratado	220	21,786	3,643	0,037	0,970
	Control		21,773	3,999		

Nota: *Coeficiente significativo al 90% / **Coeficiente significativo al 95% / ***Coeficiente significativo al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

TABLA A4. TEST DE IGUALDAD DE MEDIAS ENTRE LOS GRUPOS TRATADOS Y DE CONTROL CUANDO LA ASIGNACIÓN DE ALUMNOS A LOS GRUPOS NO ES ALEATORIA Y EL TRATAMIENTO ES *ASISTE_3*

		N	MEDIA	DESVIACIÓN TÍPICA	EST. T	P-VALOR
Asiste_3	Tratado	220	0,962	0,061	8,351	0,000***
	Control		0,902	0,086		
Mates	Tratado	220	509,6	41,4	1,918	0,056*
	Control		501,9	42,8		
Lectura	Tratado	220	505,9	41,4	1,219	0,223
	Control		501,0	43,5		
ISEC	Tratado	220	3,454	0,528	1,444	0,149
	Control		3,379	0,561		
Repetidor	Tratado	220	0,088	0,082	-1,368	0,172
	Control		0,099	0,086		
Inmigrante	Tratado	220	0,083	0,108	-3,606	0,000***
	Control		0,127	0,149		
TRI1	Tratado	220	0,237	0,097	-0,095	0,925
	Control		0,238	0,099		
Chicas	Tratado	220	0,479	0,115	-0,452	0,651
	Control		0,484	0,109		
Mono-parental	Tratado	220	0,093	0,068	-0,548	0,584
	Control		0,096	0,077		
Tamaño	Tratado	220	21,864	3,831	0,461	0,645
	Control		21,695	3,818		

Nota: *Coeficiente significativo al 90% / **Coeficiente significativo al 95% / ***Coeficiente significativo al 99%.
Fuente: Estimaciones propias a partir de la EGD 2009.

TABLA A5. ESTADÍSTICAS DESCRIPTIVAS DE LAS DIFERENCIAS MEDIAS DE LAS VARIABLES INCLUIDAS EN LOS ANÁLISIS ENTRE LOS GRUPOS TRATADOS Y CONTROL. TRATAMIENTO: *ASISTE_2*

	N	MEDIA	DESVIACIÓN TÍPICA	MÍNIMO	MÁXIMO
Dif_asiste_2	217	0,11	0,09	0,00	0,44
Dif_mates	217	-0,60	33,7	-89,2	106,8
Dif_lectura	217	2,40	32,5	-85,2	85,3
Dif_ISEC	217	0,06	0,30	-0,81	1,03
Dif_repetidor	217	-0,01	0,10	-0,34	0,25
Dif_inmigrante	217	-0,01	0,07	-0,26	0,18
Dif_TRI1	217	0,01	0,13	-0,39	0,45
Dif_monoparental	217	0,00	0,09	-0,27	0,24
Dif_chicas	217	0,01	0,11	-0,37	0,31
Dif_tamaño	217	-0,38	2,35	-8	6
Pública	217	0,65	0,48	0	1

Fuente: Elaboración propia a partir de la EGD 2009.

TABLA A6. ESTADÍSTICAS DESCRIPTIVAS DE LAS DIFERENCIAS MEDIAS DE LAS VARIABLES INCLUIDAS EN LOS ANÁLISIS ENTRE LOS GRUPOS TRATADO Y CONTROL. TRATAMIENTO: *ASISTE_3*

	N	MEDIA	DESVIACIÓN TÍPICA	MÍNIMO	MÁXIMO
Dif_asiste_3	217	0,05	0,05	0,00	0,26
Dif_mates	217	1,2	33,7	-71,0	106,8
Dif_lectura	217	6,3	31,9	-85,2	85,3
Dif_ISEC	217	0,03	0,30	-0,64	1,03
Dif_repetidor	217	-0,01	0,09	-0,34	0,25
Dif_inmigrante	217	-0,01	0,07	-0,26	0,19
Dif_TRI1	217	0,01	0,13	-0,41	0,45
Dif_monoparental	217	-0,01	0,09	-0,27	0,24
Dif_chicas	217	-0,01	0,11	-0,37	0,35
Dif_tamaño	217	0,04	2,38	-6	8
Pública	217	0,65	0,48	0	1

Fuente: Elaboración propia a partir de la EGD 2009.

LA MARATÓN DE PISA: LA PERSEVERANCIA COMO FACTOR DEL ÉXITO EN UNA PRUEBA DE COMPETENCIAS

PAU BALART
ANTONIO CABRALES
Universidad Carlos III y UCL Economics

RESUMEN

Existe una creciente atención sobre la importancia de la personalidad y las habilidades no cognitivas en el futuro personal y profesional de los estudiantes. Estudios recientes relacionan ciertos rasgos de la personalidad con unos mejores resultados académicos, laborales o de salud. Dada la no observabilidad directa de las habilidades cognitivas y no cognitivas, existe un complejo problema de identificación que constituye el principal obstáculo en este campo de investigación.

Encontramos, sin embargo, algunas estrategias para intentar diferenciar las habilidades cognitivas de las no cognitivas. Como apuntan Borghans y Schils (2012), en pruebas estandarizadas como PISA, generalmente los estudiantes responden peor las últimas preguntas que las primeras. Este decaimiento del rendimiento, no está relacionado con las competencias evaluadas en la prueba y se usa como una medida de la perseverancia de los estudiantes o de otras habilidades no cognitivas.

España es uno de los países donde los estudiantes presentan un mayor decaimiento del rendimiento. La diferencia en el porcentaje de respuestas correctas entre la fase inicial y final de la prueba es parecida a la diferencia en el porcentaje de respuestas correctas entre España y Finlandia, uno de los países con mejores resultados PISA.

En el presente capítulo se estudia el decaimiento del rendimiento de los estudiantes españoles en la prueba PISA 2009 con tres objetivos: I) determinar qué características personales y familiares afectan el decaimiento en el rendimiento y las habilidades no cognitivas; II) llevar a cabo una primera aproximación sobre la incidencia que pueden tener algunas decisiones en política educativa sobre las habilidades no cognitivas; y, III) estudiar si existen diferencias relevantes en el decaimiento en el rendimiento de los estudiantes de las distintas comunidades autónomas y ver si esto contribuye a explicar las notables disparidades en los resultados PISA que se observan entre autonomías.

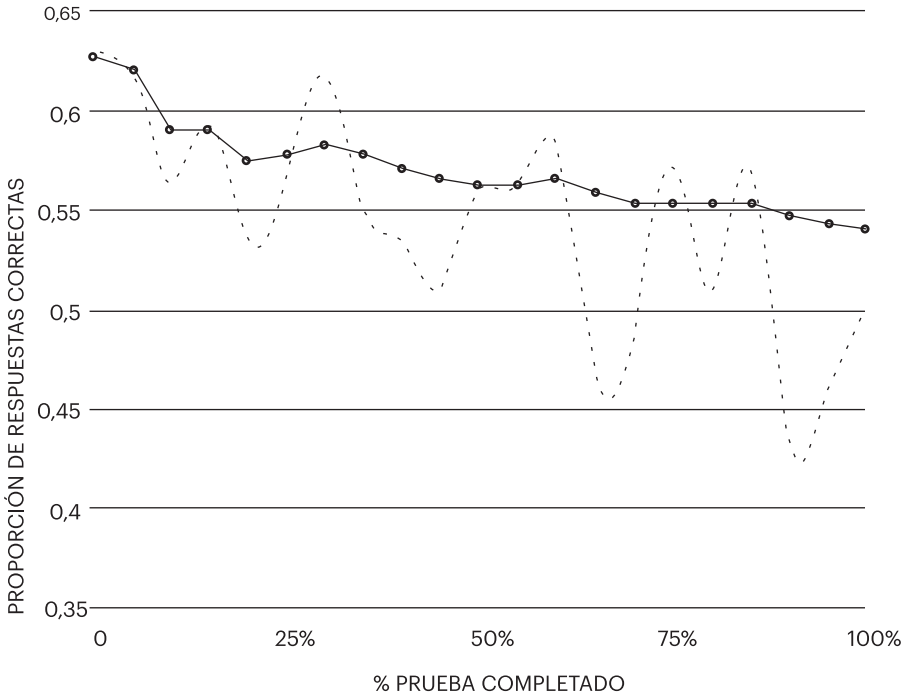
INTRODUCCIÓN

La adquisición de conocimientos juega un papel fundamental tanto en el desarrollo personal como profesional de los individuos. Además, tal y como muestran Hanushek y Woessmann (2011 y 2012), unos buenos resultados en pruebas estandarizadas de competencias, como es el *Programme for International Student Assessment* (PISA), influyen positivamente en el crecimiento económico a largo plazo de los países.

Al mismo tiempo, hay una creciente atención en entender qué miden exactamente las pruebas internacionales de competencias como PISA y los test de inteligencia. Tal y como plantean Duckworth *et al.* (2009), Eklöf (2007) o Segal (2011), el rendimiento de los estudiantes medido a partir de este tipo de pruebas refleja tanto sus habilidades cognitivas como no cognitivas. Se entiende por habilidades cognitivas aquellos conocimientos evaluados en pruebas de competencias como la habilidad lectora o los conocimientos en ciencias y matemáticas de los estudiantes. En cambio, las habilidades no cognitivas, se refieren a aspectos relacionados con la personalidad de los estudiantes, como por ejemplo su motivación intrínseca o su perseverancia. Rosen *et al.* (2010) ofrecen una clasificación de estas habilidades no cognitivas. Algunos estudios recientes señalan que las habilidades no cognitivas pueden resultar tan importantes como las cognitivas para el futuro personal y profesional de los estudiantes, véase Heckman *et al.* (2006) o Levin (2012) entre otros. El trabajo de Almlund *et al.* (2010) contiene una revisión exhaustiva de la evidencia sobre la relación entre la personalidad y las habilidades no cognitivas, así como sobre la importancia de estas en el desarrollo futuro de los individuos. Dada la no observabilidad directa de las habilidades cognitivas y no cognitivas, existe un complejo problema de identificación. Este problema de identificación constituye el principal obstáculo en este campo de investigación.

Existen sin embargo, algunas estrategias para intentar diferenciar las habilidades cognitivas de las no cognitivas. Como apuntan Borghans y Schils (2012), en pruebas estandarizadas como PISA, generalmente los estudiantes responden peor las últimas preguntas que las primeras. En la Figura 1 se puede apreciar este fenómeno usando la muestra española de la prueba PISA 2009. Tal y como refleja el gráfico, el porcentaje de respuestas correctas va disminuyendo a medida que avanza la prueba (la línea discontinua muestra el porcentaje de acierto en cada pregunta según la posición que ocupa medida en percentiles, mientras que la línea continua muestra la tendencia de estos datos). Borghans y Schils llaman a este fenómeno el decaimiento en el rendimiento y muestran como su correlación con el resultado de la prueba es baja. Esta baja correlación indica que el decaimiento del rendimiento no está relacionado con las habilidades en lectura, ciencias o matemáticas eva-

FIGURA 1 . PORCENTAJE DE RESPUESTAS CORRECTAS DE LOS ESTUDIANTES A MEDIDA QUE AVANZA LA PRUEBA PISA. ELABORACIÓN PROPIA A PARTIR DE LOS MICRODATOS DE LA MUESTRA ESPAÑOLA PISA 2009



luadas en el test. Intuitivamente, dado que los conocimientos adquiridos por los estudiantes no son susceptibles de variar durante el desarrollo de la prueba, estos ejercerán una escasa influencia sobre el decaimiento. En cambio, otras cualidades como la motivación, la perseverancia o la concentración, son susceptibles de manifestarse a través de un mayor decaimiento. Dada la baja correlación entre el decaimiento en el rendimiento y las habilidades puramente cognitivas, Borghans y Schils (2012) vinculan el decaimiento en el rendimiento a la influencia de las habilidades no cognitivas y muestran como estas medidas ejercen un efecto positivo sobre indicadores futuros del desarrollo personal y profesional de los individuos. Al mismo tiempo, dado que al inicio de la prueba no existe decaimiento, los autores utilizan el resultado al principio de la prueba como una medida de las competencias de los individuos menos influenciada por las habilidades no cognitivas.

Borghans y Schils (2012) calculan el decaimiento en el rendimiento que experimentan los estudiantes en las pruebas PISA 2006 y lo usan como medida de las habilidades no cognitivas. Estos autores explotan que el diseño de la prueba PISA sitúa las preguntas en distintas posiciones en cada versión del test donde aparece. En su trabajo estiman una regresión donde la variable dependiente es la puntua-

ción de cada estudiante en cada una de las preguntas y la variable explicativa la posición relativa de la pregunta en la versión de la prueba del estudiante en cuestión. Este procedimiento, permite usar el coeficiente de la posición relativa de la pregunta como una medida del decaimiento. En particular, informa de cuanto peor se responde una pregunta a medida que esta ocupa una posición más avanzada en el test.

Mediante este procedimiento, estos autores ofrecen una comparativa internacional del decaimiento del rendimiento. Según sus resultados, España es uno de los países de la OCDE dónde los estudiantes experimentan un mayor decaimiento del resultado. Al mismo tiempo, estos autores proporcionan una medida del resultado PISA de los distintos países corregida por el impacto del decaimiento en el rendimiento. Teniendo en cuenta esta medida, España obtiene unos resultados notablemente mejores a los habituales en esta prueba. A la vista de esta comparativa, el decaimiento en el rendimiento podría ser un elemento importante en los malos resultados cosechados habitualmente por España en la prueba PISA y una posible vía de mejora. Así, el mal rendimiento medio de los estudiantes españoles en PISA podría ser sobre todo consecuencia de una cierta carencia en habilidades no cognitivas relacionadas con la perseverancia y la constancia. El principal objetivo de este capítulo será proporcionar un estudio más exhaustivo de los factores que influyen en el decaimiento del rendimiento. Esto permitirá una primera aproximación a la distinción entre el efecto que las decisiones en política educativa pueden tener bien a nivel cognitivo o bien a nivel no cognitivo. Este enfoque también permitirá estudiar si existen diferencias relevantes en las habilidades no cognitivas de los estudiantes de las distintas Comunidades Autónomas y en qué medida estas habilidades no cognitivas permiten explicar las notables disparidades entre los resultados PISA de las Comunidades Autónomas, Villar (2012).

A diferencia de Borghans y Schils (2012) el presente estudio incluirá en la estimación la interacción de la posición relativa de cada pregunta con las características personales, familiares y del centro docente de los estudiantes. Esto, permitirá ver qué características son susceptibles de afectar el rendimiento de los estudiantes y distinguir si este efecto ocurre a través de habilidades cognitivas o no cognitivas. Además, se propone un segundo método para calcular el decaimiento basado en un enfoque individualizado para cada estudiante. Esto sirve para añadir robustez a los resultados.

Este análisis aporta novedades interesantes acerca del canal de influencia de algunas de las características personales y familiares que afectan el resultado de los estudiantes. En primer lugar se observa que el sexo, resulta muy importante a la hora de explicar el decaimiento en el rendimiento. En particular, las chicas experimentan un menor decaimiento que los chicos. Esto les lleva a obtener mejores re-

sultados en la prueba PISA 2009 a pesar de obtener los mismos resultados que los chicos al principio del test[1]. En cambio, una situación socioeconómica favorable, afecta positivamente el rendimiento de los estudiantes desde el inicio de la prueba y aunque también mitiga el decaimiento, este efecto marginal es muy inferior a su efecto inicial en el resultado. Existen también algunas variables como el mes de nacimiento, recibir más de un año de educación infantil o convivir con ambos progenitores que afectan únicamente el resultado al principio de la prueba pero no el decaimiento. Finalmente, otras variables como las interrupciones en clase, el número de libros que hay en casa o ser inmigrante podrían estar influenciando tanto el decaimiento como el resultado al principio de la prueba, no obstante estas influencias no son robustas a las distintas especificaciones del modelo.

El menor decaimiento en el rendimiento experimentado por las chicas sugiere la existencia de diferencias no cognitivas entre ambos sexos cuya repercusión hace a las chicas más competentes a la hora de llevar a cabo una actividad intelectual continuada como la prueba PISA. Hay estudios anteriores que mediante procedimientos muy distintos apuntan a la existencia de diferencias en las habilidades no cognitivas procedentes del sexo, véase Cornwell *et al.* (2013) o Duckworth y Seligman (2006).

El menor decaimiento en el rendimiento de las chicas puede tener implicaciones directas sobre algunas decisiones de política educativa referentes a la coeducación. Tal y como muestran Lavy y Schlosser (2011) o Ciccone y García-Fontes (2013) para el caso de España, existe un efecto de pares de género positivo y significativo de la chicas hacia los chicos, es decir la presencia de una mayor proporción de chicas influye positivamente sobre los resultados académicos de los chicos sin existir ningún efecto negativo en la otra dirección. Dada la diferencia entre las habilidades no cognitivas de chicos y chicas y estos efectos de pares, convendrá llevar a cabo más investigación sobre si los efectos de pares se producen como consecuencia de la transmisión de habilidades cognitivas o no cognitivas. Aunque de forma bastante preliminar, el presente estudio apunta en esta dirección al observar que una mayor proporción de chicas reduce el decaimiento en el rendimiento de los chicos. Esto añade nuevos argumentos a considerar en el debate de la coeducación de chicos y chicas.

---

1. PISA 2009 contiene una proporción parecida de preguntas en habilidad lectora (53,4%) y matemática o ciencias (46,6%). Además estas proporciones sufren pocas variaciones a lo largo de la prueba. Dado que las chicas tradicionalmente superan a los chicos en comprensión lectora y los chicos superan a las chicas en matemáticas y ciencias (Cornwell *et al.*, 2013), estos efectos explican los resultados en el inicio del test. En cambio, esta distribución equitativa entre preguntas en comprensión lectora y las demás no es suficiente para proporcionar resultados parecidos entre chicos y chicas una vez se tiene en cuenta el decaimiento en el rendimiento. Se incluyen pruebas de robustez teniendo en cuenta solo preguntas que no sean sobre comprensión lectora en la última sección de este capítulo.

Dado que tan solo el sexo, hacer la prueba en lengua materna y el estatus socioeconómico parecen tener un efecto importante en el decaimiento, se ampliarán los datos incluyendo otras características de los estudiantes y los datos referentes a las características de los centros educativos. Al hacer esto, se debe tener en cuenta, tal y como se explica en el trabajo que incluyen en este libro Jorge Calero y Oriol Escardíbul[2], que la composición anidada de los datos PISA requeriría el uso de un modelo multinivel para tener resultados más concluyentes sobre el efecto de las variables referentes al centro educativo. Por lo tanto, la inclusión de las variables de los centros deberá ser tomada con cierta precaución y como una primera aproximación de sus consecuencias sobre las habilidades no cognitivas de los estudiantes españoles. En esta versión ampliada de los modelos se observa que la mayoría de las variables consideradas no tienen ningún efecto sobre el decaimiento en el rendimiento. Sin embargo, se aprecia como el gusto por la lectura resulta clave a la hora de explicar el decaimiento en el resultado. Este resultado sugiere la importancia del fomento de la lectura como vía para reducir el decaimiento en el rendimiento de los estudiantes españoles. Sin embargo conviene ser cautelosos y mencionar la necesidad de un estudio más detallado de esta relación. En particular, existe un potencial problema de endogeneidad entre estas variables. Ciertas habilidades no cognitivas, responsables de que algunos estudiantes experimenten un menor decaimiento en el rendimiento, podrían ser también responsables de su mayor apego a la lectura y no al revés. Una buena relación entre alumno y profesor también está asociada a menores niveles de decaimiento.

Este análisis no permite extraer conclusiones sobre el efecto que el número de estudiantes por profesor tiene sobre el decaimiento. Una mala interpretación de los resultados podría concluir que un mayor número de estudiantes reduce el decaimiento del rendimiento. Sin embargo, la endogeneidad entre el rendimiento del estudiante y el tamaño del grupo que se le asigna será en gran medida el causante de este efecto.

Finalmente, algunas variables ofrecen resultados un tanto sorprendentes y que requerirán de más investigación como son el hecho que la existencia de un equipo deportivo del centro o un mayor uso de tecnologías de la información repercuten en un mayor decaimiento del rendimiento.

Tal y como se explica en el trabajo coordinado por Villar (2012) existen notables disparidades en el resultado PISA de las diferentes Comunidades Autónomas. En el presente capítulo, se puede apreciar cómo estas diferencias también existen tanto al principio de la prueba (habilidades cognitivas) como en el decaimiento (ha-

2. Ver capítulo "Resultados escolares y resultados de la educación".

bilidades no cognitivas). El decaimiento en el rendimiento explica las diferencias entre algunas comunidades con resultados similares al inicio de la prueba. Por ejemplo, Castilla y León tiene un resultado al principio de la prueba similar al de País Vasco, Asturias y Navarra. Sin embargo, su resultado en el global de la prueba es significativamente mejor que el de estas comunidades, gracias a un menor decaimiento en el resultado. Para poder considerar si estas diferencias en el decaimiento pueden ser consecuencia de características específicas de las distintas comunidades, se debe tener en cuenta las diferencias en el decaimiento después de controlar por las características de los estudiantes. Al hacer esto, las diferencias en el decaimiento de las distintas regiones no resultan significativas y por lo tanto se explican completamente a partir de las características observables de su población de estudiantes. Por lo tanto, el problema del decaimiento en el rendimiento de los estudiantes españoles es un problema generalizado y susceptible de mejora en el conjunto del país.

DATOS[3]

En 2009 la prueba PISA incluía 189 preguntas distintas (ítems), de las cuales 35 evaluaban competencias en matemáticas, 101 en comprensión lectora y 53 en ciencias[4]. Los ítems se agrupaban en 3 grupos (clústeres) en el caso de matemáticas y ciencias y en 7 para el caso de la comprensión lectora.

El diseño de la prueba emplea un sistema de rotación de los clústeres que da lugar a 13 versiones distintas de la prueba ("booklets"). El sistema de rotación funciona de manera que cada clúster aparece en cuatro "booklets" distintos y cada par de clústeres coincide en un sólo "booklet". En consecuencia, cada pregunta ocupa distintas posiciones en cada uno de los cuatro "booklets" en los que aparece. Véase OCDE (2012), para una explicación más detallada del sistema de rotación de las preguntas.

Los códigos PISA contienen información detallada sobre la posición de cada una de las preguntas en cada "booklet". Al mismo tiempo, los microdatos de PISA incluyen información sobre la puntuación del estudiante en cada pregunta. Siguiendo la metodología propuesta por Borghans y Schils (2012), en este estudio se explota esta información y el orden aleatorio de las preguntas que emerge del diseño rotativo de los "booklets" para calcular el decaimiento en el rendimiento.

3. Para el desarrollo del presente estudio se han utilizado los microdatos de la muestra española de PISA 2009. Véase el trabajo mencionado de los profesores Calero y Escardíbul para una descripción general de estos datos.
4. En este estudio no se han incluido las preguntas de actitud lectora.

EL DECAIMIENTO EN EL RENDIMIENTO

Los estudios internacionales de evaluación de competencias educativas como PISA han propiciado la aparición de multitud de estudios analizando los condicionantes y las consecuencias de sus resultados. Un fenómeno menos estudiado y conocido, que se observa a partir de un estudio de los datos PISA pregunta a pregunta, es el progresivo decaimiento del resultado a medida que avanza la prueba, tal y como documentan Borghans y Schils (2012). Como se observaba en la Figura 1 de la página 77, el porcentaje medio de respuestas correctas de los estudiantes españoles disminuye conforme aumenta el porcentaje de test completado. La línea discontinua muestra el porcentaje de respuestas correctas en cada uno de los cuantiles de orden cinco de la prueba. Aunque con oscilaciones, se aprecia una tendencia a una disminución en el porcentaje de respuestas correctas a medida que se consideran cuantiles más avanzados. La línea continua muestra esta tendencia de forma suavizada mediante el porcentaje de respuestas correctas acumulado a medida que avanza la prueba.

Al fijarse en la diferencia entre el porcentaje medio de respuestas correctas acumuladas al principio del test y al final, se observa que este varía en casi 9 puntos porcentuales desde un 62,76% al inicio de la prueba hasta un 53,82% al final de la misma. Esta diferencia es considerablemente mayor, llegando hasta los 16 puntos porcentuales si, por ejemplo, se compara el resultado al inicio del test con el resultado en el 15% final del test.

A la hora de interpretar estas variaciones conviene tener en cuenta la imposibilidad de establecer una relación directa entre el porcentaje de respuestas correctas y la puntuación en las pruebas PISA por dos motivos. En primer lugar, porque los resultados PISA, a diferencia del presente estudio proporcionan resultados por disciplina, separando entre matemáticas, lectura y ciencias. En segundo lugar, para estimar los resultados PISA se usa una ponderación basada en la teoría de la respuesta cognitiva, véase OCDE (2012)[5]. Aun así, la comparación entre países permite tener una idea aproximada de la magnitud del decaimiento. Al comparar el porcentaje total de respuestas correctas de España con el de uno de los países "top" de PISA como es Finlandia, se aprecia como la diferencia entre el porcentaje de respuestas correctas entre ambos países es de casi 10 puntos (53,8% de España por el 63,7% de Finlandia). Esta diferencia es similar a la que se observa en España al comparar el porcentaje de respuestas correctas al inicio del test con el porcentaje de respuestas correctas con el test completo.

5. La nota final de PISA usa la teoría del ítem cognitivo para crear una serie de valores plausibles para cada una de las competencias evaluadas. No obstante la correlación entre cada uno de los valores plausibles subministrados por PISA y el porcentaje se respuestas correctas es alta (entre el 83,68% y el 88,99%).

Como se ha mencionado en la sección anterior, el diseño de las pruebas PISA, ordena cada una de las preguntas en cuatro posiciones diferentes en las distintas versiones del test. Por lo tanto, una misma pregunta aparecerá tanto en posiciones avanzadas como en otras más retrasadas. Este hecho permite ver que una misma pregunta presenta un menor porcentaje de acierto a medida que ocupa una posición más avanzada en el test. Por lo tanto no se puede atribuir el decaimiento en el resultado a un aumento progresivo en la dificultad de la prueba. También es importante señalar que a la hora de calcular el porcentaje de respuestas correctas, se ha diferenciado entre preguntas no alcanzadas y preguntas incorrectas o saltadas. Las preguntas no alcanzadas se han tratado como valores no observados, por lo tanto el decaimiento antes mostrado no captura las preguntas no respondidas por falta de tiempo.

Por lo tanto, dado el diseño de las pruebas PISA y el método empleado en el cálculo de las respuestas correctas cabrá atribuir el decaimiento a lo largo de la prueba a características del individuo que realiza la prueba.

Se pueden considerar distintas medidas del decaimiento en el rendimiento. La primera consiste en la diferencia entre el porcentaje final (con todo el test completado) de respuestas correctas por estudiante y el porcentaje de respuestas correctas en el x% inicial de la prueba, con x=0, 5, 10, 20.

Esta medida infravalorará el decaimiento de aquellos estudiantes que obtienen un muy mal resultado en el test, dado que un mal resultado desde el principio impide que este decaiga. Usando una analogía se diría que cuando algo está en el suelo ya no puede caer. Por esta razón se considerará también en la tasa de decaimiento entre el porcentaje de respuestas correctas en el total de la prueba y al principio (en negativo para facilitar la explicación de los resultados).

La Tabla 1 muestra las medidas propuestas de decaimiento y su correlación con el porcentaje de respuestas correctas en la prueba PISA denotada con P_i. Se observa como todas las medidas del decaimiento son significativamente negativas y presentan una correlación muy baja con la puntuación en la prueba.

Aunque medir el decaimiento en tasas evita el posible problema de subestimar el decaimiento de los estudiantes con muy malos resultados, no informa del tamaño del decaimiento en el rendimiento. En cambio, la medida en diferencias sí permite esta interpretación. Una diferencia de 10 puntos en el porcentaje de respuestas correctas, como ocurre en el caso de España si se considera la diferencia entre la primera pregunta y el resto de la prueba, es parecido a la diferencia en el porcentaje de respuestas correctas existente entre España y Finlandia.

TABLA 1. DECAIMIENTO EN EL RENDIMIENTO COMPARANDO EL X% INICIAL DEL TEST CON EL TOTAL DEL TEST COMPLETADO

DIFERENCIA

% COMPLETADO	MEDIA	INTERVALO DE CONFIANZA		CORREL. CON P_i
0%	-0,0956533	-0,1035384	-0,0877682	0,0673
5%	-0,0840773	-0,0884325	-0,0797221	0,0007
10%	-0,0519616	-0,0550595	-0,0488637	0,0817
20%	-0,0378367	-0,0400922	-0,0355812	0,0589

TASA

% COMPLETADO	MEDIA	INTERVALO DE CONFIANZA		CORREL. CON P_i
0%	-0,2401337	-0,2619205	-0,2183469	0,1366
5%	-0,1824903	-0,1942387	-0,1707419	0,1009
10%	-0,127982	-0,1362	-0,1197641	0,1727
20%	-0,0896271	-0,0953788	-0,0838753	0,1517

En consonancia con la Figura 1, se observa como el decaimiento en el rendimiento resulta menos importante a medida que aumenta la proporción inicial del test usada para calcular el decaimiento. De este modo, el porcentaje de respuestas correctas es casi 10 puntos porcentuales mayor cuando solo se tiene en cuenta la primera pregunta o 8,4 puntos si se considera el 5% inicial de las preguntas.

Finalmente se pueden considerar medidas más pronunciadas del decaimiento, comparando el x% inicial del test con el y% final. Para evitar demasiadas combinaciones se ha tenido en cuenta: x= 5,10 e y=70, 50, 20.

La Tabla 2 proporciona medidas del decaimiento y su correlación con el porcentaje de respuestas correctas en la prueba PISA mediante una comparación de las respuestas correctas en la parte final e inicial de la prueba. En este caso se obtienen mayores medidas del decaimiento que en la Tabla 1, superando en varios casos los 10 puntos de diferencia. El tamaño del decaimiento es siempre mayor al considerar un menor porcentaje tanto inicial como final de la prueba. Se puede ver como la correlación con la puntuación es baja, especialmente cuando se tienen en cuenta las diferencias.

TABLA 2. DECAIMIENTO EN EL RENDIMIENTO COMPARANDO EL RESULTADO EN EL X% INICIAL CON EL Y% FINAL DEL TEST

DIFERENCIA

X%-Y%	MEDIA	INTERVALO DE CONFIANZA		CORREL. CON P_i
5%-70%	-0,10407	-0,1088071	-0,09934	0,0224
10%-70%	-0,07195	-0,0756111	-0,06829	0,0978
5%-50%	-0,11358	-0,118445	-0,10871	0,0257
10%-50%	-0,08139	-0,0851631	-0,07762	0,1003
5%-20%	-0,14366	-0,1490757	-0,13825	0,0166
10%-20%	-0,11081	-0,1151632	-0,10645	0,0793

TASA

X%-Y%	MEDIA	INTERVALO DE CONFIANZA		CORREL. CON P_i
5%-70%	-0,32017	-0,3376335	-0,3027133	0,1853
10%-70%	-0,26760	-0,282036	-0,2531833	0,2465
5%-50%	-0,39269	-0,4138579	-0,3715316	0,1907
10%-50%	-0,32614	-0,3428191	-0,3094684	0,2465
5%-20%	-0,31064	-0,3245231	-0,2967668	0,1519
10%-20%	-0,25530	-0,2663269	-0,244286	0,2139

Dada la baja correlación entre la puntuación (medida como porcentaje de respuestas correctas) y el decaimiento en el rendimiento, éste último no se explica por aquellas habilidades cognitivas de los estudiantes evaluadas en el test (matemáticas, ciencias y comprensión lectora). Intuitivamente, el conocimiento de los estudiantes no cambia durante el test, por lo tanto el decaimiento en el resultado debe estar relacionado con otras habilidades como la motivación, el esfuerzo, la perseverancia o la concentración. Por lo tanto, tal y como proponen Borghans y Schils (2012) se usará el decaimiento en el rendimiento como una medida de las habilidades no cognitivas de los estudiantes. Estos autores muestran como efectivamente el decaimiento está relacionado con aspectos de la personalidad. También muestran como esta distinción entre habilidades cognitivas y no cognitivas elaborada a partir del decaimiento, supone impactos distintos sobre el futuro personal y profesional de los individuos.

HABILIDADES COGNITIVAS Y NO COGNITIVAS

Dada la importancia tanto de los factores cognitivos como no cognitivos en el futuro personal y profesional de los estudiantes, véase Heckman *et al.* (2006) y Heckman *et al.* (2012), será interesante usar el decaimiento para discernir qué características personales y familiares pueden tener una mayor influencia en cada tipo de habilidades.

Al mismo tiempo, la comparación internacional llevada a cabo por Borghans y Schils (2012) revela España como uno de los países con mayor decaimiento del rendimiento. Este hecho, combinado con los malos resultados medios obtenidos en la prueba, invita a estudiar con más detalle que características afectan al decaimiento como posible vía de mejora del rendimiento de nuestros estudiantes.

Dado que el sexo y las diferencias socioeconómicas han sido señaladas como relevantes tanto en el desarrollo de las habilidades no cognitivas de los estudiantes (véase Duckworth y Seligman, 2006; Cornwell *et al.*, 2013 y Heckman, 2008) como en los resultados en pruebas de evaluación como PISA (OCDE, 2011) convendrá estudiar su influencia en el decaimiento del rendimiento. Como en otros estudios, en este capítulo se considerará también el efecto de variables referentes tanto a la institución educativa como familiares y personales, véase Woessmann (2003) entre otros.

Además de las variables del sexo (variable *chica*) y características socioeconómicas (variable ESCS[6]), también se tendrán en cuenta como variables explicativas las interrupciones en clase (variable *interrupciones*), el número de libros que hay en casa (variable *libros*), hacer el test en la lengua materna (variable *idioma*), convivir con ambos progenitores (variable *fam.nuclear*), haber atendido educación infantil más de un curso (variable *Ed.infantil*), ser inmigrante tanto de primera como de segunda generación (variable *inmigrante*) o el mes de nacimiento del estudiante (variable *mes*)[7]. Aunque también se dispone de información referente a la titularidad pública o privada de la escuela, está no será incluida dada su elevada colinealidad con la características socioeconómicas de los estudiantes[8].

---

6. El índice ESCS mide el estatus socioeconómico agregando información sobre la ocupación y el nivel educativo de los padres y los recursos disponibles en el hogar.
7. Las variables categóricas del mes de nacimiento, número de libros e interrupciones en clase se utilizan asumiendo efectos lineales de sus categorías. Los resultados sobre el decaimiento del rendimiento no se ven modificados al usar variables ficticias para cada una de las categorías de estas variables. Como se verá en la siguiente sección, estas variables no resultan significativas para el decaimiento por lo tanto esto tampoco afectará la interpretación de los resultados.
8. La titularidad de la escuela resulta significativa a la hora de analizar los resultados. Sin embargo, estudios más exhaustivos revelan que esta variable es significativa como consecuencia de la selección de los estudiantes en cada tipo de escuela (véase, Anghel y Cabrales (2010), Calero y Escardíbul (2007). Este es también el caso en nuestro estudio cuando se amplía el número de variables explicativas del centro en la sección 5.

A la luz de la notable disparidad en los resultados PISA de las distintas Comunidades Autónomas se introducirán también variables ficticias con información sobre la Comunidad Autónoma a la cual pertenece el estudiante. Esto permitirá analizar si los estudiantes de las diferentes Comunidades Autónomas experimentan diferentes niveles de decaimiento y si esto contribuye a explicar las diferencias que se observan entre comunidades. PISA 2009 contenía datos de las siguientes autonomías: Andalucía, Aragón, Asturias, Baleares, Islas Canarias, Cantabria, Castilla y León, Cataluña, Galicia, La Rioja, Madrid, Murcia, Navarra, País Vasco, Ceuta y Melilla. Para las restantes comunidades no se dispone de una muestra representativa y quedan agrupadas como "otras".

Aunque la metodología del test PISA elige las preguntas que aparecen en las distintas versiones del test con el objetivo de homogeneizar la dificultad de la prueba, esta no queda totalmente igualada. Los resultados oficiales de la prueba usan un factor para corregir estas diferencias. Dado que en este estudio no se utilizan los resultados facilitados por PISA sino con el porcentaje de respuestas correctas se deberá tener en cuenta los posibles efectos fijos provenientes de las distintas versiones de la prueba. Además, el diseño del test no consigue neutralizar la variación ordinal en la dificultad de las preguntas a lo largo de la prueba. Por esta razón, en las estimaciones de este estudio se usaran también variables ficticias de las distintas versiones del test con el objetivo controlar estos efectos.

De este modo se estimará por mínimos cuadrados la siguiente ecuación:

$$P_i = \beta X_i' + \mu CCAA_i' + \delta B_i' + u_i$$

donde P_i es la puntuación (en porcentaje de respuestas correctas obtenido por el estudiante), X_i es un vector que contiene las variables con las características del estudiante antes comentadas y finalmente $CCAA_i$ y B_i son vectores que incluyen las variables ficticias usadas para especificar la Comunidad Autónoma y la versión de la prueba ("booklet"), respectivamente. Los términos β, μ y δ denotan los vectores con los coeficientes para cada una de las variables de los vectores anteriores y el término u_i denota el error.

En la Tabla 3 se muestran los coeficientes estimados para las características personales y familiares. En primer lugar se observa como todas las variables consideradas son significativas a la hora de explicar el porcentaje final de respuestas correctas de los estudiantes. Tal y como sucede en otros estudios, al comparar las variables dicotómicas se observa como ser inmigrante y recibir educación infantil resulta especialmente relevante en la puntuación final, aumentando en casi cuatro puntos el porcentaje de respuestas correctas. Por otro lado, se puede apreciar también la importancia de la variable ESCS. Considerando el rango del índice y el coeficiente es-

timado, los estudiantes con una peor situación socioeconómica pueden tener hasta 24 puntos porcentuales menos de respuestas correctas que los más favorecidos.

TABLA 3. PORCENTAJE DE RESPUESTAS CORRECTAS Y CARACTERÍSTICAS DE LOS ESTUDIANTES

	P_i	E. ESTÁNDAR
Chica	0,0110***	(0,00209)
ESCS	0,0318***	(0,00117)
ED. Infantil	0,0384***	(0,00310)
Mes	-0,00236***	(0,00031)
Idioma	-0,0177***	(0,00344)
Libros	0,0340***	(0,00091)
Interrupciones	-0,0147***	(0,00119)
FAM. Nuclear	0,00904***	(0,00301)
Inmigrante	-0,0393***	(0,00405)
Observaciones	24.215	

* p<0,1 / ** p<0,05 / *** p<0,01.

Además, como suele ocurrir en este tipo de estudios, haber nacido en los primeros meses del año (*mes*, bajo), hacer el test en la lengua materna (*idioma*=0), poseer un elevado número de libros en casa (*libros* alto), pertenecer a un grupo-clase que sufra pocas interrupciones (*interrupciones*, bajo), y convivir con ambos padres (*fam. nuclear*=1) son factores que afectan positivamente el porcentaje de respuestas correctas. Ser chica también supone unos mejores resultados en el global de la prueba. Es bien conocido que las chicas generalmente lo hacen mejor que los chicos en comprensión lectora, mientras que lo contrario es cierto en matemáticas y ciencias (OCDE, 2011). PISA 2009, contenía una distribución bastante equitativa, tanto al principio como a lo largo de la prueba, entre las preguntas sobre comprensión lectora (53,4%) y matemáticas o ciencia (46,6%). Sin embrago esta sutil mayor proporción en preguntas de comprensión lectora no explica el mejor resultado obtenido por las chicas en PISA 2009. Tal y como se mostrará, este se explica por su menor decaimiento en el rendimiento.

HABILIDADES NO COGNITIVAS

Como se ha explicado anteriormente, en este estudio se usará el decaimiento en el rendimiento como medida de las habilidades no cognitivas y se analizará el peso de las características personales, y familiares sobre el decaimiento en el rendimiento. Esto per-

mite comprobar si aquellas variables que afectan a la puntuación de los estudiantes afectan también su decaimiento en el rendimiento. Se usarán dos modelos distintos:

MODELO 1

Este primer modelo usa medidas del decaimiento individual. Se estimará la siguiente regresión con las variables anteriormente mencionadas:

$$D_i = \beta_{m1} X_i' + \mu_{m1} CCAA_i' + \delta_{m1} B_i' + v_i$$

donde, D_i es una medida de decaimiento, v_i denota el término de error y las variables explicativas son las mismas que en la regresión anterior. Dado que se puede medir el decaimiento de muchas formas distintas, se considerarán distintas especificaciones para la variable dependiente, todas medidas en tasas de decaimiento. En la Tabla 4, se presentan los coeficientes considerando el x% inicial del test y el 100%, con x= 5, 10 y 20 (columnas 1 a 3). Se incluyen también como medida la diferencia entre la primera y la segunda mitad del test (columna 4), y las diferencias entre el 20% final de la prueba y el 5% y 10% inicial (columnas 5 y 6)[9].

MODELO 2

En segundo lugar, se usará la metodología empleada por Borghans y Schils (2012). En este modelo la variable dependiente es el resultado en cada una de las respuestas de cada uno de los estudiantes participantes en la prueba. Se denota con Y_{ij}, la puntuación del estudiante i en la pregunta j. Y_{ij} tomará valores 0, ½ o 1, para respuestas incorrectas, parcialmente correctas o totalmente correctas, respectivamente[10]. Se tendrán en cuenta como variables explicativas, las características personales y familiares de los estudiantes anteriormente mencionadas así como la Comunidad Autónoma de cada estudiante. Se incluirá además la interacción de cada una de estas variables con la posición relativa en el test de cada una de las preguntas. Se denota con Q_{ij} la posición relativa en el test de la pregunta j para el estudiante i, donde Q_{ij} tomará valores entre cero y uno, donde Q_{ij}=0 (Q_{ij}=1), significará que la pregunta j ocupa la primera (última) posición en la versión de la prueba del estudiante i. Además se incluirán efectos fijos para controlar por el nivel de dificultad de cada una de las preguntas. Por lo tanto el modelo estimado será:

$$Y_{ij} = \beta_{m2} Q_{ij} X_i' + \mu_{m2} Q_{ij} CCAA_i' + \rho_{m2} X_i' + \gamma_{m2} CCAA_i' + w_{ij}$$

9. Con otras especificaciones, como considerar el 15% inicial o el 70% final de la prueba, los resultados son similares a los aquí mostrados.
10. Varias de las preguntas de la prueba solo son puntuadas como totalmente correctas (1) o incorrectas (0). Los resultados no cambian al contabilizar las respuestas parcialmente correctas como totalmente correctas o bien como totalmente incorrectas.

donde w_{ij} representa el error y las variables explicativas son las mismas que en las estimaciones anteriores. La interpretación de los coeficientes será la siguiente; los coeficientes de las variables sin interaccionar estarán midiendo el efecto de la variable en cuestión sobre el resultado en la parte inicial de la prueba, mientras que las variables interaccionadas medirán el efecto de esa variable sobre el decaimiento del resultado[11].

La principal diferencia entre ambos reside en el cálculo del decaimiento. Mientras el Modelo 1 utiliza una medida individualizada del decaimiento en el rendimiento de cada estudiante, el Modelo 2 agrega los datos de los estudiantes y se basa en una medida de la correlación entre la posición de la pregunta y su puntuación para calcular el decaimiento. Combinar las características de ambos modelos permitirá dar mayor robustez a los resultados.

El Modelo 1, permite elegir la parte del test que se emplea en el cálculo del decaimiento. Esta flexibilidad facilita, por ejemplo usar medidas menos influenciadas por efectos como el descanso que se les concede a los estudiantes en la mitad del test[12]. Además, este modelo considerará que el decaimiento es mayor en aquellas situaciones donde los estudiantes empiecen a hacerlo mal antes. Se puede pensar en un ejemplo sencillo, donde el decaimiento es un fenómeno irreversible, es decir llegados a un punto del test los estudiantes empiezan a equivocarse. Se asumirá que la prueba tiene 6 preguntas por lo tanto Q_{ij}=0; 0,2; 0,4; 0,6; 0,8; 1 y que hay un solo estudiante y se compararán las situaciones siguientes: en la primera el estudiante responde correctamente las dos primeras preguntas, mientras que en la segunda responde correctamente las tres primeras preguntas. Cualquiera de las variables dependientes propuestas en el Modelo 1, dará un decaimiento (débilmente) mayor en la primera situación que en la segunda. En cambio, al usar el Modelo 2 incluyendo únicamente Q_{ij} como variable explicativa, el decaimiento estará representado por el coeficiente de la regresión simple entre Q_{ij} e Y_{ij} obteniendo un mayor decaimiento en la segunda situación[13].

El Modelo 2 calcula el decaimiento a partir de una agregación de las variaciones de la puntuación en cada uno de los momentos del test. La principal ventaja de este modelo consiste en evitar que el cálculo del decaimiento se vea influenciado por algunos casos extremos de estudiantes que lo hacen particularmente bien

---

11. En este modelo al inicio de la prueba (Q_{ij}=0) significa la primera pregunta. Los resultados no cambian si en su lugar se considera como parte inicial del test (Q_{ij}=0), las preguntas situadas en el 5% o 10% inicial de la prueba.
12. La prueba PISA tiene una duración de dos horas. A los estudiantes se les permite un pequeño descanso después de la primera hora.
13. En este ejemplo para la primera situación se tendría Y_{ij}=(1,1,0,0,0,0) e Y_{ij}=(1,1,1,0,0,0) en la segunda. Al estimar el coeficiente de Q se obtiene -0,95 en el primer caso y -1,07 en el segundo. Si con el modelo 1, se usara por ejemplo la diferencia con el 20% inicial del test completado y el total; en la primera situación el decaimiento es -2, mientras en la segunda es -1.

o particularmente mal. Si en el Modelo 1 se usara una parte pequeña de la prueba para medir el resultado inicial es fácil estar subestimando el decaimiento de todos aquellos estudiantes que lo han hecho mal en alguna de las preguntas iniciales. Si como caso extremo se tuviera en cuenta la tasa de decaimiento entre la puntuación en la primera pregunta y la puntuación media en el total de la prueba, todos los estudiantes que tuvieran un cero en la primera pregunta pero hubiesen respondido correctamente alguna de las demás preguntas presentarían crecimiento del 100% en su resultado, independientemente del número de respuestas correctas que haya conseguido. Al no tener en cuenta el decaimiento individualizado de los estudiantes, el Modelo 2 evita este problema[14]. En segundo lugar, el Modelo 2 hace un uso más intensivo de los datos. Mientras con el enfoque del Modelo 1 se disponía de 25.887 observaciones correspondientes a cada uno de los estudiantes, ahora habrá esta cantidad multiplicada por el número de respuestas de cada estudiante lo cual significa 1.462.622 observaciones.

Dadas las ventajas e inconvenientes de cada uno de los modelos valdrá la pena tener en cuenta ambos procedimientos a la hora de analizar la influencia en el decaimiento. La Tabla 4 muestra los coeficientes de las características personales y familiares en cada uno de los dos modelos. A pesar de las diferencias antes comentadas, ambos modelos llegan a resultados parecidos.

Se observa como algunas variables resultan significativas en todas las especificaciones. En concreto, ser chica, tener una situación socioeconómica favorable y hacer el test en la lengua materna suponen un menor decaimiento para cualquier modelo utilizado.

También se puede ver como algunas variables que resultaban significativas en la puntuación de la prueba PISA, no lo son a la hora de explicar el decaimiento del rendimiento. Exceptuando la especificación del Modelo 1 que usa la diferencia entre el 50% inicial de la prueba y el 50% final, se observa que el mes de nacimiento y tener una familia nuclear nunca resultan significativas. Además haber cursado más de un año de educación infantil, solo resulta significativa en una de las especificaciones y únicamente teniendo en cuenta un nivel de significación del 10%.

Para las demás variables conviene ser cautelosos al analizarlas dado que su significatividad depende del modelo y la especificación que se use. Aun así, y dadas las características anteriormente comentadas de cada uno de los modelos, los resultados parecen indicar que ser inmigrante resultaría en un mayor decaimiento pero concentrado en la parte final de la prueba (ya que no resulta significativo usando las

14. Obviamente, con el Modelo 1 se tienen en cuenta mayores porcentajes al inicio de la prueba para mitigar este problema. Esto no descarta que el efecto ilustrado por este caso extremo pueda afectar el resultado.

TABLA 4. DECAIMIENTO EN EL RENDIMIENTO Y CARACTERÍSTICAS DE LOS ESTUDIANTES

	MODELO 1	MODELO 1	MODELO 1	MODELO 1	MODELO 1	MODELO 1	MODELO 2
	5%-100%	10%-100%	20%-100%	50%-50%	5%-20%	10%-20%	Y_{ij}
V. Explicativa							
Chica	0,0557***	0,0441***	0,0349***	0,0442***	0,0736***	0,0601***	0,0233***
	(0,00854)	(0,00585)	(0,00406)	(0,00825)	(0,01005)	(0,00784)	(0,00260)
ESCS	0,00872*	0,0127***	0,00586**	0,0174***	0,0200***	0,0235***	0,00400***
	(0,00479)	(0,00328)	(0,00228)	(0,00464)	(0,00564)	(0,00440)	(0,00146)
Ed. Infantil	0,00807	0,00503	0,00473	0,0361***	0,0230	0,0201*	0,000916
	(0,01265)	(0,00865)	(0,00601)	(0,01223)	(0,01494)	(0,01166)	(0,00386)
Mes	-0,00107	-0,000467	-0,0000851	-0,000547	-0,00124	-0,000603	0,0000159
	(0,00125)	(0,00085)	(0,00059)	(0,00121)	(0,00147)	(0,00115)	(0,00038)
Idioma	-0,0342**	-0,0250***	-0,0300***	-0,0284**	-0,0569***	-0,0479***	-0,0175***
	(0,01403)	(0,00960)	(0,00667)	(0,01357)	(0,01659)	(0,01295)	(0,00429)
Libros	0,00599	0,00787***	0,00702***	0,0243***	0,0147***	0,0162***	0,00114
	(0,00374)	(0,00256)	(0,00177)	(0,00361)	(0,00439)	(0,00343)	(0,00114)
Interrup-ciones	-0,0116**	-0,00780**	-0,00140	-0,0179***	-0,0174***	-0,0134***	-0,00201
	(0,00485)	(0,00332)	(0,00231)	(0,00469)	(0,00570)	(0,00445)	(0,00147)
Fam. nuclear	-0,0121	-0,000667	0,00569	0,0248**	-0,00668	0,00255	-0,00351
	(0,01230)	(0,00841)	(0,00584)	(0,01191)	(0,01456)	(0,01136)	(0,00375)
Inmi-grante	-0,00990	-0,0187*	-0,00790	-0,0852***	-0,0390**	-0,0567***	-0,0188***
	(0,01656)	(0,01133)	(0,00787)	(0,01606)	(0,01987)	(0,01551)	(0,00512)
Observa-ciones	24.212	24.212	24.212	24.036	23.265	23.265	1.375.207

Errores estándar entre paréntesis. * p<0,1 / ** p<0,05 / *** p<0,01.

especificaciones que comparan el resultado al principio con el total de la prueba en el Modelo 1 pero si resulta significativo cuando se considera el Modelo 2 o bien las especificaciones del Modelo 1 que dan un mayor peso al resultado en la parte final). Finalmente, hay resultados poco concluyentes acerca de cómo afectan las interrupciones en clase o el número de libros en el decaimiento. Aunque algunas especificaciones del Modelo 1 muestran que las interrupciones en clase pueden aumentar el

decaimiento de los estudiantes y el número de libros en casa reducirlo, en otras especificaciones del Modelo 1 y en el Modelo 2, estas no resultan significativas.

Para evaluar el tamaño de estos efectos se considera el Modelo 2, dado que facilita la interpretación. Usando este modelo, y tal como se verá en la Sección 6, el decaimiento medio en España es de unos 10 puntos porcentuales. Esta diferencia es parecida a la diferencia en el porcentaje de respuestas correctas entre España y Finlandia. Además, esta diferencia se reduce en más de dos puntos para las chicas y aumenta en casi dos para aquellos estudiantes que no hacen el test en su lengua materna. La diferencia en el decaimiento del rendimiento entre la mejor situación socioeconómica y la peor es de 3,2 puntos. Finalmente, el Modelo 2 encuentra significativo el efecto de ser inmigrante resultando en un aumento en el decaimiento de casi 2 puntos.

HABILIDADES COGNITIVAS

Anteriormente se ha usado el decaimiento en el rendimiento como una medida de las habilidades no cognitivas. Al mismo tiempo, dado que el resultado al principio del test no se ve afectado por el decaimiento en el rendimiento, este proporcionará una medida más precisa de las habilidades cognitivas, es decir una medida menos influenciada por las habilidades no cognitivas que se manifiestan a través del decaimiento en el rendimiento.

El Modelo 2 permite hacer directamente esta distinción. Como se ha explicado, los coeficientes de las interacciones entre las características personales y la posición relativa de la pregunta estarán midiendo la incidencia de la característica en cuestión sobre el decaimiento. En cambio, los coeficientes de las variables sin interaccionar, estarán mostrando el caso $Q_{ij}=0$, y por lo tanto medirá el efecto sobre el resultado al principio del test.

Para hacer esto mediante el Modelo 1, se debe plantear una nueva regresión usando el porcentaje de respuestas correctas al principio del test como variable dependiente. Se presentarán los resultados obtenidos al tener en cuenta el 5%, 10% y 15% inicial de la prueba.

La Tabla 5 muestra los resultados para el Modelo 1 y 2 respectivamente. En estas tablas hay un dato que resulta llamativo. El sexo no tiene ningún impacto significativo en el resultado al principio de la prueba. Como se veía en la Tabla 3, las chicas obtienen una mayor puntuación que los chicos si se considera todo el test completado. Sin embargo, esto no es cierto al considerar únicamente el resultado al inicio de la prueba. Al mismo tiempo la proporción entre preguntas de comprensión lectora y las demás es parecida durante todo el test. Combinando estas obser-

vaciones se concluye que la mejor puntuación media de las chicas en PISA 2009 es consecuencia de su menor decaimiento en el rendimiento.

TABLA 5. PORCENTAJE DE RESPUESTAS CORRECTAS AL 5-10-15% INICIAL (MODELO 1) Y VARIABLES SIN INTERACTUAR (MODELO 2) Y CARACTERÍSTICAS DE LOS ESTUDIANTES

	MODELO 1	MODELO 1	MODELO 1	MODELO 2
% DEL TEST COMPLETADO	**5%**	**10%**	**15%**	**Y_{ij}**
Chica	-0,00493	-0,00171	-0,00166	-0,000846
	(0,00380)	(0,00290)	(0,00268)	(0,00149)
ESCS	0,0316***	0,0280***	0,0305***	0,0299***
	(0,00213)	(0,00163)	(0,00150)	(0,00083)
Ed.infantil	0,0411***	0,0392***	0,0404***	0,0385***
	(0,00563)	(0,00430)	(0,00397)	(0,00220)
Mes	-0,00198***	-0,00242***	-0,00263***	-0,00231***
	(0,00055)	(0,00042)	(0,00039)	(0,00022)
Idioma	-0,00924	-0,00847*	-0,00876**	-0,00968***
	(0,00624)	(0,00477)	(0,00440)	(0,00244)
Libros	0,0358***	0,0340***	0,0346***	0,0335***
	(0,00166)	(0,00127)	(0,00117)	(0,00065)
Interrupciones	-0,0113***	-0,0120***	-0,0131***	-0,0136***
	(0,00216)	(0,00165)	(0,00152)	(0,00084)
Fam.nuclear	0,0154***	0,0116***	0,0117***	0,0104***
	(0,00547)	(0,00418)	(0,00386)	(0,00214)
Inmigrante	-0,0436***	-0,0394***	-0,0418***	-0,0329***
	(0,00737)	(0,00562)	(0,00519)	(0,00289)
Observaciones:	24.215	24.215	24.215	1.375.207

Errores estándar entre paréntesis. * p<0,1 / ** p<0,05 / *** p<0,01.

Algo parecido podría estar sucediendo con el hecho de realizar la prueba en lengua materna, pero con resultados menos concluyentes. El Modelo 1, parece indicar que el idioma no es significativo al principio del test, en cambio sí resulta significativo en el Modelo 2. En cualquier caso resulta claro que la principal influencia del idioma ocurre como consecuencia de un mayor decaimiento. Esto encajaría si se piensa que el efecto del idioma puede implicar la necesidad de un mayor esfuerzo o tiempo en contestar cada pregunta. No obstante, en este caso no se puede hablar de habilidades no cognitivas sino más bien de la importancia de un dominio del lenguaje sobre el decaimiento en el rendimiento.

El estatus socioeconómico afecta tanto el resultado al principio del test (habilidades cognitivas) como al decaimiento en el rendimiento (habilidades no cognitivas). Sin embargo, a partir de una comparación entre los coeficientes estimados del estatus socioeconómico en las Tablas 4 y 5, se observa como la mayor parte del efecto tiene lugar al principio del test mientras que el efecto sobre el decaimiento es relativamente bajo. En cambio, ocurre lo contrario para las variables idioma y sexo, cuyo principal efecto tiene lugar a partir del decaimiento en el resultado.

Finalmente, se aprecia cómo las demás variables no siempre resultan significativas para explicar el decaimiento en el rendimiento y en cambio siempre lo son al principio de la prueba, lo cual indica que su influencia en el resultado final ocurre vía conocimientos. Este hecho se refleja también mediante el Modelo 1 al constatar que los coeficientes de estas variables son parecidos en tamaño en el 5%, 10% o 15% inicial del test (Tabla 5) y el total de la prueba completado (Tabla 3).

OTROS FACTORES RELACIONADOS CON EL DECAIMIENTO DEL RENDIMIENTO

A la vista que tan solo el sexo, ESCS y el idioma parecen tener un efecto importante y robusto a las distintas especificaciones del modelo para explicar el decaimiento en el rendimiento, se ampliará el vector de variables explicativas que se han usado en las estimaciones. Se incluirán características relacionadas con el centro de educativo y con la actitud de los estudiantes.

Convendrá ser prudentes a la hora de interpretar estos resultados ya que por un lado algunas de las variables que se incluyen dependen de respuestas subjetivas de los estudiantes como su relación con el profesor, su agrado de la lectura o el uso de tecnologías de la información. Esto supone que no se puedan dar resultados concluyentes sobre el sentido de la causalidad entre estas variables y el decaimiento del rendimiento.

En segundo lugar se tendrán en cuenta también datos procedentes de la encuesta a los centros educativos. Tal y como se explica en el capítulo de los profesores Calero y Escardíbul de este proyecto, la composición anidada de estos datos sugiere el uso de un modelo multinivel para la obtención de resultados más robustos sobre la influencia de estas variables.

Por lo tanto los resultados de esta sección deberán ser interpretados teniendo en cuenta estas dos consideraciones. Par ver el efecto de estas variables se usarán los Modelos 1 y 2 con la diferencia de que ahora el vector de características incluirá las variables detalladas en la Tabla 6.

En la Tabla 6 se muestran los coeficientes de las variables añadidas en esta sección. A parte de las nuevas variables explicativas se incluyen en la tabla aquellas variables que anteriormente habían resultado significativas, es decir, sexo, hacer el test en lengua materna y el estatus socioeconómico.

TABLA 6. DECAIMIENTO DEL RENDIMIENTO, ACTITUD DEL ESTUDIANTE Y CARACTERÍSTICAS DEL CENTRO EDUCATIVO

	MODELO 1	MODELO 1	MODELO 1	MODELO 1	MODELO 1	MODELO 2
	5%-100%	10%-100%	20%-100%	5%-20%	10%-20%	Y_{ij}
Prof. estimula lectura	0,00150	0,000130	0,000315	0,00264	0,000534	-0,000653
	(0,00518)	(0,00352)	(0,00245)	(0,00606)	(0,00472)	(0,00162)
Actitud en la escuela	-0,00252	0,00151	-0,00116	-0,000604	0,00366	-0,00235
	(0,00514)	(0,00349)	(0,00243)	(0,00602)	(0,00469)	(0,00161)
Uso de estrategias de control	0,00419	0,00521	0,00353	0,00817	0,00950*	-0,000822
	(0,00534)	(0,00362)	(0,00252)	(0,00625)	(0,00487)	(0,00167)
Gusto por la lectura	0,0354***	0,0292***	0,0192***	0,0515***	0,0463***	0,0172***
	(0,00538)	(0,00365)	(0,00254)	(0,00627)	(0,00489)	(0,00167)
Estrategias de memorización	-0,000592	0,00345	0,00448**	0,00720	0,0117***	0,00383***
	(0,00455)	(0,00309)	(0,00215)	(0,00532)	(0,00414)	(0,00142)
Relación con profesor	0,00931*	0,00702**	0,00744***	0,0106*	0,00819*	0,00744***
	(0,00510)	(0,00346)	(0,00241)	(0,00598)	(0,00466)	(0,00159)
Tecnología de información	-0,0112**	-0,0105***	-0,0109***	-0,0228***	-0,0210***	-0,00781***
	(0,00499)	(0,00339)	(0,00236)	(0,00586)	(0,00456)	(0,00156)
Centro educ. equipo deportivo	-0,0429***	-0,0319***	-0,0248***	-0,0504***	-0,0390***	-0,0154***
	(0,01424)	(0,00966)	(0,00672)	(0,01664)	(0,01296)	(0,00445)
Freq. deberes	-0,00384	-0,00582	-0,00439	-0,00310	-0,00536	-0,00323*
	(0,00575)	(0,00390)	(0,00271)	(0,00669)	(0,00521)	(0,00179)
Bajas expectativas del profesor	0,00422	0,00357	0,00355	0,00270	0,00196	0,000806
	(0,00653)	(0,00443)	(0,00308)	(0,00764)	(0,00595)	(0,00204)
Relación profesor alumnos	0,00176	-0,00386	-0,00465	0,00856	0,00509	0,00140
	(0,00785)	(0,00533)	(0,00371)	(0,00920)	(0,00717)	(0,00245)

[CONTINÚA EN PÁGINA SIGUIENTE]

TABLA 6. DECAIMIENTO DEL RENDIMIENTO, ACTITUD DEL ESTUDIANTE Y CARACTERÍSTICAS DEL CENTRO EDUCATIVO

	MODELO 1	MODELO 1	MODELO 1	MODELO 1	MODELO 1	MODELO 2
	5%-100%	10%-100%	20%-100%	5%-20%	10%-20%	Y_{ij}
Absentismo profesor	-0,0122	-0,0111**	-0,00409	-0,0129	-0,0102	-0,00134
	(0,00809)	(0,00549)	(0,00382)	(0,00955)	(0,00744)	(0,00254)
Formación continua prof.	-0,00311	0,000162	0,00438	0,00204	0,00602	0,00215
	(0,00688)	(0,00467)	(0,00325)	(0,00806)	(0,00628)	(0,00215)
Separar grupos de habilidad	-0,00880	-0,00834	-0,00316	-0,0102	-0,00871	-0,00474*
	(0,00782)	(0,00531)	(0,00369)	(0,00915)	(0,00713)	(0,00244)
Proporción chicas	0,00109**	0,000927**	0,000728***	0,00173***	0,00161***	0,000424**
	(0,00056)	(0,00038)	(0,00026)	(0,00066)	(0,00051)	(0,00017)
Ratio alumnos-profesor	0,00347***	0,00284***	0,00175***	0,00593***	0,00557***	0,00201***
	(0,00108)	(0,00073)	(0,00051)	(0,00126)	(0,00098)	(0,00034)
Act. extraecolares	0,00320	0,00215	-0,00122	0,00609	0,00502	0,000849
	(0,00617)	(0,00419)	(0,00291)	(0,00719)	(0,00560)	(0,00192)
Chica	0,0354***	0,0205***	0,0177***	0,0412***	0,0235***	0,0128***
	(0,00951)	(0,00645)	(0,00449)	(0,01114)	(0,00868)	(0,00297)
ESCS	0,000422	0,00351	0,000199	0,00350	0,00658	-0,000737
	(0,00525)	(0,00357)	(0,00248)	(0,00615)	(0,00479)	(0,00164)
Idioma	-0,0376**	-0,0272***	-0,0265***	-0,0534***	-0,0414***	-0,0128***
	(0,01534)	(0,01041)	(0,00724)	(0,01804)	(0,01405)	(0,00482)
Observaciones	20.030	20.030	20.030	19.303	19.303	1.139.659

Errores estándar entre paréntesis. * p<0,1 / ** p<0,05 / *** p<0,01.

Se observa como el estatus socioeconómico deja de ser significativo al introducir estas nuevas variables. Esto reforzaría el resultado anterior donde se veía que el efecto del estatus socioeconómico se produce principalmente al inicio del test. En cambio los efectos del sexo o el idioma no cambian demasiado con respecto a los resultados anteriores.

La mayoría de las variables usadas no tienen efectos significativos. Sin embargo hay algunas excepciones notables. En primer lugar se observa que aquellos es-

tudiantes que se declaran amantes de la lectura tienen un menor decaimiento. Este efecto es muy notable, de hasta 11 puntos al comparar el porcentaje de respuestas correctas de los estudiantes con un menor gusto por la lectura con aquellos más bibliófilos. No obstante y como se ha alertado anteriormente, no se puede establecer de manera rotunda cual es la dirección causal entre estas dos variables. Si la relación causal fuera que el gusto por la lectura favorece un menor decaimiento en el resultado, entonces se debería tener en cuenta el fomento de la lectura como vía de mejora del resultado de los estudiantes. Con todo, no se puede descartar que en realidad la misma habilidad que provoca un menor decaimiento sea responsable de un mayor interés por la lectura.

Algo parecido ocurre con la variable que recoge la buena relación entre estudiante y profesor aunque en este caso el efecto es mucho menos importante que en el caso anterior.

Se observa también como dos de las variables añadidas agravan el problema del decaimiento en el rendimiento. El primer caso es el uso de tecnologías de la información. Parece contradictorio que el uso de más medios, pueda resultar en un prejuicio en el rendimiento del estudiante. Sin embargo, aunque todavía con una investigación muy incipiente al respecto, existe una creciente atención sobre la posibilidad que las nuevas tecnologías reduzcan la capacidad de atención de sus usuarios, Purcell *et al.* (2012). En segundo lugar la existencia de un equipo deportivo en el centro también puede provocar mayores niveles de decaimiento.

Este análisis no permite extraer conclusiones sobre el efecto que el número de estudiantes por profesor tiene sobre el decaimiento. Tal y como se advertía en el capítulo de los profesores Calero y Escardíbul existe un problema de endogeneidad entre el tamaño del grupo y sus resultados. Un mejor grupo o un mejor profesor pueden provocar un efecto llamada, aumentando así el tamaño del grupo o bien la necesidad de asignar un mayor número de profesores a aquellos estudiantes con resultados especialmente malos, pueden provocar que se concentren malos resultados en grupos pequeños. Según estos autores, esta endogeneidad provocaba que contrariamente a lo que cabría esperar, se observara una relación negativa entre el tamaño del grupo y los resultados PISA. Lo mismo ocurre cuando se evalúa la relación entre el número de estudiantes y el decaimiento del rendimiento, al observar que existe un menor decaimiento cuando el tamaño del grupo es mayor.

Finalmente, la existencia de una mayor proporción de chicas favorece que haya menores niveles de decaimiento. Esto sugiere que los efectos de pares positivos de chicas a chicos documentados en Ciccone y García-Fontes (2013), podrían darse como consecuencia de una transmisión de las habilidades no cognitivas[15]. Sin embar-

15. Esta hipótesis se ve reforzado por el hecho que al tomar regresiones separadas entre chicos y chicas, la pro-

go para extraer conclusiones robustas en este aspecto convendría llevar a cabo análisis más exhaustivos basados en cohortes como hacen Ciccone y García-Fontes (2013).

DIFERENCIAS ENTRE COMUNIDADES AUTÓNOMAS

Una característica que llama la atención de los resultados de las pruebas PISA en España es la notable disparidad existente en los resultados obtenidos por las distintas Comunidades Autónomas. Además, muchas de estas diferencias no pueden ser explicadas por las diferencias en las características personales o familiares de los estudiantes, Villar (2012).

Se pueden analizar las diferencias entre comunidades teniendo en cuenta la distinción entre resultado al principio de la prueba y el decaimiento. Se usarán los coeficientes de las variables ficticias de cada una de las comunidades en el Modelo 2 para mostrar las diferencias entre comunidades al principio de la prueba y en el decaimiento del rendimiento antes y después de controlar por las características individuales y familiares de los estudiantes. Los intervalos de confianza de estos indicadores permitirán ver si las diferencias entre comunidades existen también en el resultado al principio de la prueba y en el decaimiento.

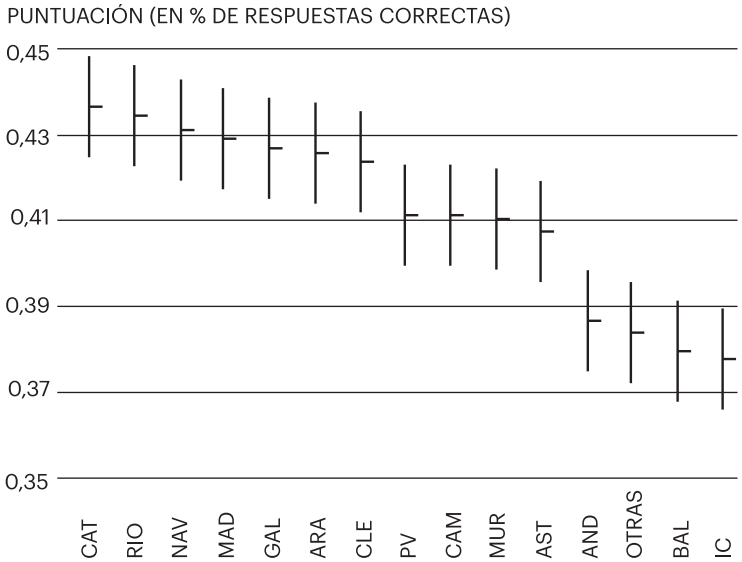
En Figura 2[16], se muestra el porcentaje medio de respuestas correctas, el porcentaje medio de respuestas correctas al inicio de la prueba y el decaimiento de cada una de las comunidades sin controlar por las características personales y familiares de los estudiantes. Por un lado se observa que las diferencias en el porcentaje de respuestas correctas son elevadas. Entre la mejor clasificada, Castilla y León, y una de las peores, Islas Canarias, la diferencia en el porcentaje de respuestas correctas es superior a la existente entre España y Finlandia[17].

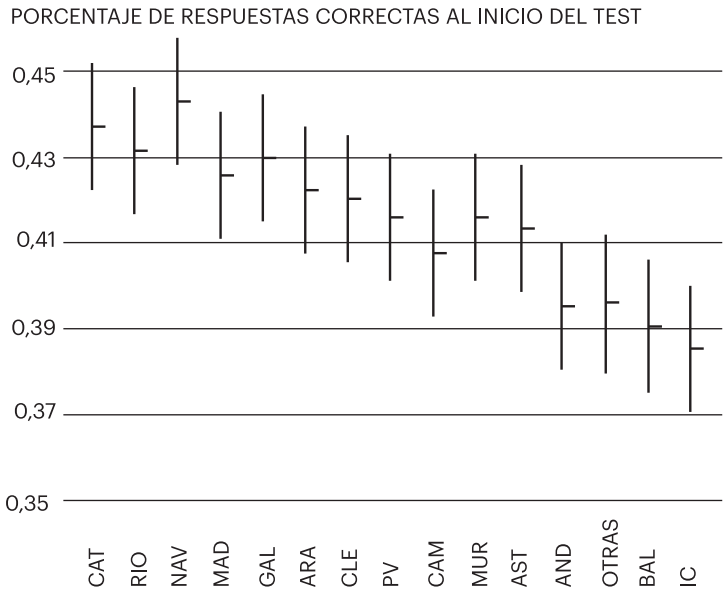
Por otro lado, existen diferencias significativas entre algunas comunidades tanto en el porcentaje de respuestas correctas al principio como en el decaimiento del rendimiento. La diferencia en el resultado al inicio del test puede variar hasta en 10 puntos entre algunas comunidades, mientras que el decaimiento lo puede hacer hasta en 4 puntos, como se observa al comparar Castilla y León con Baleares (cabe recordar que cuanto más negativo sea el valor del decaimiento de una comunidad mayor es su decaimiento). La comparación entre comunida-

porción de chicas sólo tiene un efecto significativo en el decaimiento de los chicos mientras resulta no significativa sobre el decaimiento de las chicas.

16. Las ciudades autónomas de Ceuta y Melilla no han sido incluídas en algunos de los gráficos para facilitar una elaboración más compacta de las figuras.
17. Algo parecido ocurre cuando se tiene en cuenta la puntuación PISA en lectura, matemáticas y ciencias en lugar del porcentaje total de respuestas correctas.

FIGURA 2. DIFERENCIAS ENTRE COMUNIDADES EN EL PORCENTAJE TOTAL DE RESPUESTAS CORRECTAS, PORCENTAJE DE RESPUESTAS CORRECTAS AL INICIO DE LA PRUEBA Y DECAIMIENTO DEL RENDIMIENTO. ELABORACIÓN PROPIA A PARTIR DE LOS MICRODATOS PISA 2009

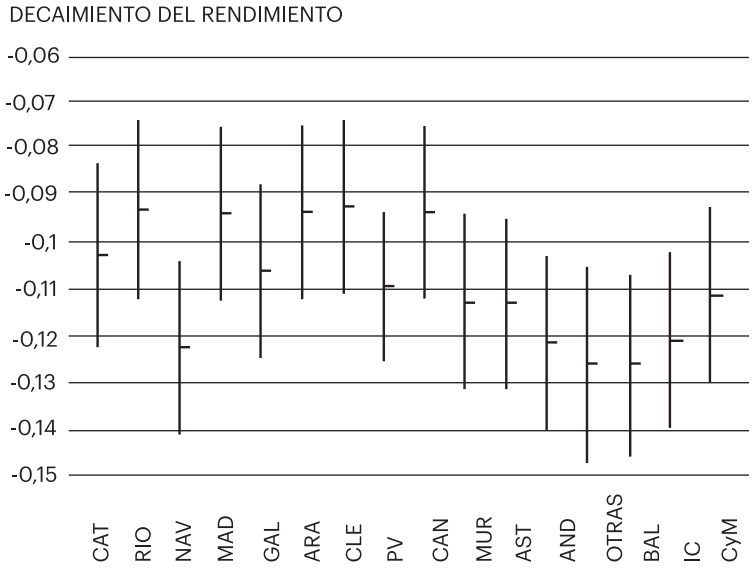




Nota: Las líneas verticales representan intervalos de confianza del 95% de cada valor.

[CONTINÚA EN PÁGINA SIGUIENTE]

FIGURA 2. DIFERENCIAS ENTRE COMUNIDADES EN EL PORCENTAJE TOTAL DE RESPUESTAS CORRECTAS, PORCENTAJE DE RESPUESTAS CORRECTAS AL INICIO DE LA PRUEBA Y DECAIMIENTO DEL RENDIMIENTO. ELABORACIÓN PROPIA A PARTIR DE LOS MICRODATOS PISA 2009



Nota: Las líneas verticales representan intervalos de confianza del 95% de cada valor.

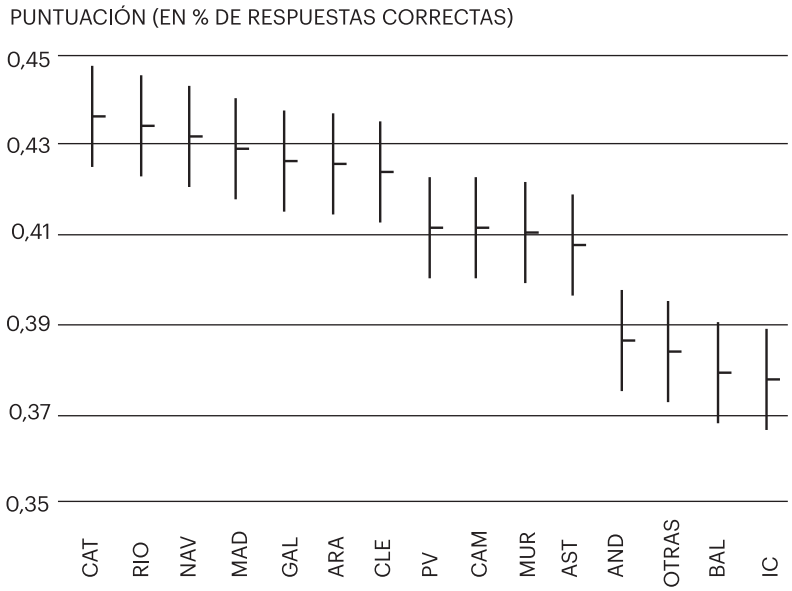
des difiere notablemente según la variable que se esté considerando. Por ejemplo, Castilla y León que tiene un porcentaje de respuestas correctas significativamente superior a todas las comunidades clasificadas por debajo de la sexta posición, deja de presentar diferencias significativas con las comunidades de País Vasco, Asturias y Navarra cuando se considera solo el resultado al inicio de la prueba. En cambio, se observa como precisamente estas comunidades, muestran niveles de decaimiento significativamente superiores a los de Castilla y León. Por lo tanto, la diferencia entre el porcentaje total de respuestas correctas entre estas comunidades se explica por diferencias en el decaimiento. En particular el decaimiento en el rendimiento es de especial relevancia para explicar el buen resultado de las Comunidades Autónomas a la cabeza de la clasificación (Castilla y León, Aragón, La Rioja y Madrid).

Se observa también que aunque en muchos casos las diferencias no resultan significativas el resultado al principio de la prueba ordena las comunidades de una forma más parecida al orden obtenido al tener en cuenta el porcentaje total de respuestas correctas. Por lo tanto esto indica que el resultado al principio de la prueba tiene más peso a la hora de explicar la clasificación entre las distintas comunidades.

Sin embargo y dado la importancia de las características familiares y personales, valdrá la pena analizar estas diferencias entre comunidades controlando estos aspectos. La Figura 3 muestra las diferencias entre comunidades después de controlar por las características personales.

Llama la atención que la clasificación de las distintas comunidades basada en el porcentaje medio de respuestas correctas, cambia notablemente cuando se tienen en cuenta las características personales de los estudiantes. Sin embargo cabe señalar que las Comunidades Autónomas que intercambian sus posiciones no presentan diferencias significativas ni antes ni después de usar los controles, por lo tanto este resultado es estadísticamente irrelevante. Las diferencias encontradas en el porcentaje de respuestas correctas entre comunidades siguen siendo considerables después de controlar por las características de los estudiantes llegando a ser de hasta casi 6 puntos.

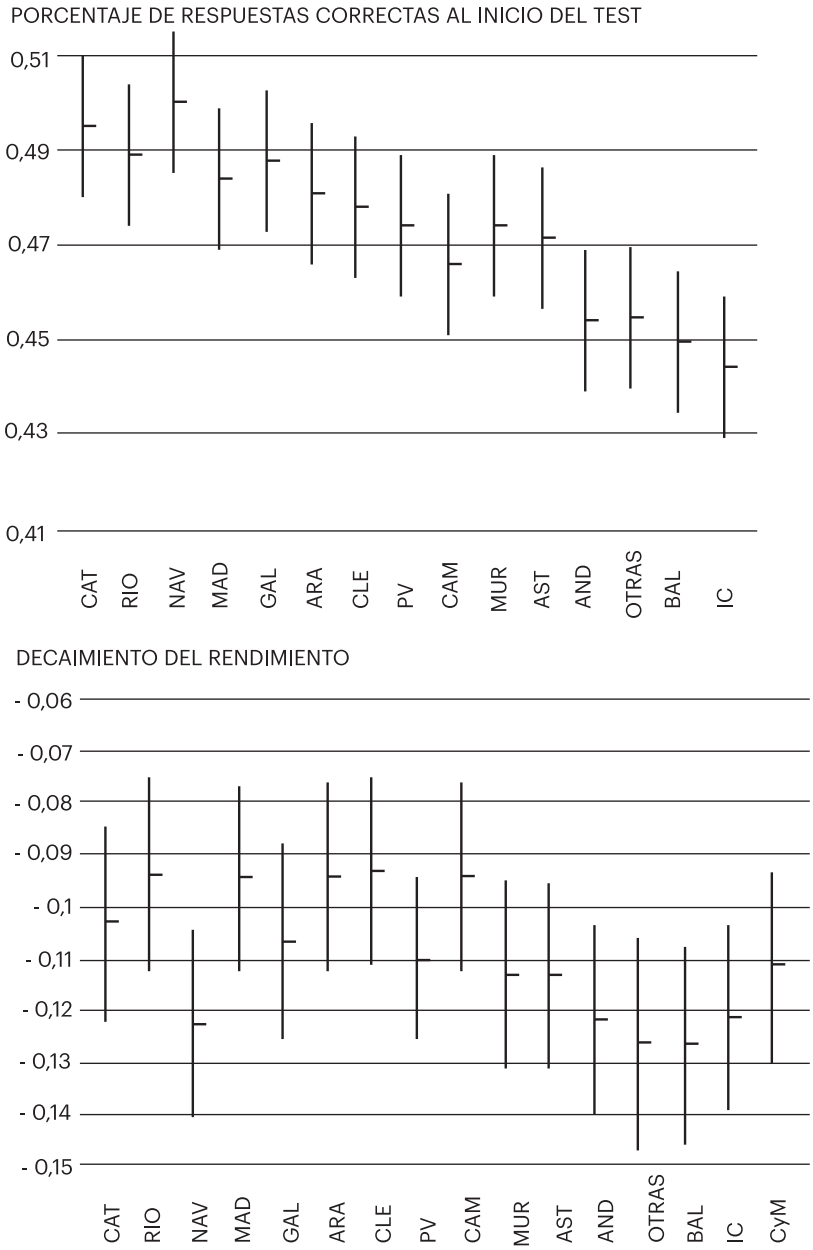
FIGURA 3. DIFERENCIAS ENTRE COMUNIDADES EN EL PORCENTAJE TOTAL DE RESPUESTAS CORRECTAS, PORCENTAJE DE RESPUESTAS CORRECTAS AL INICIO DE LA PRUEBA Y DECAIMIENTO DEL RENDIMIENTO DESPUÉS DE CONTROLAS POR CARACTERÍSTICAS DEL ESTUDIANTE. ELABORACIÓN PROPIA A PARTIR DE LOS MICRODATOS PISA 2009



Nota: Las líneas verticales representan intervalos de confianza del 95% de cada valor.

[CONTINÚA EN PÁGINA SIGUIENTE]

FIGURA 3. DIFERENCIAS ENTRE COMUNIDADES EN EL PORCENTAJE TOTAL DE RESPUESTAS CORRECTAS, PORCENTAJE DE RESPUESTAS CORRECTAS AL INICIO DE LA PRUEBA Y DECAIMIENTO DEL RENDIMIENTO DESPUÉS DE CONTROLAS POR CARACTERÍSTICAS DEL ESTUDIANTE. ELABORACIÓN PROPIA A PARTIR DE LOS MICRODATOS PISA 2009



Nota: Las líneas verticales representan intervalos de confianza del 95% de cada valor.

Hay un segundo resultado que conviene destacar. Una vez controladas las características personales, desaparece cualquier diferencia significativa en términos de decaimiento entre las comunidades. Por lo tanto, las diferencias en el decaimiento en el rendimiento no aportan información adicional sobre los resultados de las distintas comunidades que no se pueda explicar por la composición de la población de estudiantes en función de sus características personales y familiares[18]. Estos resultados no cambian si se permite que la influencia de los controles sea distinta en cada una de las comunidades, tomando regresiones separadas para cada comunidad.

El alto nivel de decaimiento en el rendimiento es un problema generalizado en el conjunto de regiones españolas una vez se han tenido en cuenta aspectos alejados del ámbito de la política educativa como son el sexo, el origen o el estatus socioeconómico de los estudiantes.

CONCLUSIONES

En este trabajo se ha utilizado la muestra española del estudio PISA 2009 para distinguir entre la puntuación en el test de los estudiantes y su decaimiento en el rendimiento. Tal y como muestran Borghans y Schils (2012) en una comparativa internacional, los malos resultados que obtiene España en las pruebas PISA son en gran parte consecuencia de su alto decaimiento en el rendimiento.

Estos resultados, sugieren la posibilidad que una de las debilidades de los estudiantes españoles esté relacionada con sus habilidades no cognitivas como podrían ser el esfuerzo, la motivación o la perseverancia.

Existe una creciente atención sobre la importancia de la personalidad y las habilidades no cognitivas en el futuro personal y profesional de los estudiantes, véase entre otros Levin (2012). Algunos estudios relacionan ciertos rasgos de la personalidad con unos mejores resultados académicos (Cunha, Heckman y Schennach, 2010; Heckman *et al.*, 2011; Heckman, Stixrud y Urzua, 2006; Lleras, 2008: Poropat, 2009); laborales (Barrick y Mount, 1991; Drago, 2011; Gallo *et al.*, 2003; Hogan y Holland, 2003; Judge y Hurst, 2007; Nyhus y Pons, 2005; Salgado,1997) o de salud (Conti y Heckman, 2010; Hampson *et al.*, 2007; Roberts *et al.*, 2007). Paralelamente, se pueden encontrar numerosos trabajos que resaltan la educación y las intervenciones tempranas como piezas clave en el desarrollo de rasgos de personalidad y las

18. Borghans y Schils (2012) realizan una comparativa del decaimiento del rendimiento de distintos países sin controlar por las características de los estudiantes. Al incluir estos controles en su comparativa se puede ver cómo, aunque se reducen considerablemente, siguen habiendo diferencias significativas entre algunos países.

habilidades no cognitivas de los individuos, véase Cunha *et al.* (2006) o Knudsen *et al.* (2006) entre otros. Además, como muestran Heckman *et al.* (2012) a través de la evaluación del *Perry Preschool Program*, las importantes mejoras de los resultados obtenidos a través de educación temprana en carácter ocurren principalmente como consecuencia de una mejora en las habilidades no cognitivas de los individuos[19]. Tal y como muestran estos autores la inversión en la mejora de las habilidades no cognitivas de los individuos puede comportar beneficios enormes y deberá ser muy tenida en cuenta a la hora de intentar mejorar los resultados educativos de cualquier país.

En este capítulo se ha considerado el decaimiento en el rendimiento como indicador de las habilidades no cognitivas y se ha analizado qué características personales o familiares y de los centros educativos pueden explicar diferencias en el decaimiento del rendimiento de los estudiantes españoles.

Se aprecia como muchas de las variables que afectan al resultado en las prueba PISA de los estudiantes no afectan necesariamente el decaimiento de su rendimiento como por ejemplo las interrupciones en clase, el mes de nacimiento o el número de libros en casa. La educación infantil tampoco supone un menor decaimiento en el rendimiento. Este resultado, es un tanto sorprendente dado la conocida incidencia de las intervenciones tempranas en el desarrollo de las habilidades no cognitivas. Sin embargo, no se puede descartar que esto ocurra a través de habilidades no cognitivas distintas a las que se manifiestan a través del decaimiento en el rendimiento. Por otro lado, el estatus socioeconómico de los estudiantes incide tanto en la puntuación de los estudiantes al inicio de la prueba como en el decaimiento de su rendimiento, aunque la mayor parte de este efecto ocurre a través del resultado al inicio de la prueba. En cambio sexo o hacer el test en lengua materna afectan al resultado de los estudiantes, principalmente como consecuencia de diferencias en el decaimiento del rendimiento.

Haciendo una evaluación preliminar y que requerirá de investigación futura se puede analizar el efecto de características susceptibles de ser modificadas a través de la política educativa. Así, el gusto por la lectura está fuertemente correlacionado con un menor decaimiento en el rendimiento. Esto sugiere que el fomento de la lectura podría ser una buena medida contra el elevado decaimiento en el resulta-

19. El *Perry Preschool Program* consiste en una intervención experimental llevada a cabo en Estados Unidos en niños y niñas afroamericanos de entre tres y cuatro años de edad con capacidades cognitivas por debajo de la media. Tenía por objetivo mejorar el entorno social y emocional de niños y niñas, obteniendo como resultado notables mejoras académicas. El uso de métodos aleatorios en la asignación de este programa permite concluir que las mejoras en el resultado académico de los receptores del programa no ocurre por diferencias en su capacidad cognitiva (medida mediante tests de inteligencia) sino como consecuencia de una potenciación de sus habilidades no cognitivas.

do de nuestros estudiantes, aunque queda por verificar la dirección causal de esta relación. En segundo lugar, parece que parte de los efectos de pares positivos que reciben los chicos mediante la coeducación están relacionados con las habilidades no cognitivas ya que una mayor proporción de chicas reduce el decaimiento de los chicos. Por otra parte, el uso de tecnologías de la información, parece tener un efecto negativo sobre el decaimiento del rendimiento, lo cual va en la línea de algunas advertencias recientes sobre los posibles problemas de atención que podría provocar un (mal) uso de la tecnología. En este sentido, sería conveniente tener en cuenta los efectos sobre el rendimiento en el decaimiento a la hora de evaluar la efectividad de programas como "Escuela 2.0" llevado a cabo entre los años 2009 y 2011.

Finalmente se observa como existen diferencias en el decaimiento del rendimiento de las distintas Comunidades Autónomas. Sin embargo estas diferencias se explican completamente por las diferencias en la composición por características personales de sus estudiantes. De este modo, el alto nivel de decaimiento en el rendimiento es un problema generalizado en el conjunto del país que deberá ser tenido en cuenta con el objetivo de mejorar el rendimiento educativo de los estudiantes.

REFERENCIAS

Almlund, M., Duckworth, A. L., Heckman, J. J. y Kautz, T. D. (2011). *Personality psychology and economics* (No. w16822). National Bureau of Economic Research.

Anghel, B. y Cabrales, A. (2010). "The Determinants of Success in Primary Education in Spain". *Documentos de trabajo* (FEDEA), núm. 20, pp. 1-65.

Barrick, M. R. y Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, *44* (1), 1-26.

Borghans, L. y T. Schils (2012), "The leaning tower of Pisa: decomposing achievement test scores into cognitive and noncognitive components", *Unpublished manuscript*.

Ciccone, A y García-Fontes, W. (2013). "Efectos de género en las escuelas, un enfoque basado en cohortes de edad" en *La educación en España. Una visión académica. Monografía de la Fundación de Estudios de Economía Aplicada*.

Calero, J. y Escardíbul, J. O. (2007). "Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003". *Hacienda Pública Española*, vol. 183 (4), pp. 33-66.

Conti, G. y Heckman, J. J. (2010). "Understanding the Early Origins of the Education-Health Gradient A Framework That Can Also Be Applied to Analyze Gene-Environment Interactions". *Perspectives on Psychological Science*, 5(5), 585-605.

Cornwell, Christopher M., David B. Mustard y Jessica Van Parys. 2009. "How Does theNew SAT Predict Academic Achievement in College?" WorkingPaper. University of Georgia.

Cunha, F., Heckman, J. J., Lochner, L. y Masterov, D. V. (2006). "Interpreting the evidence on life cycle skill formation". *Handbook of the Economics of Education*, 1, 697-812.

Cunha, F., Heckman, J. J. y Schennach, S. M. (2010). "Estimating the technology of cognitive and noncognitive skill formation". *Econometrica*, 78(3), 883-931.

Drago, F. (2011). "Self-esteem and earnings". *Journal of Economic Psychology*, 32(3), 480-488.

Duckworth, A. L., Quinn, P. D., Lynam, D., Loeber, R., Stouthamer-Loeber, M., Moffit, T. E. y Caspi, A. (2009). "What intelligence tests test: Individual differences in test motivation and IQ". Unpublished Manuscript. University of Pennsylvania.

Duckworth, A. L. y Seligman, M. E. (2006). "Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores". *Journal of educational psychology*, 98(1), 198.

Eklöf, H. (2007). "Test-taking motivation and mathematics performance in TIMSS 2003". *International Journal of Testing*, 7(3), 311-326.

Gallo, W. T., Endrass, J., Bradley, E. H., Hell, D. y Kasl, S. V. (2003). "The influence of internal control on the employment status of German workers". *Schmollers Jahrbuch*, 123(1), 71-81.

Hanushek, E. A. y Woessmann, L. (2011). "How much do educational outcomes matter in OCDE countries?" *Economic Policy*, 26(67), 427-491.

Hanushek, E. A. y Woessmann, L. (2012). "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation". *Journal of Economic Growth*, 17(4), 267-321.

Heckman, J. J. (2008). "Schools, skills, and synapses". *Economic inquiry*, 46(3), 289-324.

Heckman, J.J., Humphries, J.E., Mader, N., (2011) "The GED". En: Hanushek, E.A., Machin,S., y Wößmann, L.,(Ed.), *Handbook of the Economics of Education*. Elsevier, Amsterdam, pp. 423-484.

Hampson, S. E., Goldberg, L. R., Vogt, T. M. y Dubanoski, J. P. (2007). "Mechanisms by which childhood personality traits influence adult health status: educational attainment and healthy behaviors". *Health psychology*, 26(1), 121.

Heckman, JJ., Humphries, J.E., Urzua, S. y Veramendi, G.(2011). "The Effects of Schooling on Labor Market and Health Outcomes". *Unpublished Manuscript*. Department of Economics, University of Chicago.

Heckman, J. J., Pinto, R. y Savelyev, P. A. (2012). "Understanding the mechanisms

through which an influential early childhood program boosted adult outcomes" (No. w18581). *National Bureau of Economic Research.*

Heckman, JJ., Stixrud, J. y Urzua, S., (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics, vol.* 24(3), pages 411-482, July.

Hogan, J. y Holland, B. (2003). "Using theory to evaluate personality and job-performance relations: a socioanalytic perspective". *Journal of Applied Psychology*, 88(1), 100.

Judge, T. A. y Hurst, C. (2007). "Capitalizing on one's advantages: role of core self-evaluations". *Journal of Applied Psychology*, 92(5), 1212.

Knudsen, E. I., Heckman, J. J., Cameron, J. L. y Shonkoff, J. P. (2006). "Economic, neurobiological, and behavioral perspectives on building America's future workforce". *Proceedings of the National Academy of Sciences*, 103(27), 10155-10162.

Lavy, V. y Schlosser, A. (2007). "Mechanisms and impacts of gender peer effects at school (No. w13292)". *National Bureau of Economic Research.*

Levin, H. M. (2012). "More than just test scores". *Prospects*, 42(3), 269-284.

Lleras, C. (2008). "Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings". *Social Science Research*, 37(3), 888-902.

Nyhus, E. K. y Pons, E. (2005). "The effects of personality on earnings". *Journal of Economic Psychology*, 26(3), 363-384.

OCDE (2011), PISA 2009 at a Glance, OCDE Publishing.

OCDE (2012), PISA 2009 Technical Report, PISA, OCDE.

Poropat, A. E. (2009). "A meta-analysis of the five-factor model of personality and academic performance". *Psychological bulletin*, 135(2), 322.

Purcell, K., Rainie, L., Heaps, A., Buchanan, J., Friedrich, L., Jacklin, A., ... y Zickuhr, K. (2012). "How teens do research in the digital world". *Pew Research Center's Internet & American Life Project.*

Roberts, B. W. (2007). "Contextualizing personality psychology". *Journal of personality*, 75(6), 1071-1082.

Rosen, J. A., Glennie, E. J., Dalton, B. W., Lennon, J. M. y Bozick, R. N. (2010). "Non cognitive skills in the classroom: New perspectives on educational research". *RTI Press publication.*

Salgado, J. F. (1997). "The Five Factor Model of personality and job performance in the European Community". *Journal of Applied psychology*, 82(1), 30.

Segal, C. (2012). "Working when no one is watching: Motivation, test scores, and economic success". *Management Science*, 58(8), 1438-1457.

Villar, A. (2012). "Educación y desarrollo: PISA 2009 y el sistema educativo español". *Estudios BBVA*.

Wößmann, L. (2003). "Schooling resources, educational institutions and student performance: the international evidence". *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.

RECURSOS ESCOLARES Y RESULTADOS DE LA EDUCACIÓN

JORGE CALERO
J. ORIOL ESCARDÍBUL
Universidad de Barcelona e IEB

RESUMEN

En el presente artículo nos planteamos, como pregunta de investigación, cuál es el efecto de los recursos materiales y humanos de los centros sobre los resultados educativos. Esta pregunta resulta extremadamente relevante en los procesos de diseño y ejecución de las políticas educativas. En el artículo pretendemos evitar dos posiciones opuestas: en la primera de ellas se sostiene que la expansión de los recursos puede generar mejoras continuas en el rendimiento; la segunda se define por un escepticismo extremo con respecto a la capacidad de los recursos a la hora de proponer mejoras de rendimiento. La respuesta que nosotros proponemos, como más razonable en función del conocimiento actual, es que el efecto de los recursos de los centros educativos sobre los resultados depende de factores como el tipo de gasto, el nivel inicial de la inversión y el colectivo afectado por ésta.

En el artículo efectuamos una aproximación por dos vías a la pregunta propuesta. Por una parte efectuamos una revisión en profundidad de la evidencia empírica previa, que es muy extensa y cuyos resultados dependen en gran medida de las metodologías empleadas y del entorno de aplicación concreto. Por otra parte, presentamos una serie de análisis empíricos propios, basados en los datos de España de PISA. En ellos se explora la distribución de los recursos de los centros para diferentes grupos de alumnos y se identifican, mediante modelos de regresión multinivel, qué papel juegan los recursos educativos de los centros a la hora de determinar la adquisición de la competencia de lectura.

INTRODUCCIÓN

¿Pueden mejorarse los resultados de nuestro sistema educativo por medio de inversiones adicionales en recursos materiales y humanos? La respuesta a esta pregunta resulta crucial a la hora de diseñar las intervenciones de política educativa. Sin embargo, el conocimiento actual no nos permite formular una respuesta taxativa, un sí o un no. Más bien, como veremos a lo largo de este artículo, la respuesta más razonable actualmente sería un "depende". Ese "depende" se contrapone a dos visiones opuestas pero que provienen, a nuestro juicio, de una cierta ingenuidad. La primera de ellas, hegemónica hasta el Informe Coleman (Coleman *et al.*, 1966) y, en bastantes contextos hasta mucho después, sostiene que los recursos materiales y humanos son muy relevantes para los sistemas educativos y que su expansión puede generar mejoras continuas en el rendimiento. La segunda visión, que ha pasado a ser hegemónica durante la última década, da énfasis en cómo las inversiones adicionales en recursos materiales y humanos generan efectos nulos sobre el rendimiento, no permiten mejoras en la calidad del sistema cuando ésta es medida en términos de resultados (véase, por ejemplo, Hanushek, 2003).

En este artículo nos planteamos como objetivo explicar el por qué de ese "depende" que hemos adelantado como respuesta alternativa a las dos visiones que calificamos de "ingenuas". Veremos que el efecto de la inversión de recursos depende, al menos, de tres factores: el tipo de gasto (algunas inversiones específicas sí tienen efectos sobre el rendimiento); el nivel inicial de la inversión (cuando se parte de niveles bajos de inversión sí se generan efectos significativos) y, finalmente, el colectivo afectado por la inversión (algunos colectivos de alumnos desfavorecidos pueden mejorar su rendimiento gracias a la inversión en recursos específicos).

La respuesta a la pregunta inicial, por tanto, resulta extremadamente dependiente del contexto concreto en el que se lleve a cabo el análisis. Si analizamos el efecto de una política orientada a facilitar el acceso a la Educación Primaria a niños sin recursos en un país en desarrollo tendremos más probabilidades de encontrar una respuesta positiva que si analizamos el efecto de reducir la ratio alumnos/profesor de 22 a 19, de forma generalizada, en un país desarrollado.

Por otra parte, al plantearse nuestra pregunta de investigación, es preciso tener en cuenta el papel determinante que tienen los aspectos metodológicos al decantarse por una u otra respuesta. Los problemas de endogeneidad, por ejemplo, se presentan muy frecuentemente en los análisis empíricos sobre los efectos de los recursos. Las estrategias orientadas a superar estos problemas suelen ser muy exigentes en cuanto al tipo y calidad de los datos empíricos necesarios, por lo que la evidencia empírica de alta calidad y desprovista de problemas metodológicos es muy escasa y está referida a ámbitos muy específicos (Schlotter *et al.*, 2011).

En el caso concreto de España las evaluaciones rigurosas sobre los efectos de los programas de gasto educativo han sido, y siguen siendo, muy poco abundantes. Ello fue un problema en el periodo de expansión presupuestaria, hasta 2008: a las políticas educativas de las Comunidades Autónomas y la Administración central se fueron incorporando una diversidad de programas basados en los recursos materiales y humanos cuya eficacia a la hora de incrementar el rendimiento es todavía desconocida o muy poco conocida. Pero resulta, quizás, un problema todavía mayor en el periodo, más reciente, marcado por los recortes presupuestarios. En este periodo se han retirado recursos de diferentes actividades educativas, que afectaban a colectivos muy variados, sin que existiera un criterio estable que permitiera establecer "líneas rojas" que delimitaran ámbitos o personas especialmente merecedoras de protección. Adicionalmente, en ocasiones se han utilizado argumentos basados en la supuesta nula eficacia de los recursos materiales y humanos (argumentos "a la Hanushek") para restar importancia al efecto de los recortes, sin tener en cuenta al menos dos cuestiones. Por una parte, un efecto medio reducido puede ser compatible con un efecto más significativo pero localizado en un colectivo determinado, de estudiantes desaventajados, por ejemplo. Y, por otra parte, el reconocimiento de que puede existir un cierto margen dentro del cual puedan contraerse los recursos sin que resulten aparentes los efectos negativos sobre el rendimiento pero que, una vez superado un determinado umbral, la escasez de recursos puede traducirse en pérdidas significativas de rendimiento.

La aproximación que efectuamos en el artículo a nuestra pregunta inicial es doble, basada inicialmente en la evidencia empírica previa y, posteriormente, en un análisis empírico propio. Ello se proyecta sobre la estructura del resto del artículo, que es la siguiente. A continuación se presentan y discuten una serie de estudios referidos al efecto del gasto educativo y la inversión en recursos materiales y humanos en los centros educativos sobre el rendimiento de los estudiantes. En el siguiente apartado se expone un breve análisis empírico, en el que se utilizan los datos referidos a España de la evaluación de competencias de PISA para el año 2009, con dos vertientes: en primer lugar, exploramos la distribución de los recursos materiales y humanos entre diferentes colectivos de estudiantes. En segundo lugar presentamos los resultados de un modelo en el que se estiman los determinantes de la adquisición de competencias de lectura, prestándose especial atención al efecto, en el modelo, de las variables relativas a los recursos materiales y humanos. Un último apartado contiene las conclusiones.

EVIDENCIA EMPÍRICA SOBRE LOS EFECTOS DE LOS RECURSOS ECONÓMICOS, MATERIALES Y HUMANOS SOBRE EL RENDIMIENTO ACADÉMICO

Desde el pionero Informe Coleman para los Estados Unidos (véase Coleman *et al.*, 1966), buena parte de los estudios empíricos ha destacado el escaso efecto que los recursos educativos (medidos mediante variables como el gasto en educación, los recursos humanos o materiales) tienen en el logro escolar. En este sentido, Hanushek (2003) señala que, comparando los resultados de generaciones diferentes, y teniendo en cuenta el fuerte aumento de recursos del sistema educativo a lo largo del siglo XX, no parece, en principio, que dicho incremento haya servido para mejorar los resultados, de modo que su aportación a la calidad de la educación sería estéril. Ahora bien, dos elementos deben tenerse en consideración. Por un lado, como se expone más adelante, en la literatura existente se observan múltiples excepciones a la afirmación de que los recursos educativos no inciden en el rendimiento académico, tanto para el conjunto de individuos como, especialmente, para determinados colectivos (como los socialmente desfavorecidos). Por otro lado, debe tenerse en cuenta que en el proceso de crecimiento de recursos se ha producido la universalización de la escolarización, de modo que se ha tenido que atender a muchos individuos de entornos socioeconómicos desfavorables que, anteriormente, no serían en su mayoría escolarizados; por tanto, las comparaciones en el tiempo deben considerarse con suma cautela.

En este apartado se discuten las evidencias empíricas existentes en torno al efecto de diversos factores vinculados con los recursos escolares disponibles sobre el logro académico. Este último viene definido, fundamentalmente, por los resultados en un test de contenidos o competencias adquiridas (*proxies* del concepto de calidad de la educación), aunque en ocasiones puede también referirse a la reducción del abandono educativo prematuro o del fracaso escolar. La discusión se centra en los niveles educativos primarios y secundarios (con alguna referencia a la educación infantil y superior). En consecuencia, se excluyen del análisis las capacidades no cognitivas que pueden adquirirse en la escuela, a pesar de su relevancia en la trayectoria educativa y laboral de los individuos (Heckman, 2008). Asimismo, la revisión de la literatura se realiza mayoritariamente desde la Economía de la Educación, motivo por el que predominan estudios basados en métodos de regresión en torno a la función de producción educativa, que relaciona los recursos educativos citados anteriormente (denominados *inputs*) con el rendimiento académico (*output*). Se presta especial atención a los trabajos más recientes, que han intentado resolver los problemas que las investigaciones con regresiones generan, como la endogeneidad de la variable dependiente, y que han aplicado técnicas de análisis

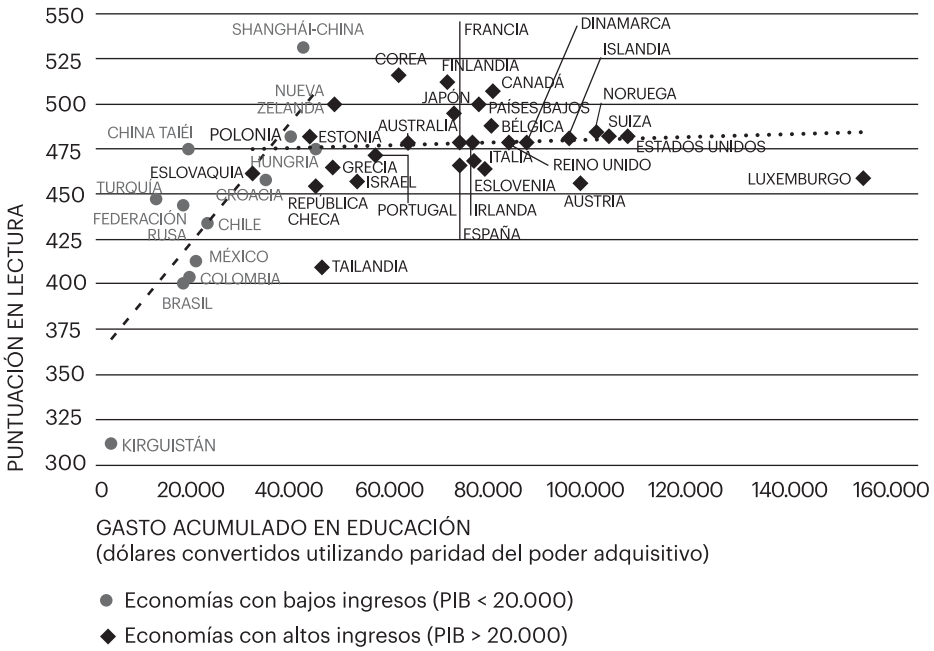
más avanzadas (véase, por ejemplo, una revisión de metodologías cuasiexperimentales en Schlotter *et al.*, 2011).

Se consideran, en primer lugar, una serie de aportaciones que relacionan el gasto en educación con el logro educativo. En segundo lugar, se analiza el papel de los recursos físicos; finalmente, se tienen en cuenta los recursos humanos, básicamente el profesorado, aunque también la dirección de los centros. Los trabajos discutidos no sólo se refieren al alumnado medio sino también a determinados colectivos socialmente desfavorecidos.

GASTO EDUCATIVO

Los resultados de las distintas evaluaciones sobre la adquisición de competencias de los estudiantes, desarrolladas por la OCDE en el *Programme for International Student Assessment* (PISA), permiten concluir a dicha organización que un mayor gasto en educación no garantiza un mejor rendimiento del alumnado (OECD, 2009). Como se observa en el Gráfico 1, por encima del umbral de los 35.000 dólares de gasto educativo acumulado por alumno, es decir, invertido en un estudiante de los 6 a los 15

GRÁFICO 1. PROMEDIO DEL RENDIMIENTO EN PISA-2009 Y DEL GASTO POR ALUMNO 6-15 AÑOS



Fuente: OECD (2012).

años (situación que se da mayoritariamente en los países ricos), no existe una relación entre el nivel de inversión y el rendimiento. En este contexto, se señala que el gasto en educación es menos importante que la forma en que se emplean dichos recursos (OECD, 2012). Sin embargo, por debajo de dicho nivel de gasto (al que pertenecen de manera mayoritaria economías con bajos ingresos) sí se constata una clara relación positiva entre gasto en educación y adquisición de competencias.

Más allá del simple análisis de correlación mostrado entre gasto educativo y adquisición de competencias, a continuación se exponen los principales resultados de análisis de causalidad entre gasto educativo y rendimiento académico. La revisión de trabajos se agrupa en distintos ámbitos: en primer lugar, se exponen investigaciones realizadas con grandes muestras de países; en segundo lugar, se detallan resultados de meta-análisis; en tercer lugar, se evalúan estudios empíricos específicos referidos a países desarrollados, en especial a Estados Unidos (EE.UU.); finalmente, se evalúan políticas educativas focalizadas en alumnos socialmente desfavorecidos, ya sea con intervenciones sobre el sistema escolar o las familias.

Investigaciones realizadas sobre grandes muestras de países

El primer grupo de trabajos analizado señala mayoritariamente una ausencia de efecto del gasto educativo sobre el rendimiento de los estudiantes. Así, Hanushek y Kimko (2000) subrayan que los recursos asociados al gasto educativo (ya sea medido por alumno o como porcentaje del PIB) no inciden significativamente sobre los resultados de los alumnos en pruebas internacionales con muestras de hasta 70 países. A la misma conclusión llegan Lee y Barro (2001), con datos de panel y una mayor muestra de países, quienes también señalan el nulo papel del gasto público por alumno. Asimismo, en un estudio que considera seis evaluaciones de competencias internacionales, durante el periodo 1964 a 2005, Altinok (2010) confirma la no significatividad del gasto por alumno y del gasto educativo con respecto al PNB en la adquisición de competencias. También el análisis de Al-Samarrai (2006) para 90 países (la mayoría en desarrollo) concluye que el gasto en educación no incide significativamente sobre el logro educativo (medido como matriculación y finalización de estudios). Asimismo, con datos de la evaluación de TIMSS (*Trends in International Mathematics and Science Study*), Woessmann (2003) y Hanushek y Luque (2003) concluyen que el gasto por alumno no resulta relevante con respecto al rendimiento del alumnado (de hecho, Woessmann (2003) afirma que el gasto educativo incide negativamente en el rendimiento). También sostienen que el gasto por alumno no incide de forma significativa sobre los resultados Fuchs y Woessmann (2007) y Hanushek y Woessmann (2010), con muestras provenientes de la evaluación de PISA. Ahora bien, en el caso de Fuchs y Woessmann

(2007) sí resulta significativo el gasto educativo si se incluye a los países con menor nivel de inversión.

Los estudios anteriores conviven con trabajos que presentan una evidencia contraria. Así, en su análisis para 138 países, entre 1960 y 1985, Chaudhuri y Maitra (2008) revelan que el gasto público en educación incide positivamente en que los alumnos finalicen la Educación Secundaria; dicho efecto es mayor en el caso de las alumnas. Además, Gupta *et al.* (1999) apuntan los efectos positivos del gasto educativo sobre el PIB respecto a la matriculación y progresión en Educación Primaria y secundaria en unos 50 países en desarrollo. Asimismo, con muestras de hasta 50 países, a inicios de la década de 1990, McMahon (1999) concluye que el gasto en educación por estudiante favorece la finalización de la Educación Primaria tanto de alumnos como de alumnas.

Resultados de meta-análisis

Un segundo grupo de estudios se basa en meta-análisis realizados a partir de los resultados de otras investigaciones. En este contexto, los artículos de Hanushek (1986, 1997, 2003) han tenido una gran repercusión. Este autor ha insistido sobre el nulo papel del gasto en educación a la hora de aumentar el rendimiento de los estudiantes. Ahora bien, el meta-análisis de Krueger (2003), que parte de los mismos estudios revisados por Hanushek, cuestiona los resultados de este último y señala que si a los trabajos se les concede la misma importancia (en la ponderación) los recursos están relacionados sistemáticamente con el rendimiento del alumnado. De hecho, ya anteriormente, Hedges *et al.* (1994a, 1994b) habían cuestionado los resultados de Hanushek y destacado el efecto positivo de los recursos educativos en el logro académico. Asimismo, la revisión de 46 trabajos de Dewey *et al.* (2000), entre los que se incluyen 28 análisis evaluados por Hanushek (1986), refuerza el papel positivo del gasto educativo por alumno. Finalmente, un meta-análisis anterior de Fuller (1985) había destacado la importancia del gasto por alumno y el gasto en escuelas en el rendimiento académico en los países en desarrollo.

Estudios empíricos referidos a países desarrollados

Con respecto a los estudios en países desarrollados, con resultados del *National Assessment of Educational Progress* (Educación Secundaria), en EE.UU., Hanushek (1989) también señala la inexistencia de una relación robusta entre gasto educativo y rendimiento académico. Asimismo, en evaluaciones de cambios en las políticas de financiación de la educación a nivel de diversos estados de EE.UU., Downes (1992), en California, y Clark (2003), en Kentucky, no encuentran efectos de mayores inversio-

nes en educación sobre el rendimiento en exámenes externos de conocimientos. Contrariamente, otros estudios de este tipo sí apuntan a la importancia del gasto educativo en los resultados de los estudiantes. Así, Guryan (2001) analiza cambios en el sistema de financiación de escuelas de Massachusetts (a través de los recursos monetarios que reciben los distritos del Estado), en la década de 1990, y señala que el aumento de la financiación mejoró las puntuaciones de los alumnos de cuarto grado (si bien no las de aquellos en octavo). En un análisis similar para Michigan, Papke (2005) destaca que la mejora de resultados en matemáticas derivada de la mayor inversión educativa en primaria se dio especialmente en las escuelas con peores recursos iniciales, un resultado similar al de Roy (2011) para ese mismo estado en su análisis a nivel de distritos. En otro contexto geográfico, Lavy (2012) revela cómo una mayor dotación de recursos financieros a las escuelas con más alumnos (en especial si son socialmente desaventajados) y más tiempo de instrucción mejoraron los resultados (en varios ámbitos) de los alumnos de Educación Primaria en Israel. Este estudio destaca no sólo por sus resultados, sino porque con la metodología aplicada (que soluciona los problemas de endogeneidad) obtiene resultados opuestos a los que se obtienen utilizando una regresión por MCO, lo que pone de relieve la importancia del método de análisis escogido.

Políticas educativas focalizadas en alumnos socialmente desfavorecidos

Además de analizar sus efectos agregados, resulta relevante analizar el efecto del gasto educativo sobre el rendimiento académico de determinados colectivos (en particular los alumnos socialmente desfavorecidos)[1]. Diversos estudios referidos al Reino Unido constatan un pequeño efecto positivo provocado por la asignación adicional de recursos financieros a las escuelas sobre el rendimiento de todo el alumnado y, en especial, de los estudiantes desaventajados socialmente. Así, en su análisis para la Educación Primaria, Holmlund *et al.* (2010) describen las mejoras provocadas por un mayor gasto por alumno sobre el resultado de los estudiantes en los test nacionales existentes al final de dicho ciclo. Gibbons *et al.* (2011), por su parte, señalan el mejor rendimiento de los alumnos en los centros que reciben más recursos cuando se comparan escuelas similares. A esta conclusión llegan también los análisis para Educación Secundaria de Machin *et al.* (2010) y Levacic *et al.* (2005). Como indican los primeros, los recursos adicionales son relevantes para los niños en las escuelas más pobres, aunque pequeños cambios en los recursos tienen poco

1. Obviamente, el éxito en las intervenciones en centros con bajo rendimiento de sus alumnos no puede basarse sólo en una mayor dotación de recursos. Cambios en la cultura, organización y liderazgo del centro, en las prácticas docentes, así como en el entorno del mismo y la relación del centro con dicho entorno son relevantes (Corallo y McDonald, 2002). Aquí se trata, sin embargo, de conocer el efecto del gasto sobre el rendimiento.

efecto en los menos capacitados. Levacic *et al.* (2005), por su parte, concluyen que la disponibilidad de más recursos mejora los resultados en las pruebas de matemáticas y ciencias, pero sólo en lenguaje (inglés) mejoran aquellos alumnos socialmente más desaventajados (definidos como los elegibles para recibir becas-comedor). A este respecto, Jenkins *et al.* (2006) encuentran efectos positivos del aumento de recursos sobre las competencias adquiridas por los alumnos en las pruebas al final de la secundaria (*General Certificate of Secondary Education*) en ciencias, pero no en matemáticas ni en lenguaje.

Una línea de trabajo alternativa evalúa los efectos de los programas de financiación de la "demanda", es decir, de ayudas financieras a las familias, especialmente de grupos desfavorecidos, para que los hijos alcancen niveles superiores del sistema escolar (usualmente son las madres las que reciben las transferencias monetarias y las que asumen los compromisos del programa). La mayoría de este tipo de intervenciones se implementa en países en desarrollo, que a veces cuentan con el apoyo del Banco Mundial y el Banco Interamericano de Desarrollo. Se pueden destacar como programas pioneros en este campo los que entran en funcionamiento en diversos municipios y en el Distrito Federal de Brasil a mediados de la década de 1990, así como el programa "Progresa" de México, iniciado en 1997 (véase un resumen de intervenciones en Patrinos (2002), Cardoso y Portela Souza (2003), Slavin (2010) y Cecchini y Madariaga (2011). Como señalan estos últimos, en el año 2011 se estaban aplicando programas de transferencias condicionadas en dieciocho países iberoamericanos que beneficiaban a 25 millones de familias (alrededor de 113 millones de personas), con un gasto aproximado del 0,4% del total del PIB de América Latina y el Caribe. Este gasto tiene un importante efecto redistributivo dado que, pese a su modestia relativa, está focalizado en poblaciones de ingresos muy bajos. Los compromisos que establecen los programas de transferencias condicionadas en el ámbito de la educación consisten, en la mayor parte de los programas, en asistir a la escuela con regularidad; esta regularidad se establece mediante un porcentaje del total de clases o como un máximo de inasistencias. En algunos programas, además, pueden ponerse otras condiciones vinculadas con una mayor actividad de los padres en la escuela, la participación de adultos en programas de alfabetización o la retirada de los menores de actividades de trabajo remunerado.

Sin embargo, también se han dado iniciativas de este tipo en países desarrollados (como Israel, diversas ciudades y estados de Estados Unidos o el Reino Unido). Entre estos últimos, a continuación consideramos sólo aquellos programas orientados no sólo a la mejora de la asistencia al centro escolar sino a la mejora del rendimiento académico[2].

2. En una breve revisión de programas en países en desarrollo, destacan el programa "Progresa", llevado a cabo

Angrist y Lavy (2009) muestran el éxito parcial del programa israelí de ayuda financiera orientada a que los alumnos en escuelas de bajos resultados puedan aprobar el examen para acceder a la universidad (*Bagrut*). Dicho programa ofrece transferencias de dinero para escolares que progresen satisfactoriamente durante algunos cursos antes del último de secundaria, se presenten a la prueba señalada y aprueben (las ayudas son acumulativas). Los resultados revelan un efecto positivo sobre la probabilidad de aprobar el *Bagrut* en las chicas (aunque el efecto es pequeño), pero no en los chicos. Asimismo, en EE.UU., el *Monthly Grade Stipend* es un programa federal que entrega dinero a estudiantes de renta baja (con una composición multiétnica donde predominan los afroamericanos, con un 45%), si éstos obtienen buenas notas en los estudios de secundaria superior. El programa genera un aumento de 10 puntos porcentuales en la probabilidad de obtener buenos resultados en el grupo que recibe la ayuda respecto a grupos de control similares (la probabilidad pasa del 51% al 61%), si bien las tasas de graduación fueron similares (Spencer *et al.*, 2005). Finalmente, Slavin (2010) analiza un programa similar al "Progresa", implantado en la ciudad de Nueva York ("*Opportunity* NYC") que, entre otras medidas, paga a los estudiantes (en primaria) directamente en función de sus resultados en test estatales de competencias. Una primera evaluación no ofrece resultados significativos de dicho programa sobre la mejora del éxito de los alumnos en las pruebas, si bien se hizo sobre alumnos que realizaron dichos exámenes sólo 6 meses después, como máximo, de haberse incorporado al programa.

En general, este tipo de intervenciones tienen un escaso éxito, en términos de mejora de la asistencia de los alumnos y su logro académico, en los países desarrollados, a diferencia de los programas similares aplicados en países en desarrollo, (Slavin, 2010). Ahora bien, los matices son importantes y, en consecuencia, las características del programa y a quién va dirigido pueden ser determinantes, de modo que debe seguir examinándose cada caso en particular. Además, como puntualiza Lupton (2005), al evaluar el éxito de los programas normalmente se mide la calidad educativa en términos de resultado, pero no en términos del progreso que consiguen los estudiantes que empiezan con altos niveles de desventaja, un elemento que debe tenerse en cuenta en este tipo de actuaciones.

---

por el gobierno de México y el Banco Mundial y el programa "Bolsa Escola" de Brasil. "Progresa" ha servido de modelo para otros países de América Latina. Consiste en un conjunto de medidas de lucha contra la pobreza, entre las que destaca el pago de subsidios en efectivo a las familias, condicionado a que sus hijos vayan a la escuela. Como indican Leclercq (2005) y Slavin (2010), el programa aumentó la matrícula para todos los estudiantes (aunque especialmente en las niñas) en los grados 1 a 8. En el segundo caso, el programa brasileño se ha dirigido a las familias con miembros entre los 7 y 14 años con el objetivo, que se consigue con éxito, de que incrementaran su logro educativo mediante la recepción de una prestación si se daba un mínimo de asistencia al centro escolar (Cardoso y Portela Souza, 2003).

Un último ámbito de análisis con respecto a la importancia del gasto en educación en grupos desfavorecidos debe considerar las becas y ayudas al estudio. El efecto positivo de dicho tipo de ayudas (en particular de las becas) sobre el acceso y el rendimiento universitario (definido usualmente como permanencia y finalización de la carrera) está claramente demostrado –véase, entre otros, Dynarski (2003), Deming y Dynarski (2009), Castleman y Long (2012) y Dynarski y Scott-Clayton (2013)–. Ahora bien, como subrayan Hossler *et al.* (2009) y Arendt (2013), en la mayoría de aportaciones se observa que los efectos sobre la permanencia son menores que sobre el acceso.

Asimismo, las características de las ayudas son relevantes: se constata un mayor efecto beneficioso de las becas que de los préstamos, tanto en referencia al acceso a la universidad como a la permanencia (aunque especialmente en este segundo caso). También se aprecia un mayor efecto sobre los estudiantes de menor nivel de renta (Heller, 2008). Con respecto a las becas, aquellas con una estructura más sencilla, fáciles de comprender, de más cobertura y cuantitativamente más importantes son más efectivas (Dynarski, 2003; Dynarski y Scott-Clayton, 2006; Mundel, 2008; Hossler *et al.*, 2009). Además, para los EE.UU., Stater (2009) muestra el efecto beneficioso sobre las notas obtenidas que tienen tanto las becas basadas en la renta como aquellas basadas en el mérito (definido en función del rendimiento académico u otras características no relacionadas necesariamente con la renta), si bien Scott-Clayton (2012) y Dynarski y Scott-Clayton (2013) señalan el mayor efecto beneficioso de las becas vinculadas con la consecución del éxito académico en la universidad. Teniendo en cuenta los distintos colectivos a los que se dirigen las ayudas, Paulsen y St. John (2002) y St. John *et al.* (2005) encuentran diferencias en los efectos de las becas, en EE.UU., en función del grupo étnico: las becas inciden más en las decisiones de los estudiantes afroamericanos que en las de los blancos. Finalmente, para el caso de Italia, Mealli y Rampichini (2012) muestran que las becas reducen la probabilidad de que los alumnos de bajo nivel de renta abandonen los estudios, como es de esperar, pero no en los niveles de renta más bajos. Por último, para el caso español, Mediavilla (2013) analiza el efecto de las becas sobre la probabilidad de que los alumnos finalicen con éxito el nivel secundario postobligatorio. Este autor destaca el efecto beneficioso de las becas en el conjunto de alumnos y, en especial, en aquellos socialmente más desfavorecidos.

RECURSOS MATERIALES Y TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN (TIC)

Más allá de conocer si alguna variable referida al gasto educativo es significativa a la hora de explicar el rendimiento, parte de los estudios empíricos también analiza

la efectividad de recursos materiales concretos disponibles en el proceso. A continuación revisamos dos tipos de análisis. En primer lugar, investigaciones que se centran en el estudio de cualquier tipo de recurso material; en segundo lugar, literatura que específicamente analiza la inversión en tecnologías de la información y la comunicación.

Recursos materiales

Los estudios que evalúan el efecto de los recursos materiales escolares sobre el rendimiento de los alumnos permiten concluir que existe un efecto positivo en los países en desarrollo pero un nulo efecto en los países desarrollados (Heyneman y Loxley, 1982, 1983; Leclerq, 2005). El resultado resulta razonable debido a que en los segundos se dan dotaciones de recursos aceptables y bastante similares (Mancebón y Muñiz, 2003). Como se indica en OECD (2010), no se trata de que los recursos no sean importantes en el logro escolar, sino de que si los centros tienen las necesidades materiales necesarias para la docencia cubiertas los recursos materiales adicionales apenas incidirán en el resultado.

A modo de ejemplo, Fuller (1985) detalla los factores que inciden positivamente en el rendimiento académico de los alumnos en los países en desarrollo: gasto en libros, materiales de texto y equipamientos docentes, así como en alimentación de los estudiantes. Por su parte, Glewwe *et al.* (2011) reducen el papel de los elementos pedagógicos (como los libros de texto) pero no la importancia del resto de recursos materiales. Así, para estos países, una escuela en pleno funcionamiento (con una infraestructura de calidad, con escritorios, mesas y sillas, y con una biblioteca escolar) es un recurso que afecta positivamente al aprendizaje del alumnado.

Ahora bien, en este ámbito también existen divergencias, aunque muy minoritarias. Entre éstas destaca el análisis de Lietz (2010), para 21 países y regiones europeas, donde se subraya la existencia de un efecto positivo de los recursos materiales escolares sobre el rendimiento del alumnado en matemáticas, concluyendo que dicho efecto no varía en función del nivel de renta de los territorios considerados, contrariamente a la hipótesis Heyneman-Loxley. En el ámbito español, sin embargo, la evidencia empírica sigue la corriente mayoritaria: con datos de evaluaciones de PISA, López *et al.* (2009) y Cordero *et al.* (2012) concluyen que los recursos materiales no inciden en el rendimiento de los estudiantes.

Tecnologías de la información y la comunicación (TIC)

Los resultados de la evidencia empírica existente en torno a la relación causal entre la disponibilidad y el uso de las TIC en los centros escolares y el rendimiento

académico de los estudiantes no son concluyentes. Entre los estudios con resultados negativos destacan los siguientes. Angrist y Lavy (2002) analizan un programa de dotación de recursos para adquirir ordenadores en escuelas que imparten Educación Primaria o secundaria en Israel, y encuentran un efecto negativo de estos recursos sobre los resultados de los alumnos en matemáticas en primaria y nulo en secundaria. Asimismo, Leuven *et al.* (2007) evalúan un programa de ayuda para comprar ordenadores y programas informáticos en escuelas desaventajadas de primaria en los Países Bajos, y tampoco constatan efecto alguno sobre los resultados de los estudiantes en lenguaje y matemáticas. Además, Goolsbee y Guryan (2006) analizan los subsidios para que las escuelas inviertan en Internet y comunicaciones en EE.UU. y constatan que dicha inversión no tiene impacto alguno significativo en ninguna medida analizada relacionada con el logro académico. Tampoco Rouse y Krueger (2004), en su estudio de los efectos de un programa educativo informático sobre el conocimiento del lenguaje y la lectura, encuentran resultados satisfactorios para el conjunto del alumnado ni para aquellos con dificultades de aprendizaje. En Colombia, Barrera-Osorio y Linden (2009) revelan que el programa de introducción de ordenadores en escuelas no ha mejorado el rendimiento en matemáticas y lenguaje de los alumnos, si bien la falta de éxito se asocia al propio uso de los ordenadores, que se dedicaron más a enseñar a los alumnos a utilizarlos que a la enseñanza de materias. Finalmente, el análisis internacional de Fuchs y Woessman (2004), con alumnos que participaron en la evaluación de PISA-2000, concluye que el uso de ordenador no tiene un impacto positivo en el rendimiento de los alumnos.

Sin embargo, otros trabajos sí revelan efectos positivos de la inversión en TIC. Así, Machin *et al.* (2007) en Inglaterra, Barrow *et al.* (2009) en EE.UU., y Banerjee *et al.* (2007) en India muestran un efecto causal positivo en el uso de las TIC y el logro escolar. En concreto, Machin *et al.* (2007) describieron el efecto beneficioso de aumentar el gasto en TIC sobre los resultados en inglés y ciencias, aunque no en matemáticas, en alumnos de Educación Primaria. Asimismo, Barrow *et al.* (2009) señalan el éxito de un programa informático para aprender matemáticas en varios distritos de Los Ángeles, de modo que, en su estudio aleatorio, los alumnos asistidos por ordenador obtuvieron mejores resultados que aquellos que siguieron métodos de enseñanza tradicional. Los autores suponen que la eficacia del programa se debe a mayor instrucción individualizada, ya que los efectos son mayores para los estudiantes en las clases con más alumnos y con altos índices de absentismo. Finalmente, Banerjee *et al.* (2007) enfatizan el mayor beneficio sobre el rendimiento de alumnos rezagados, en escuelas urbanas en la India, que genera la utilización de un programa de aprendizaje de las matemáticas asistido por ordenador frente a la

contratación de maestras jóvenes (aunque con ambos métodos los estudiantes mejoraron sus resultados).

En el ámbito español, los estudios con datos de distintas evaluaciones de PISA alcanzan también resultados no coincidentes en cuanto al efecto de las TICs sobre los resultados. Por ejemplo, algunos estudios (Calero y Escardíbul, 2007; Cordero *et al.*, 2012) muestran la no incidencia de la ratio ordenadores por alumno en el rendimiento académico e incluso la existencia de un efecto negativo (Choi y Calero, 2013). En esta misma línea, en su análisis de la evaluación de distintas competencias llevada a cabo por la Comunidad de Madrid, entre 2006 y 2009, Anghel y Cabrales (2010) concluyen que la variable asociada a las TIC no es significativa para explicar los resultados de los estudiantes. Por el contrario, Cabras y Tena (2013), con datos de la evaluación de PISA de 2012, muestran evidencia del efecto positivo del uso de ordenadores sobre la adquisición de competencias matemáticas de los alumnos, con una mayor incidencia en estudiantes de grupos socioeconómicos más desfavorecidos.

En suma, las investigaciones presentan resultados dispares. Dado que la mayor parte de los alumnos que hoy están en las aulas tendrán trabajos en los que el conocimiento de las TIC será capital parece innegable que la incorporación de las TIC en las aulas es imprescindible. De hecho, como señala OECD (2011), su presencia en las escuelas es masiva, con porcentajes por encima del 90% en el acceso a un ordenador y a Internet en la escuela, tanto en España como en el promedio de países de la OCDE. El debate, aún no resuelto, es cómo deben implantarse las TIC para que no sólo se dé un cambio tecnológico en el aula sino, además, una mejora en el método de enseñanza-aprendizaje y, en consecuencia, en la adquisición de competencias por parte de los estudiantes.

PROFESORADO

Con referencia al profesorado, en primer lugar consideramos su importancia en el logro académico de los estudiantes en términos cuantitativos, es decir, teniendo en cuenta la relación de alumnos por profesor o el número de alumnos por clase (los autores anglosajones acostumbran a clasificar ambos conceptos bajo el término *class-size*). En segundo lugar, nos centraremos en aspectos cualitativos, definidos en función de las características de los docentes (su formación y competencias, su experiencia y salarios). Asimismo, abordaremos, brevemente, el papel de la dirección de los centros.

Profesorado: análisis cuantitativo

El efecto de la variable *class-size* sobre el rendimiento educativo es, a priori, difícil de determinar, ya que si bien menos alumnos por aula permiten una mejor atención del profesorado a cada uno, también se reduce el número de compañeros existentes, que pueden ayudarse en el proceso de aprendizaje, como revela la literatura sobre los efectos compañero o *peer effects* –véase Hoxby (2000a), Zimmer y Toma (2000), Hanushek *et al.* (2003) y Dills (2005)–.

En los análisis macroeconómicos internacionales más conocidos, presentados anteriormente, mientras que Lee y Barro (2001) muestran un efecto positivo asociado a los menores ratios de alumnos por profesor, Hanushek y Kimko (2000) no señalan efecto alguno. Ahora bien, en este ámbito destacan dos cuestiones. Por una parte, el debate entre Hanushek y Krueger a partir de las revisiones de otros estudios (meta-análisis) efectuados por ambos autores; por otra, la evaluación del proyecto STAR (*Student-Teacher Achievement Ratio*) implantado en Tennessee (EE.UU.).

Con referencia al meta-análisis, Hanushek (1986, 2003) concluye que no puede establecerse una relación entre *class-size* y rendimiento educativo, ya que algunos trabajos exponen una relación estadísticamente significativa y positiva, otros significativa y negativa, pero para la gran mayoría la relación es no significativa. Ahora bien, la revisión de Krueger (2003) de los mismos estudios analizados por Hanushek sí concluye que la reducción del valor de dicha variable mejora el rendimiento de los estudiantes.

Esta posición más favorable se constata también en los estudios basados en el Proyecto STAR en Estados Unidos, que tiene la virtud metodológica de ser un experimento aleatorio a gran escala, con 79 escuelas, 340 profesores y 11.600 alumnos durante todo el período de duración. Mediante dicho programa, que se inició en 1985, aplicado a alumnos del último curso de educación infantil, y que duró cuatro años, se redujo el tamaño de clase en unos grupos de tratamiento que pueden ser comparados con grupos de control de mayor tamaño. Como señalan Card y Krueger (1992), Krueger (1999, 2003) y Krueger y Whitmore (2001, 2002) se constata que un menor número de alumnos por clase mejora el rendimiento de los estudiantes, si bien el mayor efecto se da durante el primer año. Además, los beneficios son mayores entre los alumnos afroamericanos y los alumnos económicamente desfavorecidos. Ahora bien, Hanushek (2003, 2011) y Rivkin *et al.* (2005) minimizan los resultados positivos expuestos del proyecto STAR y, además, relativizan su "exportabilidad" a otros ámbitos y enfatizan que las ganancias en rendimiento exigen políticas muy costosas económicamente orientadas a reducir el valor de la variable *class-size*.

Otro tipo de investigaciones aplicadas a entornos específicos de Estados Unidos también presentan resultados no concluyentes. En su evaluación de poco más de

30.000 alumnos en tercero de primaria en Kentucky, Borland *et al.* (2005) concluyen que la reducción del número de alumnos por aula tiene un efecto positivo decreciente, situándose el óptimo entre 21 y23 alumnos. Asimismo, Cho *et al.* (2012) advierten de los beneficios en el rendimiento académico de los estudiantes al inicio de la Educación Primaria (en Minnesota) de la reducción de *class-size*, si bien los resultados son cuantitativamente mucho más reducidos que en el proyecto STAR. También aporta resultados positivos, aunque se desvanecen a medida que los alumnos van finalizando la primaria, el estudio para Texas de Rivkin *et al.* (2005), así como las evaluaciones del programa de reducción del tamaño escolar en educación infantil en California de Jepsen y Rivkin (2000) y Sims (2009). Sin embargo, Chingos (2012) concluye que las reducciones del número de alumnos por aula en la Educación Primaria de Florida no mejoran su rendimiento. En este sentido apuntan también Dee y West (2011) en su análisis de alumnos de octavo grado con datos del *National Educational Longitudinal Study* de 1988 y Hoxby (2000b), quien no encuentra efecto alguno de la variable *class-size* en el rendimiento en primaria en su análisis en Connecticut.

En otros ámbitos geográficos se evidencian beneficios de la reducción del número de alumnos por clase, aunque con efectos reducidos. Así, Bosworth (2011) y Ding y Lehrer (2011) realizan distintos análisis, tanto para Estados Unidos como para algunos países europeos, donde se constata un mayor efecto beneficioso de la reducción del *class-size* para los grupos más desfavorecidos (en términos de renta, etnia y grupo social). En su estudio internacional, básicamente europeo, Woessman y West (2006) también revelan efectos beneficiosos de reducir el *class-size* pero sólo en países con un profesorado de bajos salarios. Asimismo, Vignoles *et al.* (2000) proporcionan evidencia para Estados Unidos y el Reino Unido de que el beneficio de la reducción del número de alumnos por aula sobre el rendimiento escolar es muy reducido con respecto al coste que supone este tipo de política. Además, con una metodología equivalente a los experimentos aleatorios, Angrist y Lavy (1999) concluyen que un menor número de alumnos por clase mejora los resultados de matemáticas y lectura en los cursos finales de la Educación Primaria en Israel[3]. Finalmente, con datos de diversas olas del *National Child Development Study* de 1958 en Gran Bretaña, Iacovou (2002) muestra el efecto beneficioso de reducir el tamaño escolar en educación infantil y primaria.

Entre los estudios más recientes no referidos a los Estados Unidos, los principales resultados son los siguientes. En Noruega, Leuven *et al.* (2008) no encuentran

3. El estudio de Angrist y Lavy (1999) y el anteriormente citado de Hoxby (2000b) revelan la importancia de la metodología utilizada. En ambos casos, simples regresiones por MCO dan resultados opuestos a los finalmente obtenidos mediante un análisis con variables instrumentales. Asimismo, tras aplicar a su estudio la metodología utilizada por Angrist y Lavy (1999), Hoxby (2000b) alerta de los distintos resultados que pueden producir diferentes métodos de creación de instrumentos.

efecto positivo alguno de reducir el número de alumnos por aula. Asimismo, considerando sólo escuelas con dificultades en los Países Bajos, Leuven *et al.* (2007) no constatan ninguna incidencia de un programa de incremento de recursos de profesorado para escuelas con un alto porcentaje de alumnos inmigrantes de familias socialmente desaventajadas. Ahora bien, los propios autores señalan que este resultado puede venir motivado por las características del propio plan, ya que en dicho país ya se dan más recursos a las escuelas con este tipo de alumnado, de modo que el complemento económico para profesorado, que debía destinarse a contratar nuevo profesorado (poco útil en centros con clases reducidas) o a aumentar el salario del profesorado (de manera no permanente ni vinculado a ningún sistema de incentivos) puede tener poco sentido. Sin embargo, Fredriksson *et al.* (2013) destacan el efecto beneficioso de tener clases menos numerosas en Suecia, no sólo en primaria sino también posteriormente, ya que dicha reducción permite alcanzar mayores niveles educativos en el futuro. Asimismo, Iversen y Bonesronning (2011) encuentran un efecto positivo para aquellos alumnos socialmente desaventajados (en función del nivel de estudios de los padres y, especialmente, en función de no vivir en familias nucleares). De hecho, Nusche (2009) aporta evidencia empírica para diversos países sobre el efecto beneficioso que supone la reducción de la variable *class-size* de los alumnos desaventajados (definidos en función de la renta, o su condición de minoría étnica o inmigrante). También Brunello y Checchi (2005) indican que la reducción de la ratio de alumnos por profesor mejora el resultado y el logro escolar de los alumnos. En el estudio consideran una muestra representativa de italianos nacidos entre 1941 y 1970 y revelan que el impacto positivo ha sido particularmente importante para las personas nacidas en las regiones más pobres y, a su vez, en los grupos con menor nivel de renta. Asimismo, Valdenaire (2006) expone los efectos beneficiosos de la reducción del tamaño escolar en los resultados escolares de alumnos en primaria y secundaria inferior, en Francia, y destaca el mayor efecto positivo en escuelas socioeconómicamente desfavorecidas (en las denominadas *Zones d'Education Prioritaire*). Ahora bien, para este tipo de centros en zonas ZEP, Bénabou *et al.* (2009) concluyen que la financiación extraordinaria dada para pagar profesores y aumentar las horas de clase no mejoraron los resultados en evaluaciones. Finalmente, Levacic *et al.* (2005) y Jenkins *et al.* (2006) apuntan la mayor eficiencia de reducir específicamente la ratio alumnos/profesor en lugar de mejorar de forma general el gasto por alumno en el Reino Unido.

Para el caso español, en un análisis regional, Mora, Escardíbul y Espasa (2010) destacan el efecto positivo de la relación de alumnos por profesor en el abandono escolar, mientras que el número de estudiantes por aula apenas resulta significativo. Asimismo, en los estudios con datos de PISA de diversas evaluaciones, Zinovyeva

et al. (2008), López *et al.* (2009), Calero *et al.* (2010), Cordero *et al.* (2012), Choi y Calero (2013) y Anghel *et al.* (2013) señalan la no significatividad de la variable *class-size* en el rendimiento del alumnado. Sin embargo, Calero y Escardíbul (2007) sí muestran que un aumento en la ratio alumnos por profesor reduce el rendimiento de los alumnos en las pruebas de matemáticas. En este sentido, con datos de la evaluación de TIMSS, Woessman y West (2006) indican los beneficios de la reducción de dicha variable en el aprendizaje de ciencias pero no en matemáticas. Finalmente, con datos de la evaluación de competencias de alumnos en sexto curso de primaria de la Comunidad de Madrid, para diversos años y diversos ámbitos, Anghel y Cabrales (2010) subrayan la no significatividad de las variables asociadas tanto al alumnado por aula como a la ratio alumnos por profesor en el resultado en dichas pruebas.

En suma, los trabajos con mejores metodologías permiten concluir que las variables asociadas a *class-size* son importantes a la hora de explicar el rendimiento educativo, en especial en educación infantil y primaria, si bien sus efectos no parecen de gran magnitud y, en consecuencia, para ser efectiva deben aplicarse políticas costosas de fuerte reducción de alumnos por aula o profesor (Vignoles *et al.*, 2000; Chingos, 2012).

Profesorado: análisis cualitativo (características)

Numerosas investigaciones que consideran el papel de los *inputs* en la función de producción educativa se han centrado en la calidad del profesorado. De hecho, existen dos tipos de aproximaciones. Por un lado, autores que introducen una variable referida al profesorado en funciones de producción educativa (en general, se incluyen efectos fijos relacionados con los docentes en ecuaciones de rendimiento educativo del alumnado). Por otro lado, estudios que profundizan en el análisis mediante el examen de qué características del profesorado mejoran la calidad educativa. El primer tipo de estudios (efectos fijos del profesorado) revela claramente que los docentes sí importan en el resultado de los alumnos (véase Rockoff, 2004; Hanushek *et al.*, 2005; Rivkin *et al.*, 2005; Aaronson *et al.*, 2007; Leigh, 2010; Hanushek, 2011). En cuanto al segundo tipo, a continuación se presentan los principales resultados de la evidencia empírica existente en torno a las características del profesorado referidas a su formación y certificación, conocimientos, experiencia y salarios.

En cuanto a la formación del profesorado, la posesión de un título de máster no se relaciona con el rendimiento de los alumnos –véanse los estudios para EE.UU. de Hanushek (2003, 2011), Rivkin *et al.* (2005), Hanushek y Rivkin (2006), Clotfelter *et al.* (2006), Harris y Sass (2011) y Chingos y Peterson (2011)–. Sin embargo Fuchs y Woessmann (2007) sí muestran la importancia de tener un máster en el conocimiento en el que se enseña, o bien en pedagogía, en su análisis para 32 países con da-

tos de la evaluación de PISA-2000. Este efecto positivo se da especialmente en matemáticas; para esta competencia, en el caso de EE.UU., la evidencia empírica también muestra una relación positiva entre el logro educativo de los alumnos (en secundaria) y la posesión de titulación referida a las matemáticas por parte del profesorado, tales como una licenciatura o cursos específicos (Darling-Hammond, 2000; Akiba *et al.*, 2007; Goe y Sticler, 2008; Darling-Hammond, 2010).

En este ámbito, la literatura para Estados Unidos ha analizado el rendimiento de los alumnos en función de si el profesorado tiene una certificación o "licencia" estatal para ejercer[4]. Unos estudios cuestionan el efecto de la certificación sobre el rendimiento de los estudiantes –véase Murnane y Steele (2007), Angrist y Guryan (2008), Kane *et al.* (2008) y Chingos y Peterson (2011)–, mientras que otros sí indican sus beneficios: Clotfelter *et al.* (2006), Darling-Hammond (2010) y Akiba *et al.* (2007). Entre estos últimos, la mayoría concluye que la certificación mejora el resultado de los alumnos en el caso de la Educación Secundaria pero no en la primaria. En un ámbito internacional, la certificación sí aparece como determinante positiva de los resultados (véase el análisis internacional citado de Fuchs y Woessmann, 2007). Asimismo, Agasisti y Longobardi (2012) señalan la importancia de la calidad del profesorado (en términos de certificación) en el éxito del alumnado resiliente en Italia, es decir, aquellos alumnos con buenos resultados en la evaluación de PISA de 2009 a pesar de pertenecer a un grupo socialmente desfavorecido.

Diversos autores evalúan también el efecto que pueden tener sobre el rendimiento del alumnado los conocimientos del profesorado, medidos usualmente en función de los test obtenidos en su titulación o en pruebas para poder ejercer como profesores (los estudios se localizan en Estados Unidos). En su revisión de artículos, Hanushek y Rivkin (2006) muestran resultados dispares, con un predominio de trabajos donde la puntuación en dichos test es irrelevante. Resultados similares se exponen en Harris y Sass (2011) para Florida. Sin embargo, sí aparece un efecto positivo de los conocimientos mostrados por el profesorado en investigaciones en otros estados y ciudades norteamericanas –véase Clotfelter *et al.* (2006), Boyd *et al.* (2008), Cantrell *et al.* (2008), Chingos y Peterson (2011)– así como en Perú (Metzler

4. En Estados Unidos no existe una regulación federal única para ejercer de maestro o profesor. Los estados han ido exigiendo obtener una "licencia" o certificación que se logra mediante unas pruebas de conocimientos, pero existen otras vías de acceso. Asimismo, puede solicitarse disponer de ciertas titulaciones para determinadas enseñanzas, o que las mejoras laborales se vinculen con la posesión de dichas titulaciones (por ejemplo de máster, o estar graduado en alguna titulación en concreto). Así, un profesor de matemáticas no necesariamente ha cursado una *major* o asignatura en matemáticas ni en enseñanza de las matemáticas. De hecho, Akiba *et al.* (2007), en su análisis de TIMSS-2003, apuntan que sólo el 47,3% de los alumnos de matemáticas de octavo curso en Estados Unidos ha tenido un profesor con el primer tipo de formación y un 55,3% con el segundo tipo (dichas cifras para el promedio de 46 países analizados, con distintos niveles de desarrollo, son 70,9% y 50,2%, respectivamente).

y Woessmann, 2010). Para el caso de Kentucky, Kukla-Acevedo (2009), además, relaciona como positivos y complementarios los conocimientos y la experiencia de los docentes.

Con respecto precisamente a los años de experiencia, los estudios de Rice (2003), Rivkin *et al.* (2005), Boyd *et al.* (2006), Clotfelter *et al.* (2006), Kane *et al.* (2008), Harris y Sass (2008) y Chingos y Peterson (2011), así como las revisiones de Vignoles *et al.* (2000), Hanushek (1986, 2003), Hanushek y Rivkin (2006) y Metzler y Woessmann (2010) permiten concluir que la experiencia sí resulta relevante, con un efecto positivo sobre el rendimiento de los estudiantes mayoritariamente no lineal en el tiempo. De hecho, el efecto de la mejora provocada por la experiencia parece concentrarse muy al inicio de la práctica docente (a los 2-5 años y, en algunos casos, tan sólo a los 2-3 años) para decrecer a continuación. Sin embargo, una excepción es el análisis de Hanushek y Luque (2003) para diversos países en la prueba TIMSS, quienes apenas encuentran significatividad de esta variable en los diferentes países examinados.

Otro de los elementos de política educativa evaluado en una diversidad de estudios se refiere al salario del profesorado. En las revisiones ya citadas de Hanushek (1986, 2003) se constata que en el 73% de los casos analizados la variable salarial asociada al profesorado no resulta significativa. Ahora bien, cuando sí lo es, en un 20% incide positivamente y sólo en un 7% de manera negativa. En trabajos aplicados a otros países se obtienen resultados similares a los anteriores, aunque con una mayor significatividad del salario del profesorado sobre el rendimiento del alumnado en términos relativos (si bien disminuye la distancia entre estudios que muestran un efecto positivo y los que señalan uno negativo): en el 54% de los casos dicha variable no resulta significativa, en el 31% la incidencia es positiva y en un 15% negativa. Dicha escasez de significatividad en la mayoría de investigaciones es destacada también en las revisiones de Vignoles *et al.* (2000) y Hanushek y Rivkin (2006). Sin embargo, como indican Dolton y Marcenaro-Gutiérrez (2011), debe resaltarse que estos análisis utilizan datos de Estados Unidos, con escasa variabilidad en calidad y salarios del profesorado como para observar diferencias entre los salarios de los docentes y los resultados del alumnado. Asimismo, Loeb y Page (2000) introducen correcciones en dichas limitaciones y concluyen que existe una relación positiva entre salarios del profesorado y rendimiento educativo (medido en este caso como reducción del abandono escolar).

A nivel internacional, con muestras de una diversidad de países, Lee y Barro (2001) concluyen que el salario del profesorado no siempre incide significativamente en el rendimiento del alumnado. Ahora bien, una réplica de dicho estudio efectuada por Sequeira y Robalo (2008) sí revela la significatividad positiva del sa-

lario del profesorado sobre el rendimiento. Además, en un estudio que extiende el realizado por Lee y Barro (2001), con datos de 39 países provenientes de las evaluaciones de TIMSS y PISA, Dolton y Marcenaro (2011) destacan la existencia de un importante efecto positivo del salario del profesorado sobre el rendimiento de los alumnos en todas las disciplinas analizadas, es decir, matemáticas, ciencias y comprensión lectora; dicho resultado persiste si se tienen en cuenta los incrementos salariales en el período analizado en lugar del nivel salarial. A modo de ejemplo, un aumento del 10% en el salario del profesorado mejora un 5-10% los resultados de los alumnos en los test. Finalmente, para el caso de la India, Kingdon y Teal (2010) concluyen que el salario del profesorado incide positivamente en el rendimiento de los estudiantes, pero sólo en centros privados (no en los públicos). Dichos autores suponen que este efecto parcial se debe a que las mejoras salariales en el sector privado se relacionan con sistemas de pago al profesorado relacionados con la actuación (*performance-related pay*). Precisamente, en el ámbito de los salarios, algunos autores se han centrado en el análisis de dichos sistemas de pago, como se detalla a continuación.

Woessmann (2011) muestra la existencia de un incremento de resultados en países participantes en la evaluación de PISA-2003 con sistemas de pago al profesorado relacionados con la actuación, ya sean éstos a nivel de centro, autoridad local o un ámbito superior. En su análisis, los alumnos en países con dichos sistemas de incentivos obtuvieron una puntuación adicional del 24,8% de una desviación estándar en el test de matemáticas y lectura de PISA (y de un 15% en ciencias). Este resultado es similar al observado en la India (Muralidharan y Sundararaman, 2009) y aproximadamente el doble que el señalado por Angrist y Lavy (2009) para Israel. También se constata que este tipo de retribución mejora significativamente el rendimiento del alumnado en Inglaterra (Atkinson *et al.*, 2009) y en Arkansas (Winters *et al.*, 2006). Asimismo, considerando diferentes tipos y grados de aplicación de sistemas de incentivos salariales, Figlio y Kenny (2007) destacan la incidencia positiva de la existencia de incentivos salariales en las escuelas en el rendimiento de alumnado para una muestra de centros en EE.UU.; sin embargo, estos autores destacan que dichos incentivos deben vincularse con determinadas actuaciones del profesorado, ya que si se generalizan a la práctica totalidad de los docentes y son demasiado fáciles de conseguir se vuelven ineficaces. Finalmente, Eberts *et al.* (2000) destacan el éxito en el incremento de la tasa de finalización de estudios y la reducción del abandono escolar de una experiencia en Dallas (EE. UU.) consistente en proporcionar complementos salariales por desempeño para docentes que trabajan en escuelas con una alta proporción de alumnos de minorías étnicas o inmigrantes.

Un último elemento de análisis se refiere a la distribución del profesorado entre centros. Para EE.UU., Murnane y Steele (2007) destacan el problema que supone la desigual distribución del profesorado de mayor calidad entre los centros, de modo que los niños más pobres y los afroamericanos son desproporcionadamente asignados a profesores con una menor preparación y conocimiento. Bénabou (2009) en Francia, Karsten *et al.* (2006) en los Países Bajos, y Field *et al.* (2007), para diversos países de la OCDE describen cómo el profesorado más cualificado y con más experiencia prefiere centros con un porcentaje menor de inmigración. En la revisión de estudios de Akiba *et al.* (2007) se añade el problema de asignar a estos centros a los profesores más jóvenes y los no certificados, con resultados más bajos en pruebas de conocimientos de certificación docente, o sin estudios cursados (*majors*) en las materias que enseñan. Como indican Ingersoll (2001) y Murnane y Steele (2007) la rotación de profesorado en centros con una mayoría de alumnos pobres o no blancos es alta, dadas las más difíciles condiciones de trabajo en estos centros (con respecto al alumnado, a los padres, etc.), y la menor dotación de recursos de los mismos. Para el Reino Unido, Lupton (2004, 2005) y Horgan (2009) subrayan, además, la escasa voluntad que tienen los profesores de residir en las zonas en las que se ubican este tipo de centros, así como los problemas organizativos de funcionamiento que aparecen en los mismos.

Dirección de centros

Un último factor a considerar en el papel de los recursos humanos escolares sobre el rendimiento de los estudiantes es la dirección de los centros escolares (que en algunos países es ejercida por profesores de los mismos).

Un grupo de autores desarrolla un análisis de efectos fijos para los directores en las funciones de producción educativa, como el descrito anteriormente para el caso del profesorado, que revela resultados positivos, de modo que el papel de la dirección de las escuelas aparece como relevante en el rendimiento académico de los estudiantes (véase una revisión en Coelli y Green, 2012).

Sin embargo, autores como Lavy (2008), Clark *et al.* (2009) y Branch *et al.* (2012) sí explicitan qué factor relacionado con la dirección de centros puede incidir en el logro académico. A este respecto, se señalan como positivos el salario, la experiencia laboral y la antigüedad de la dirección en el centro y como negativo su rotación entre escuelas. Asimismo, en un estudio previo referido a Estados Unidos, Brewer (1993) ya mostraba el papel positivo de los directores en el rendimiento del alumnado del centro, en este caso a través de su papel en la selección de los profesores y el establecimiento de objetivos en la escuela.

EFECTO DE LOS RECURSOS SOBRE LA ADQUISICIÓN DE COMPETENCIAS EN ESPAÑA. ANÁLISIS CON DATOS DE LA EVALUACIÓN DE PISA DE 2009

El *Programme for International Student Assessment* (PISA) es una evaluación de competencias que desarrolla la OCDE cada tres años, desde el año 2000, a estudiantes de 15 años de edad. En la edición de 2009 participaron 34 países de la OCDE y 41 asociados. En España fueron evaluados casi 26.000 alumnos, con muestras representativas ampliadas para catorce Comunidades Autónomas.

En este apartado, en primer lugar exponemos una información referida a la dotación de recursos para el conjunto de centros. Posteriormente, con objeto de conocer el grado de desigualdad de la dotación de recursos entre diferentes tipos de centros en el sistema educativo español, presentaremos una serie de resultados bivariantes que relacionan variables relacionadas con los recursos y variables socioeconómicas de los estudiantes[5]. Finalmente, se presentan los resultados de una serie de análisis de regresión relativos al efecto de los recursos escolares en los resultados de PISA del alumnado español en la evaluación de la competencia lectora (la competencia analizada en detalle en la evaluación de 2009).

En el Cuadro 1 aparecen los valores agregados referidos a la disponibilidad de recursos en los centros sostenidos con fondos públicos (titularidad pública o privada concertada). Destacaríamos los siguientes resultados:

- Resulta mayoritaria la percepción de falta de personal de apoyo (afecta al 72,7% del alumnado). Sin embargo, la casi totalidad de escuelas españolas cuentan con profesorado de lenguaje suficiente, así como con biblioteca en el centro.
- La calidad de recursos materiales se calcula mediante un índice. Este índice mide la percepción de la dirección del centro de determinados aspectos que dificultan el aprendizaje: existencia de materiales didácticos, escasez o inadecuación de los equipos de laboratorio de ciencias, equipos para la enseñanza, conectividad a Internet, programas informáticos para la enseñanza, materiales de la biblioteca y recursos audiovisuales. El signo negativo para España denota que la percepción acerca de la dotación de recursos es ligeramente inferior a la media de la OCDE.
- En promedio, el número de alumnos por clase es de casi 22 y la ratio entre alumnos y profesores es de 1 docente por cada 11 alumnos. Asimismo, el número de ordenadores por alumno de 15 años es 0,6.

5. En el análisis se excluyen los centros escolares privados independientes. Asimismo, los cálculos tienen en cuenta los valores de las variables ponderados poblacionalmente.

CUADRO 1. RECURSOS ESCOLARES EN LOS CENTROS ESPAÑOLES FINANCIADOS PÚBLICAMENTE

VARIABLES	TOTAL MUESTRA
Falta de personal de apoyo	72,7%
Falta profesorado lenguaje	1,0%
Calidad de los recursos materiales (índice)	-0,037
Existencia de una biblioteca	95,7%
Ratio alumnado/profesorado	11,3
Ordenadores por alumno (15 años)	0,578
Alumnos por clase	21,6

Fuente: Elaboración propia a partir de los microdatos de la evaluación de PISA-2009 de la OCDE.

Con objeto de explorar las pautas en la distribución de los recursos presentamos resultados en los que dividimos la muestra en grupos específicos. En primer lugar se toman en consideración los centros en función de su titularidad y composición social; en segundo lugar, se divide a las familias en función de su nivel de renta y nacionalidad. La evidencia presentada permite concluir que, en general, se dan escasas diferencias en función de los grupos considerados; de hecho, la dotación de recursos es, en algunos casos, ligeramente compensatoria, dando más recursos a centros con una composición económico-social más desfavorecida.

Si se divide la muestra en función de la titularidad del centro escolar, en el Cuadro 2 se constata una dotación similar de recursos entre centros públicos y privados concertados. Ahora bien, se observan algunas leves diferencias. En favor de los centros concertados, la menor presencia de falta de dotación de personal de apoyo, así como una leve mejora en la calidad de los recursos materiales. Sin embargo, los centros concertados presentan una menor dotación de biblioteca escolar, así como un mayor número de alumnos por profesor y por clase.

CUADRO 2. RECURSOS ESCOLARES EN LOS CENTROS ESPAÑOLES FINANCIADOS PÚBLICAMENTE POR TITULARIDAD

VARIABLES	CENTROS PÚBLICOS	CENTROS CONCERTADOS
Falta de personal de apoyo	75,7%	64,7%
Falta profesorado lenguaje	1,1%	0,8%
Calidad de los recursos materiales (índice)	-0,051	0,001
Existencia de una biblioteca	97,2%	91,9%
Ratio alumnado/profesorado	9,4	16,4
Ordenadores por alumno (15 años)	0,58	0,55
Alumnos por clase	20,7	23,7

Fuente: Elaboración propia a partir de los microdatos de la evaluación de PISA-2009 de la OCDE.

El análisis mediante una *proxy* de la clase social del centro escolar (se divide a los centros en función del valor de la moda de la ocupación de los padres de los alumnos) evidencia leves diferencias entre centros (Cuadro 3). De hecho, los centros donde predominan padres y madres con ocupaciones de cuello azul y cuello blanco no cualificado se parecen más entre sí que con respecto a los centros con predominio de progenitores en trabajos de cuello blanco cualificado. Este último tipo de centros tienen una mayor ratio de alumnos por profesor y alumnos por clase y una peor ratio de ordenadores por alumno. En buena medida, estas diferencias son atribuibles a la mayor presencia de los centros concertados entre aquellos en los que predomina la categoría de "cuello blanco cualificado".

CUADRO 3. RECURSOS ESCOLARES EN FUNCIÓN DE LA COMPOSICIÓN SOCIOECONÓMICA DEL CENTRO

VARIABLES	ESCUELA CUELLO AZUL	ESCUELA CUELLO BLANCO NO CUALIFICADO	ESCUELA CUELLO BLANCO CUALIFICADO
Falta de personal de apoyo	72,7%	72,5%	73,1%
Falta profesorado lenguaje	1,9%	1,0%	0,5%
Calidad de los recursos materiales (índice)	0,047	-0,062	-0,035
Existencia de una biblioteca	97,9%	95,0%	95,7%
Ratio alumnado/profesorado	10,0	10,5	13,9
Ordenadores por alumno (15 años)	0,57	0,62	0,49
Alumnos por clase	19,9	20,9	24,1

Fuente: Elaboración propia a partir de los microdatos de la evaluación de PISA-2009 de la OCDE.

El Cuadro 4 contiene los valores de las variables escogidas en función de la nacionalidad de los estudiantes. Pocas diferencias significativas pueden señalarse: por un lado, existe un leve mayor porcentaje de alumnado inmigrante, respecto al nativo, en centros con una menor calidad de los recursos materiales y una mayor falta de personal de apoyo, pero los alumnos inmigrantes se escolarizan mayoritariamente en centros con menores niveles de alumno por profesor y por clase (es decir, en centros públicos). Tampoco se observan, en el Cuadro 5, diferencias relevantes entre alumnos con mayores y menores niveles de renta (se divide al total de estudiantes en cinco quintiles y se consideran los dos grupos extremos). En este caso, los alumnos con menor nivel de renta familiar, respecto a los de mayor nivel, acuden a centros con una mayor falta de profesorado de lenguaje, pero con mayor calidad de recursos materiales, así como más profesorado respecto al total de alumnos y menos alumnos por clase.

CUADRO 4. ACCESO A LOS RECURSOS ESCOLARES SEGÚN NACIONALIDAD DE LOS ESTUDIANTES

VARIABLES	NATIVO	INMIGRANTE
Falta de personal de apoyo	72,3%	76,9%
Falta profesorado lenguaje	1,1%	0,9%
Calidad de los recursos materiales (índice)	-0,033	-0,066
Existencia de una biblioteca	95,8%	96,5%
Ratio alumnado/profesorado	11,4	10,6
Ordenadores por alumno (15 años)	0,58	0,57
Alumnos por clase	21,8	19,6

Fuente: Elaboración propia a partir de los microdatos de la evaluación de PISA-2009 de la OCDE.

CUADRO 5. ACCESO A LOS RECURSOS ESCOLARES SEGÚN LA RENTA FAMILIAR

VARIABLES	RENTA. QUINTIL MAYOR	RENTA.QUINTIL MENOR
Falta de personal de apoyo	72,0%	73,5%
Falta profesorado lenguaje	0,7%	1,6%
Calidad de los recursos materiales (índice)	-0,035	-0,028
Existencia de una biblioteca	95,2%	95,6%
Ratio alumnado/profesorado	12,0	10,6
Ordenadores por alumno (15 años)	0,55	0,60
Alumnos por clase	22,6	20,1

Fuente: Elaboración propia a partir de los microdatos de la evaluación de PISA-2009 de la OCDE.

Como se ha mostrado, el análisis bivariante muestra una similitud importante en cuanto a la dotación de recursos entre centros y alumnos diferentes en términos socioeconómicos. A continuación se presentan los resultados de la regresión multivariante de los factores determinantes de la adquisición de competencias en comprensión lectora. El análisis se realiza mediante una regresión multinivel, dada la composición anidada de los datos en PISA (véase una explicación detallada del método de análisis en Calero y Escardíbul, 2007). En este tipo de estudios, los resultados de los alumnos dependen de sus características personales y familiares (nivel 1 de análisis) y de las características de los centros escolares (nivel 2), tal y como se expone en las ecuaciones (1) a (4):

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^{n} \beta_{1j} X_{kij} + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0,\sigma^2) \qquad (1)$$

$$\beta_{0j} = \gamma_{00} + \sum_{l} \gamma_{0l} Z_{lj} + \mu_{0j} \qquad \mu_{0j} \sim N(0, \tau_0) \qquad (2)$$

$$\beta_{1j} = \gamma_{10} \qquad \mu_{1j} \sim N(0, \tau_1) \qquad (3)$$

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{kij} + \gamma_{0l} Z_{lj} + \mu_{0j} + \varepsilon_{ij} \qquad (4)$$

El modelo econométrico expuesto indica que Y_{ij} se refiere a los resultados esperados en comprensión lectora del estudiante "i" en la escuela "j"; X_{kij} es un vector de "k" características del estudiante "i" en la escuela "j" (variables explicativas al nivel 1); Z_{lj} es un vector de "l" características de la escuela "j" (variables al nivel 2). Los efectos aleatorios son μ_j (a nivel de escuela) y ε_{ij}, (a nivel de alumno). Los parámetros estimados se anotan como β. La ecuación (4) se obtiene al introducir las ecuaciones (2) y (3) en la ecuación (1).

El Cuadro 6 recoge los resultados del análisis de regresión para el conjunto de la muestra. Como resulta habitual en este tipo de estudios, las variables personales y familiares son muy relevantes a la hora de explicar los resultados de los estudiantes en las evaluaciones de competencias. Entre las características personales de los jóvenes relacionadas con la escuela podemos destacar que la repetición de curso reduce claramente los resultados obtenidos en la evaluación de PISA, mientras que haber cursado educación infantil más de un curso los incrementan. En el ámbito familiar la disponibilidad de recursos educativos y culturales, así como pertenecer a una familia de estatus social más alto incide positivamente en los resultados obtenidos, como también no ser inmigrante.

En el ámbito de los recursos escolares apenas resulta significativa ninguna variable. Ahora bien, la existencia de biblioteca en la escuela o cierto apoyo a los alumnos inmigrantes incrementan el rendimiento de los estudiantes. Sin embargo, de forma a primera vista contraintuitiva, el número de alumnos por aula tiene un efecto *positivo* decreciente (a mayor número de alumnos por clase mejores resultados, aunque el efecto decrece a medida que aumenta el valor de esta variable). También se asocian *mejoras* de los resultados con un mayor número de alumnos por profesor. Así, parece que las variables relacionadas con la cantidad de docentes se relacionan inversamente con los resultados. A nuestro entender, dicho efecto puede deberse a dos posibles motivos, ambos vinculados con la existencia de formas de causalidad inversa. El primer motivo se produciría cuando los alumnos con peores resultados tengan a su disposición una mejor atención (menos alumnos por aula y más profesores) para poder proseguir en el sistema educativo. El segundo motivo estaría relacionado con la atracción de una mayor cantidad de demanda por parte de los centros más atractivos para los padres, que suelen ser los que alcanzan mejores resultados; en este segundo caso hablamos de causalidad inversa en tanto que no se trata de que los centros alcancen mejo-

res resultado debido a tener más alumnos por profesor, sino que tienen más alumnos por profesor, en última instancia, debido a tener mejores resultados.

De entre los ámbitos tratados en el apartado dedicado al profesorado, el Cuadro 6 señala un leve efecto positivo si la dirección del centro ejerce cierto liderazgo. Otras variables de centro resultan más significativas. Entre éstas destacamos la composición social del mismo (definida en función de los años de escolarización de los padres de los alumnos), así como algunas características asociadas al comportamiento de los estudiantes y la existencia de una buena relación entre profesores y alumnos. También inciden positivamente el desarrollo de actividades extracurriculares y la presencia de competencia de otros centros.

Una vez expuesto el análisis para el conjunto de la muestra, a continuación se presentan resultados para determinadas submuestras seleccionadas a partir de las variables socioeconómicas del análisis bivariante: en el Cuadro 7 se exponen resultados para diferentes centros en función de la titularidad de los mismos y en el 8 teniendo en cuenta su categoría socioeconómica; asimismo, en el Cuadro 9 las muestras se dividen en función del estatus de inmigrante del alumnado y en el Cuadro 10 según la renta familiar[6].

CUADRO 6. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA EL CONJUNTO DE LA MUESTRA

ÁMBITO	VARIABLE	LECTURA	
		COEFICIENTE	ERROR ESTÁNDAR
	Constante	207,689***	28,745
PERSONAL			
	Edad	8,250***	1,511
	Mujer	19,848***	0,940
	Repetición sólo en primaria	-70,390***	3,232
	Repetición sólo en secundaria	-60,119***	1,241
	Repetición en primaria y secundaria	-109,375***	1,991
	Ha cursado educación infantil más de un curso	9,854***	1,308

[CONTINÚA EN PÁGINA SIGUIENTE]

6. Se excluyen como variables explicativas aquellas relacionadas con la atención idiomática a los inmigrantes, debido a su especificidad.

CUADRO 6. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA EL CONJUNTO DE LA MUESTRA

ÁMBITO	VARIABLE	LECTURA	
FAMILIAR			
	Familia nuclear	-0,409	1,255
	Recursos educativos del hogar	2,593***	0,533
	Nativo	19,060***	1,891
	Años de escolarización de la madre	0,192	0,134
	Padre activo	-0,325	1.904
	Hogar cuello azul cualificado	-0,461	1,684
	Hogar cuello blanco no cualificado	5,619***	1,570
	Hogar cuello blanco cualificado	15,610***	1,672
	Hogar con más de 100 libros	22,000***	0,973
RECURSOS DE LA ESCUELA			
	Falta de personal de apoyo	1,019	2,147
	Falta profesorado lenguaje	-3,516	5,822
	Calidad de los recursos materiales	0,761	1,200
	Existencia de una biblioteca	13,956***	2,793
RECURSOS DE LA ESCUELA			
	Ratio alumnado/profesorado	0,596*	0,318
	Ordenadores por alumno (15 años)	-3,536	3,246
	Alumnos por clase	3,252***	0,404
	Alumnos por clase al cuadrado	-0,058***	0,010
	Alumnos lengua madre extranjera: asisten clase regular y sesiones de refuerzo de lenguaje	2,293	2,145
	Alumnos lengua madre extranjera: están en programa de competencias en lenguaje antes de ser transferidos a clases con el resto de alumnos	3,029	2,263
	Alumnos lengua madre extranjera: reciben mucha instrucción en su idioma para dominar ambas lenguas	8,067**	3,346
	Alumnos lengua madre extranjera: se reduce tamaño clase para atender necesidades de estos alumnos	2,249	2,897

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 6. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA EL CONJUNTO DE LA MUESTRA

ÁMBITO	VARIABLE	LECTURA	
PROCESO EDUCATIVO EN LA ESCUELA			
	Autonomía contratación profesorado	-10,835	6,537
	Autonomía fijación salario	-6,302	6,139
	Autonomía incrementos salariales	4,851	5,958
	Autonomía elaboración del presupuesto	-2,861	2,476
	Autonomía uso del presupuesto	2,512	5,548
	Autonomía establecimiento sistemas de evaluación	-1,080	2,465
	Autonomía elección libros de texto	-1,654	4,954
	Autonomía elección contenidos	0,685	2,139
	Agrupación alumnos por niveles entre clases	1,287	2,128
	Agrupación alumnos por niveles en las clases	3,670	2,198
	Evaluación alumnado usado para comparar escuela con promedio nacional/regional	-0,271	2,570
	Evaluación alumnado usada para comparar escuela con otras escuelas	1,888	2,578
	Se da información a los padres del rendimiento del alumnado con respecto promedio nacional/regional	0,062	2,859
	Se da información a los padres del rendimiento del alumnado con respecto a otras escuelas	-4,786	4,002
	Buen comportamiento del alumnado	2,999**	1,273
	Buena relación alumnado-profesorado	3,700***	0,475
	Clima de disciplina	2,390***	0,473
	Concertada	5,966	6,778
	Independiente	6,968	8,020
	Directora de centro	-0,854	1,998

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 6. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA EL CONJUNTO DE LA MUESTRA

ÁMBITO	VARIABLE	LECTURA	
CARACTERÍSTICAS DE LA ESCUELA Y SU ALUMNADO			
	Liderazgo de la dirección	2,080*	1,178
	Número de alumnos	-0,006*	0,003
	Municipio 100.000-1.000.000 hab.	-2,372	2,592
	Municipio >1.000.000 hab.	5,881**	2,779
	Prácticas de selección en la admisión: competencias y escuela previa	-2,669	1,886
	Se da publicidad a los logros del alumnado del centro	-3,809	3,778
	Actividades extracurriculares	2,943**	1,289
	Competencia 1 escuela	6,885*	3,481
	Competencia dos o más escuelas	6,884**	3,300
	Presión de minoría de padres para mejorar rendimiento	-1,069	2,125
	Presión constante de muchos padres para mejorar rendimiento	-0,082	3,348
CARACTERÍSTICAS DE LA ESCUELA Y SU ALUMNADO			
	Años escolarización padres	5,493***	0,967
	Centro con mayoría de padres cuello blanco no cualificado	2,647	3,578
	Centro con mayoría de padres cuello blanco cualificado	0,060	4,797
	Proporción de inmigrantes	10,883	8,749
	Proporción de chicas	0,150	0,135

Nota: Errores estándar robustos. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente.

Como aparece en el Cuadro 7, las diferencias en la incidencia de las variables de recursos por tipo de centro denotan que el efecto beneficioso de la existencia de biblioteca sólo afecta a los centros públicos, mientras que la ratio alumnos por profesor resulta sólo estadísticamente significativa (con signo positivo) en los centros concertados; de este modo, se refuerza la idea de la causalidad inversa de esta variable. Si dividimos los centros por categoría socioeconómica (Cuadro 8), y compara-

CUADRO 7. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA MUESTRAS EN FUNCIÓN DE LA TITULARIDAD

VARIABLE	C. PÚBLICOS		C. CONCERTADOS	
	COEF.	E.S.	COEF.	E.S.
Falta de personal de apoyo	0,703	3,243	-6,925	5,673
Falta profesorado lenguaje	-2,246	10,851	16,516	19,943
Calidad recursos materiales	-0,169	1,882	2,153	3,021
Existencia de una biblioteca	17,347**	8,303	8,501	5,210
Ratio alumnado/profesorado	0,207	1,028	1,463***	0,549
Ordenadores por alumno	-8,236	5,514	-11,633	7,207
Alumnos por clase	2,831***	0,797	4,411***	1,042
Alumnos por clase al cuadrado	-0,051**	0,019	-0,107***	0,026
Variables personales, familiares y escolares	Sí			

Nota: Errores estándar robustos. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente.

mos aquellos en situaciones más dispares, la diferencia más significativa es, como en el caso de los centros concertados, que en aquellos centros con predominio de padres con ocupaciones de cuello blanco cualificado la variable asociada a la ratio de alumnos por profesor es significativamente positiva.

CUADRO 8. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA MUESTRAS EN FUNCIÓN DE LA CATEGORÍA SOCIAL DE LOS CENTROS

VARIABLE	C. CUELLO AZUL		C. C. BLANCO CUALIFICADO	
	COEF.	E.S.	COEF.	E.S.
Falta de personal de apoyo	-5,156	7,136	4,263	4,374
Falta profesorado lenguaje	3,641	15,970	0,453	4,735
Calidad recursos materiales	0,069	4,138	0,845	2,516
Existencia de una biblioteca	-9,936	20,946	11,965	7,458
Ratio alumnado/profesorado	-2,216	1,362	1,213**	0,518
Ordenadores por alumno	-0,078	8,606	-4,462	6,762
Alumnos por clase	0,912***	0,328	3,832***	1,331
Alumnos por clase al cuadrado			-0,085***	0,031
Variables personales, familiares y escolares	Sí			

Nota: Errores estándar robustos. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente.

En cuanto al efecto de los recursos en función de la nacionalidad de los estudiantes, en el Cuadro 9 se constata que la ratio alumnado/profesorado sólo incide con signo positivo en el caso de los nativos. De nuevo, parece existir causalidad inversa entre los resultados en las pruebas de evaluación de lectura y dicha variable.

CUADRO 9. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA MUESTRAS EN FUNCIÓN DE LA NACIONALIDAD DE LOS ESTUDIANTES

VARIABLE	NATIVO		INMIGRANTE	
	COEF.	E.S.	COEF.	E.S.
Falta de personal de apoyo	-4,012	3,253	0,578	6,754
Falta profesorado lenguaje	-0,793	8,281	-18,178	21,402
Calidad de los recursos materiales	-0,526	1,867	2,844	3,807
Existencia de una biblioteca	9,331*	5,279	41,272**	16,194
Ratio alumnado/profesorado	1,004**	0,478	0,585	1,203
Ordenadores por alumno (15 años)	9,016	4,929	-19,881**	10,005
Alumnos por clase	3,696***	0,778	0,971**	0,435
Alumnos por clase al cuadrado	-0,080***	0,020		
Variables personales, familiares y escolares	Sí			

Nota: Errores estándar robustos. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente.

Lo mismo sucede con la variable ordenadores por alumno, sólo significativa (con signo negativo) en el caso de los inmigrantes. Para distintos niveles de renta (Cuadro 10), se constatan los mismos efectos de causalidad inversa que los señalados anteriormente para la variable alumnos por profesor y ordenadores por alumno: en este caso, la primera sólo incide con signo positivo en aquellos alumnos en el quintil de mayor nivel de renta y la segunda con signo negativo en aquellos en el quintil de renta menor.

En suma, no parece que existan diferencias importantes entre grupos en cuanto a los determinantes de los resultados de los estudiantes en la evaluación de PISA sobre comprensión lectora asociados a los recursos escolares. De hecho, cuando se producen resulta razonable suponer la existencia de causalidad inversa. Ello se debe a la dotación de recursos de las escuelas en España, bastante igualitaria, e incluso compensadora en algunos factores. Como se constata en Instituto de Evaluación (2010), España es uno de los países con una menor variación de resultados entre centros (19,5%), junto con Suecia y Canadá y por detrás de Finlandia (el promedio para la OCDE es del 41,7%).

CUADRO 10. RESULTADOS DEL ANÁLISIS DE REGRESIÓN PARA MUESTRAS EN FUNCIÓN DE LA RENTA

VARIABLE	RENTA QUINTIL MAYOR		RENTA QUINTIL MENOR	
	COEF.	E.S.	COEF.	E.S.
Falta de personal de apoyo	-7,830	5,060	-2,327	4,411
Falta profesorado lenguaje	4,493	14,302	5,992	10,950
Calidad de los recursos materiales	-1,262	3,015	3,213	2,537
Existencia de una biblioteca	27,634***	8,521	23,716*	12,337
Ratio alumnado/profesorado	5,289**	2,338	1,148	0,719
Ratio alumnado/profesorado al cuadrado	-0,150*	0,079		
Ordenadores por alumno (15 años)	-2,987	6,277	-12,712*	6,591
Alumnos por clase	0,487	0,337	4,006***	1,159
Alumnos por clase al cuadrado			-0,090***	0,031
Variables personales, familiares y escolares	Sí			

Nota: Errores estándar robustos. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente.

CONCLUSIONES

Nos referiremos en estas conclusiones a las dos partes diferenciadas de este artículo, la primera referida a la revisión de la literatura y la segunda a los análisis propios basados en PISA-2009.

Por lo que respecta a la revisión de múltiples investigaciones en torno al efecto de los recursos en el rendimiento académico de los estudiantes no es concluyente, pero sí permite señalar una serie de elementos mayoritarios.

En primer lugar, el gasto educativo tiene cierta incidencia positiva en el logro académico. Así sucede claramente para los niveles más bajos de gasto, de modo que en esos niveles los incrementos del mismo permiten aumentos significativos del rendimiento de los estudiantes. La evidencia no está tan clara para los niveles de gasto elevados. Ello permite concluir, sin embargo, que no estamos tratando con una variable irrelevante y que no resulta aconsejable reducir el gasto educativo en países cercanos a umbrales de riesgo establecidos por la OCDE.

Una vez considerado el nivel total de gasto, obviamente resulta más relevante conocer en qué elementos es mejor invertir si se desea mejorar el rendimiento. La

introducción de TIC en las aulas no parece un destino muy relevante para la inversión. Dado que dicho aumento de recursos informáticos se mantendrá en el futuro cabe reflexionar en torno a que no sólo se introduzcan los mismos sino que se aproveche dicho cambio para variar las metodologías de aprendizaje, ya que algunas experiencias al respecto sí han resultado positivas en la mejora del rendimiento de los estudiantes. Para el conjunto de recursos materiales, la evidencia revisada indica su escaso o nulo papel en el rendimiento del alumnado en los países desarrollados. Dicho fenómeno está acorde con lo indicado anteriormente para el total de gasto, de modo que una vez que el sistema educativo está suficientemente dotado de recursos un aumento de los mismos no resulta relevante en el rendimiento de los alumnos. Una vez más, cabe hacer la misma reflexión que hacíamos al referirnos al total de gasto en aquellos países que están sufriendo severos recortes en la dotación de recursos educativos y pueden alcanzar umbrales que afecten negativamente al rendimiento de los alumnos. No debe olvidarse que la evidencia analizada siempre muestra efectos mucho más favorables de los recursos en centros y alumnos desfavorecidos, motivo por el cual en caso de reducción de la financiación en educación debe tenerse en cuenta no afectar a este tipo de centros.

Con respecto al profesorado, de la revisión de estudios puede concluirse que reducir la cantidad de alumnos por profesor (o el número de alumnos en el aula) puede tener un efecto positivo sobre el logro académico de los estudiantes, en especial en las etapas iniciales del sistema educativo (infantil y primaria). Ahora bien, en general, para obtener beneficios significativos debe producirse una elevada inversión que, seguramente, sería mejor destinar a mejorar la calidad del profesorado. En cualquier caso, parece razonable suponer que existe un intervalo de *class-size*, que puede oscilar por países o entornos determinados, por encima del cual un mayor número de alumnos incide negativamente en el rendimiento, y por debajo del cual disponer de más recursos en profesorado no aporta beneficio alguno (y, por el contrario, introduce un efecto negativo a través de la reducción de *peers* en el aula). A nuestro entender, la búsqueda de evidencia empírica debería insistir en el estudio de dichos intervalos, objetivo planteado en alguno de los análisis revisados.

La evidencia empírica existente sobre diversos aspectos relacionados con el profesorado permite concluir, con alguna salvedad, que los conocimientos y capacidades del profesorado, así como su salario (en especial si se vincula con el rendimiento del estudiante), y parcialmente los años de experiencia docente (muy al inicio de la vida laboral), inciden positivamente sobre el rendimiento del alumnado. En cualquier caso, las investigaciones sugieren que, para mejorar el rendimiento educativo de los estudiantes, es mejor invertir en la calidad del profesorado que en aumentar su cantidad en las escuelas. Esta consideración puede ser aún más rele-

vante si se trata de alumnos desfavorecidos (especialmente si son inmigrantes) ya que el efecto beneficioso de reducir el número de alumnos por profesor puede desaparecer si los docentes no tienen conocimientos suficientes para tratar con este tipo de alumnado. Asimismo, cabe considerar también determinadas características asociadas a las direcciones de los centros que pueden permitir mejoras en el rendimiento académico.

En suma, la revisión de la literatura muestra cómo los recursos educativos no parecen irrelevantes si se desea mejorar el rendimiento académico del alumnado. De hecho, tras la evaluación de PISA de 2009, en OECD (2012) se indica que los sistemas escolares con mayores niveles de rendimiento invierten fuertemente en educación. Esta conclusión no implica que cualquier inversión sea acertada (por ejemplo, en OECD, 2012, se sugiere priorizar el salario del profesorado sobre la reducción del número de estudiantes por aula), ya que no sólo la cantidad invertida sino el diseño de los programas y el contexto en el que se aplican son muy relevantes. Para nuestro país, resulta imprescindible evaluar las políticas que se practican al respecto y, además, que las nuevas iniciativas ya contemplen en su ejecución la realización de evaluaciones que eliminen la endogeneidad de los estudios para poder aislar qué factores causan los resultados observados (véase al respecto, en este mismo volumen, Santín y Sicilia, en prensa). Dicho tipo de análisis sobre los recursos debe ser complementado, además, con investigaciones sobre los métodos de enseñanza que inciden en el resultado del alumnado (véase De Witte y Van Klaveren, 2011; Schwerdt y Wuppermann, 2011; Lavy, 2011).

Los análisis empíricos, con datos de la evaluación de PISA-2009, que hemos presentado en el apartado anterior, se han orientado a dos objetivos. Por una parte, a describir la distribución de los recursos educativos, en España, entre diferentes colectivos de alumnos. Por otra, a identificar, mediante modelos de regresión multinivel, qué papel juegan los recursos educativos a la hora de determinar la adquisición de la competencia de lectura. Por lo que respecta al primer objetivo, la distribución resulta considerablemente igualitaria (al menos, en función de los grupos que hemos distinguido); de hecho, en algunos casos la dotación de recursos parece ligeramente compensatoria, asignándose más recursos a los centros con un alumnado más desfavorecido. Una pauta que quisiéramos destacar, en tanto que se proyecta sobre el acceso a los recursos de diferentes colectivos de alumnos, es la que aparece al comparar los centros públicos con los centros privados concertados: en los primeros se percibe una mayor escasez de personal de apoyo pero, sin embargo, las ratios de alumnos por profesor y alumnos por clase son menores.

En los análisis relativos al segundo objetivo, efectuados mediante regresiones multinivel, destacan, como es habitual, las variables personales y relativas al origen familiar como muy relevantes a la hora de explicar el nivel de competencias

de los alumnos. Resulta especialmente destacable la fuerte incidencia positiva de la educación infantil.

Centrándonos en el efecto de los recursos escolares de los centros, foco de nuestro interés en este artículo, no resultan significativas la mayor parte de las variables que hemos incorporado al análisis con objeto de identificar los recursos disponibles en el centro. Resulta destacable, sin embargo, un resultado en principio contraintuitivo, debido, a nuestro juicio, a la existencia de causalidad inversa: tanto la ratio de alumnos por aula como la ratio de alumnos por profesor tienen un efecto *positivo* sobre los resultados; es decir, se asocian ratios más altos (más desfavorables) con *mejores* resultados. Conviene señalar que este efecto se produce después de haber controlado por la titularidad del centro y que, por tanto, no es únicamente consecuencia de la existencia de ratios más bajos en los centros de titularidad pública. La posible causalidad inversa viene generada por dos vías: por una parte, la asignación de una mayor cantidad de recursos a los alumnos con peores resultados y, por otra, la atracción de una mayor cantidad de alumnos (que empeoran los ratios) por parte de las escuelas identificadas por los padres como "mejores" y que suelen alcanzar mejores resultados.

Hemos presentado también una réplica del análisis de factores determinantes de los resultados aplicada a diferentes grupos de alumnos (en función del tipo de centro, la categoría socioeconómica familiar, nacionalidad y renta). Las pautas básicas que encontramos para el conjunto de la muestra vuelven a repetirse, con ligeras variaciones, para estos diferentes grupos de alumnos: un efecto no significativo de la mayor parte de variables relativas a los recursos del centro junto con un efecto positivo, atribuible a la causalidad inversa, de las dos variables relativas a la disponibilidad del profesorado.

BIBLIOGRAFÍA

Aaronson, D., Barrow, L. y Sander, W. (2007). "Teachers and student achievement in the Chicago Public High Schools". *Journal of Labor Economics*, vol. 25(1), pp. 95-135.

Agasisti, T. y Longobardi, S. (2012). "Inequality in education: can Italian disadvantaged students close the gap? A focus on resilience in the Italian school system". *Documentos de Trabajo del* IEB 2012/39. Barcelona: Instituto de Economía de Barcelona (IEB).

Akiba, M., LeTendre, G. K. y Scribner, J. P. (2007). "Teacher quality, opportunity gap, and national achievement in 46 countries". *Educational Researcher*, vol. 36(7), pp. 369-387.

Al-Samarrai, S. (2006). "Achieving education for all: how much does money matter?" *Journal of International Development*, vol. 18(2), pp. 179-206.

Altinok, N. (2010). "Do School Resources Increase School Quality?" IREDU *Working Papers* 2010/3. Dijon Cedex: Institut de Recherche sur l'Education (IREDU).

Andersson, C., Johansson, P. y Waldenström, N. (2010). "Do you want your child to have a certified teacher?" *Economics of Education Review*, vol. 30, pp. 65-78.

Anghel, B. y Cabrales, A. (2010). "The Determinants of Success in Primary Education in Spain". *Documentos de trabajo* (FEDEA), núm. 20, pp. 1-65.

Anghel, A., Cabrales, A., Sainz, J. y Sanz, I. (2013). "Publicación de los resultados de las pruebas estandarizadas externas: ¿Tiene ello un efecto sobre los resultados escolares?" *Documentos de trabajo* (FEDEA), núm. 7, pp. 1-33.

Angrist, J. D. y Lavy, V. (1999). "Using Maimonides' rule to estimate the effect of class size of scholastic achievement". *Quarterly Journal of Economics*, vol. 114, pp. 533-576.

Angrist, J. D. y Lavy, V. (2002). "New evidence on classroom computers and pupil learning". *Economic Journal*, vol. 112(482), pp. 735-765.

Angrist, J. D. y Guryan, J. (2008). "Does teacher testing raise teacher quality? Evidence from state certification requirements". *Economics of Education Review*, vol. 27, pp. 483-503.

Angrist, J. D. y Lavy, V. (2009). "The effects of high stakes high school achievement awards: Evidence from a randomized trial". *American Economic Review*, vol. 99(4), pp. 1384-1414.

Arendt, J. N. (2013). "The impact of public student grants on drop-out and completion of higher education evidence from a student grant reform". *Empirical Economics*, vol. 44(3), pp. 1545-1562.

Banerjee, A., Cole, S., Duflo, E. y Linden, L. (2007). "Remedying education: Evidence from two randomized experiments in India". *Quarterly Journal of Economics*, 122(3), pp. 1235-1264.

Barrera-Osorio, F. y Linden, L.L. (2009). "The use and misuse of computers in education: Evidence from a randomized experiment in Colombia". *The World Bank Human Development Network Education Team, Policy Research Working Paper* 4836.

Barrow, L., Markman, L. y Rouse, C. E. (2009). "Technology's edge: The educational benefits of computer-aided instruction". *American Economic Journal. Applied Economics*, vol. 1(1), pp. 52-74.

Bénabou, R., Kramarz, F. y Prost, C. (2009). "The French zones d'éducation prioritaire: Much ado about nothing?" *Economics of Education Review*, vol. 28(3), pp. 345-356.

Bloom, H., Thompson, S. L. y Unterman, R. (2010). *Transforming the high school experience how new york city's new small schools are boosting student achievement and graduation rates*. New York: MDRC. http://www.mdrc.org/sites/default/files/full_589.pdf.

Borland, M. V., Howsen, R. M. y Trawick, M. W. (2005). "An investigation of the effect of class size on student academic achievement". *Education Economics*, vol. 13(1), pp. 73-83.

Bosworth, R. (2011). "Class size, class composition, and the distribution of student achievement". *Education Economics*, DOI: 10.1080/09645292.2011.568698.

Boyd, D., Grossman, P., Lankford, H., Loeb, S. y Wyckoff, J. (2006). "How changes in entry requirements alter the teacher workforce and affect student achievement". *Education Finance and Policy*, vol. 1(2), pp. 176-215.

Boyd, D., Lankford, H., Loeb, S., Rockoff, J. y Wyckoff, J. (2008). "The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools". *Journal of Policy Analysis and Management*, vol. 27(4), pp. 793-818.

Branch, G. F., Hanushek, E.A. y Rivkin S.G. (2012). "Estimating the effect of leaders on public sector productivity: The case of school principals". NBER *Working Paper* 17803. Cambridge, Mass.: National Bureau of Economic Research.

Brewer, D. J. (1993). "Principals and student outcomes: Evidence from U.S. High schools". *Economics of Education Review*, vol. 12(4), pp. 281-292.

Brunello. G. y Checchi, D. (2005). "School quality and family background in Italy". *Economics of Education Review*, vol. 24, pp. 563-577.

Cabras, S. y Tena, J. D. (2013). "Estimación del efecto causal del uso de ordenadores en los resultados de los estudiantes en el test PISA 2012". En INEE (ed.), *PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen* II: *Análisis secundario*. Madrid. INEE.

Calero, J. y Escardíbul, J. O. (2007). "Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003". *Hacienda Pública Española*, vol. 183(4), pp. 33-66.

Calero, J., Choi, Á. y Waisgrais, S. (2010). "Determinantes del riesgo de fracaso escolar en España: una aproximación a través de un análisis logístico multinivel aplicado a PISA-2006". *Revista de Educación*, núm. extraordinario, pp. 225-256.

Cantrell, S., Fullerton, J., Kane, T. J. y Staiger, D. O. (2008). "National Board certification and teacher effectiveness: Evidence from a random assignment experiment". NBER *Working Paper* 14608. Cambridge, MA: National Bureau of Economic Research.

Card, D. y Krueger, A. (1992). "Does school quality matter? Returns to education and the characteristics of public schools in the United States". *Journal of Political Economy*, vol. 100, pp. F1-F40.

Cardoso, E. y Portela Souza, A. (2003). *The impacts of cash transfers on child labor and school attendance in Brazil*. Sao Paolo: Department of Economics. University of Sao Paolo.

Castleman, B. J. y Long, B.T. (2013). "Looking beyond enrolment: the causal effect of need-based grants on college access, persistence, and graduation". NBER *Working Paper* 19306. Cambridge, MA: National Bureau of Economic Research.

Cechini, S. y Madariaga, A. (2011). "Programas de transferencias condicionadas. Balance de la experiencia reciente en América Latina y el Caribe". *Cuadernos de la* CEPAL 95. Santiago de Chile: Naciones Unidas-CEPAL.

Chaudhuri, K. y Maitra, P. (2008). "School attainment, completion, and economic development: A cross-country analysis". *Review of Development Economics*, 12(1), pp. 90-105.

Chingos, M. M. (2012). "The impact of a universal class-size reduction policy: Evidence from Florida's state wide mandate". *Economics of Education Review*, vol. 31, pp. 543-562.

Chingos, M. M. (2013). "Class size and student outcomes: research and policy implications". *Journal of Policy Analysis and Management*, vol. 32(2), pp. 411-438.

Chingos, M. M. y Peterson, P. E. (2011). "It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness". *Economics of Education Review*, vol. 30, pp. 449-465.

Cho, H., Glewwe, P. y Whitler, M. (2012). "Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools". *Economics of Education Review*, vol. 31, pp. 77-95.

Choi, Á. y Calero, J. (2013). "Determinantes del riesgo de fracaso escolar en España en PISA-2009 y propuestas de reforma". *Revista de Educación*, 362, pp. 562-593.

Clark, M. A. (2003). *Education Reform, Redistribution, and Student Achievement: Evidence from the Kentucky Education Reform Act*. Mimeo, Princeton University, January.

Clark, D., Martorell, P. y Rockoff, J. (2009). "School Principals and School Performance". CALDER *Working Paper* 38. Urban Institute. Center for Analysis of Longitudinal Data in Education Research.

Clotfelter, T., Ladd, H. F. y Vigdor, J. L. (2006). "Teacher-student matching and the assessment of teacher effectiveness". *Journal of Human Resources*, vol. 41(4), pp. 778-820.

Coelli, M.B. y Green, D.A. (2012). "Leadership effects: School principals and student outcomes". *Economics of Education Review*, vol. 31(1), pp. 92-109.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., Mcpartlet, J., Mood, A. M., Weinfeld, F.D. y York R. L. (1966). *Equality of Educational Opportunity*. Washington, D.C.: Government Printing Office.

Corallo, C. y McDonald, D. H. (2002). *What Works with Low-Performing Schools: A Review of Research*. http://edvantia.ehclients.com/site-assets/WhatWorks.pdf.

Cordero, J. M., Crespo, E. y Pedraja, F. (2013): "Rendimiento educativo y determinantes según PISA: Una revisión de la literatura en España" *Revista de Educación, n.* 362, pp. 273-297.

Cordero, J. M., Manchón, C. y Simancas, R. (2012). "Análisis de los condicionantes del rendimiento educativo de los alumnos españoles en PISA 2009 mediante técnicas multinivel". *Presupuesto y Gasto Público*, vol. 67(2), pp. 71-96.

Darling-Hammond, L. (2000). "Teacher quality and student achievement: A review of state policy evidence". *Education Policy Analysis Archives*, vol. 8(1), pp. 1-44.

Darling-Hammond, L. (2010). "Recruiting and Retaining Teachers: Turning around the race to the bottom in high-need schools". *Journal of Curriculum and Instruction*, vol. 4(1), pp. 16-32.

Dee, T. S. y West, M. R. (2011). "The non-cognitive returns to class size". *Education Evaluation and Policy Analysis*, vol. 33, pp. 23-46.

Deming, D. y Dynarski, S. (2009). Into college, out of poverty? Policies to increase the postsecondary attainment of the poor. *NBER Working Papers* 15387. Cambridge, MA: National Bureau of Economic Research.

De Witte, K. y Van Klaveren, C. (2011). "How are teachers teaching? A nonparametric approach". *Education Economics*, DOI:10.1080/09645292.2011.560448.

Dewey, J., Husted, T.A. y Kenny, L.W. (2000). "The ineffectiveness of school inputs: A product of misspecification?" *Economics of Education Review*, vol. 19(1), pp.27-45.

Dills, A. K. (2005). "Does cream-skimming curdle the milk? A study of peer effects". *Economics of Education Review*, 24(1), pp. 19-28.

Ding, W. y Lehrer, S. F. (2011). "Experimental estimates of the impacts of class size on test scores: robustness and heterogeneity". *Education Economics*, vol. 19(3), pp. 229-252.

Dolton, P. y Marcenaro-Gutiérrez, O. D. (2011). "If you pay peanuts do you get monkeys? A cross country analysis of teacher pay and pupil performance". *Economic Policy*, vol. 26(65), pp. 5-55.

Downes, T. A. (1992). "Evaluating the impact of school finance reform on the provision of public education: The California case". *National Tax Journal*, vol. 45(4), pp. 405-419.

Dynarski, S. M. (2003). Does aid matter? Measuring the effect of student aid on college attendance and completion. *American Economic Review*, vol. 93(1), March, pp. 279-288.

Dynarski, S. M. y Scott-Clayton, J. E. (2006). The cost of complexity in federal student aid: Lessons from optimal tax theory and behavioral economics. *National Tax Journal*, vol. 59(2), June, pp. 319-356.

Dynarski, S. M. y Scott-Clayton, J. E. (2013). Financial aid policy: Lessons from research. NBER *Working Paper Series* 18710. Cambridge, MA: National Bureau of Economic Research.

Eberts, R., Hollenbeck, K. y Stone, J. (2000). "Teacher performance incentives and student outcomes". *Journal of Human Resources*, vol. 37(4), pp. 913-927.

Fertig, M. (2003), "Who's to Blame? The determinants of German students' achievement in the PISA 2000 study". *IZA Discussion Paper Series* 739. Bonn: Institute for the Study of Labor.

Field, S., Kuczera, M. y Pont, B. (2007). *No more failures: Ten steps to equity in education*. Paris: OECD.

Fredriksson, P., Öckert, B. y Oosterbeek, H. (2013). "Long-term effects of class size". *Quarterly Journal of Economics*, vol. 128(1), pp. 249-285.

Fuchs, T. y Woessmann, L. (2004). "Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and at school". *Brussels Economic Review*, vol. 47(3-4), pp. 359-385.

Fuchs, T. y Woessmann, L. (2007). "What accounts for international differences in student performance? A reexamination using PISA data". *Empirical Economics*, vol. 32(2-3), pp. 433-462.

Fuller, B. (1985). "Raising school quality in developing countries: What investments boost learning?" *The World Bank Discussion Paper. Education and Training Series* EDT7. Washington, D. C.: The World Bank.

Gibbons, S., McNally, S. y Viarengo, M. (2011). "Does additional spending help urban schools? An evaluation using boundary discontinuities". *CEE Discussion Papers* 128. London: Centre for the Economics of Education, LSE.

Glewwe, P., Ilias, N. y Kremer, M. (2003). "Teacher Incentives". NBER *Working Paper* 9671. Cambridge, Mass.: National Bureau of Economic Research.

Glewwe, P., Hanushek, E.; Humpage, S. D. y Ravina, R. (2011). "School resources and educational outcomes in developing countries: a review of the literature from 1990 to 2010", NBER *Working Paper* 17554. Cambridge, Mass.: National Bureau of Economic Research.

Goe, L. y Stickler, L. (2008). "Teacher quality and student achievement: making the most of recent research". *Research and Policy Brief*, March. Washington, D. C.: National Comprehensive Center for Teacher Quality (NCCTQ).

Goolsbee, A. y Guryan, J. (2006). "The impact of internet subsidies in public schools". *The Review of Economics and Statistics*, vol. 88(2), pp. 336-347.

Gupta, S., Verhoeven, M. y Tiongson, E. (1999). "Does higher government spending buy Better results in education and health care?" *IMF Working Paper* 99/21. Washington, D. C.: International Monetary Fund.

Guryan, J. (2001). "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts". *NBER Working Papers* 8269, Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A. (1986). "The economics of schooling: Production and efficiency in public schools". *Journal of Economic Literature*, vol. 24(3), pp.1141-1177.

Hanushek, E. A. (1989). "The effect of differential expenditures on school performance". *Education Researcher*, vol. 18(4), pp. 45-51.

Hanushek, E. A. (1997). "Effects of school resources on economic performance". *Education Evaluation and Policy Analysis*, vol. 19(2), pp.141-164.

Hanushek, E. A. (2003). "The failure of input-based schooling policies". *Economic Journal*, vol. 113 (485), pp. F64-F98.

Hanushek, E. A. (2011). "The economic value of higher teacher quality". *Economics of Education Review*, vol. 30, pp. 466-479.

Hanushek, E. A. y Kimko, D. D. (2000). "Schooling, labor-force quality, and the growth of nations". *American Economic Review*, vol. 90(5), pp. 1184-1208.

Hanushek, E. A. y Luque, J. A. (2003). "Efficiency and equity in schools around the world". *Economics of Education Review*, 22, pp. 481-502.

Hanushek, E. A. y Rivkin, S. G. (2006). "Teacher quality". En: E. A. Hanushek, S. Machin y L. Woessmann (eds.): *Handbook of the Economics of Education*, vol. 2, pp. 1051-1078.

Hanushek, E. A. y Woessmann, L. (2010). "The economics of international differences in educational achievement". *NBER Working Paper* 15949. Cambridge, Mass.: National Bureau of Economic Research.

Hanushek, E. A., Kain, J. F., Markman, J. M. y Rivkin, S. G. (2003). "Does peer ability affect student achievement?" *Journal of Applied Econometrics*, vol. 18(5), pp. 527-544.

Hanushek, E., Kain, J., O'Brien, D. y Rivkin, S. (2005). "The market for teacher quality". *NBER Working Paper* 11154. Cambridge, Mass.: National Bureau of Economic Research.

Harris, D. N. y Sass, T. R. (2011). "Teacher training, teacher quality and student achievement". *Journal of Public Economics*, vol. 95(7-8), pp. 798-812.

Heckman, J. J. (2008). "Schools, skills, and synapses". *Economic Inquiry*, vol. 46(3), pp. 289-324.

Hedges, L. V., Laine, R. D. y Greenwald, R. (1994a). "Does money matter? a me-

ta-analysis of studies of the effects of differential inputs on student outcomes". *Educational Researcher*, vol. 23, pp. 5-14.

Hedges, L. V., Laine, R. D. y Greenwald, R. (1994b). "Money does matter somewhere: a reply to Hanushek". *Educational Researcher*, vol. 23, pp. 9-10.

Heller, D. E. (2008). The impact of student loans on college access. En: S. Baum, M. McPherson y P. Steele (eds.) *The effectiveness of student aid policies: What the research tell us*. Lumina Foundation for Education, pp. 39-67.

Heyneman S. P. y Loxley, W. A. (1982). "Influences on academic performance across high and low-income countries: A re-analysis of IEA data". *Sociology of Education*, vol. 55, pp. 13–21.

Heyneman S. P. y Loxley, W. A. (1983). "The effect of primary school quality on academic achievement across twenty-nine high- and low-income countries". *American Journal of Sociology*, vol. 88, pp. 1162–1194.

Holmlund, H., McNally, S. y Viarengo, M. (2010). "Does money matter for schools?" *Economics of Education Review*, vol. 29, pp. 1154-1164.

Horgan, G. (2009). "'That child is smart because he's rich': the impact of poverty on young children's experiences of school". *International Journal of Inclusive Education*, vol. 13(4), pp. 359-376.

Hossler, D., Ziskin, M., Gross, J. P. K., Kim, S. y Cekic, O. (2009). "Student aid and its role in encouraging persistence". En: J. C. Smart (ed.) *Higher Education: Handbook of Theory and Research*. Springer, pp. 389-426.

Hoxby, C. M. (2000a). "Peer effects in the classroom: Learning from gender and race variation". NBER *Working Paper* 7867. Cambridge, Mass.: National Bureau of Economic Research.

Hoxby C. M. (2000b). "The effects of class size on student achievement: new evidence from population variation". *Quarterly Journal of Economics*, vol. 115(4), pp. 1239-1285.

Iacovou, M. (2002). "Class size in the early years: is smaller really better?" *Education Economics*, vol. 10(3), pp. 261-290.

Ingersoll, R. M. (2001). "Teacher turnover and teacher shortages". *American Educational Research Journal*, vol. 38(3), pp. 499-534.

Instituto de Evaluación (2010). *PISA 2009. Informe español*. Madrid: Ministerio de Educación. Instituto de Evaluación.

Iversen, J. M. V. y Bonesrønning, H. (2013). "Disadvantaged students in the early grades: will smaller classes help them?" *Education Economics*, vol. 21(4), pp. 305-324.

Jenkins, A., Levacic, R. y Vignoles, A. (2006). "Estimating the relationship between school resources and pupil attainment at GCSE". *Research report* RR727, Department for Education and Skills.

Jepsen, C. y Rivkin, S. (2009). "Class size reduction and student achievement. The potential tradeoff between teacher quality and class size". *Journal of Human Resources*, vol. 44(1), pp. 223-250.

Kane, T. J., Rockoff, J. E. y Staiger, D. O. (2008). "What does certification tell us about teacher effectiveness? Evidence from New York City". *Economics of Education Review*, vol. 27(6), pp. 615-631.

Karsten, S., Felix, C., Meijness, W., Roeleveld, J. y van Schooten, E. (2006). "Choosing segregation or integration? The extent and effects of ethnic segregation in Dutch cities". *Education and Urban Society*, vol. 38, pp. 228-247.

Kingdon, G. y Teal, F. (2010). "Teacher unions, teacher pay and student performance in India: A pupil fixed effects approach". *Journal of Development Economics*, vol. 91(2), pp. 278-288.

Krueger, A. B. (1999). "Experimental estimates of education production functions". *Quarterly Journal of Economics*, vol. 114(2), pp. 497-532.

Krueger, A. B. (2003). "Economic considerations and class size". *Economic Journal*, vol. 113 (485), pp. F34-63.

Krueger A. B. y Whitmore D. M. (2001). "The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project STAR". *Economic Journal*, vol. 111, pp. 1-28.

Krueger A. B. y Whitmore D. M. (2002). "Would smaller classes help close the black-white achievement gap?" en: J. Chubb y T. Loveless (eds.): *Bridging the Achievement Gap*. Washington, DC: Brookings Institute Press.

Kukla-Acevedo, S. (2009). "Do teacher characteristics matter? New results on the effect of teacher preparation on students achievement". *Economics of Education Review*, vol. 28, pp. 49-57.

Lavy, V. (2008). "Does raising the principal's wage improve the school's outcomes? Quasi-experimental evidence from an unusual policy experiment in Israel". *Scandinavian Journal of Economics*, vol. 110(4), pp. 639-662.

Lavy, V. (2011). "What makes an effective teacher? Quasi-experimental evidence". NBER *Working Paper* 16885. Cambridge, Mass.: National Bureau of Economic Research.

Lavy, V. (2012). "Expanding school resources and increasing time on task: effects of a policy experiment in Israel on student academic achievement and behavior". NBER *Working Papers* 18369. Cambridge, Mass.: National Bureau of Economic Research.

Leclercq, F. (2005). "The Relationship between educational expenditures and outcomes". DIAL *Working Paper* DT/2005-05. Paris: Développement Institutions & Analyses de Long Terme.

Lee, J. W. y Barro, R. J. (2001). "Schooling quality in a cross-section of countries". *Economica*, vol. 68(272), pp. 465-488.

Leigh, A. (2010). "Estimating teacher effectiveness from two-year changes in students' test scores". *Economics of Education Review*, vol. 29(3), pp. 480-488.

Leuven, E., Lindahl, M., Oosterbeek, H. y Webbink, D. (2007). "The effect of extra funding for disadvantaged pupils on achievement". *Review of Economics and Statistics*, Vol. 89(4), pp. 721-736.

Leuven E., Oosterbeek, H. y Ronning, M. (2008). "Quasi-experimental estimates of the effect of class size on achievement in Norway". *Scandinavian Journal of Economics*, vol. 110(4), pp. 663-93.

Levacic, R., Jenkins, A., Vignoles, A., Steele, F. y Allen, R. (2005). "Estimating the relationship between school resources and pupil attainment at Key Stage 3". *Research report* RR679, Department for Education and Skills.

Lietz, P. (2010). "School quality and student achievement in 21 European countries". En: D. Hastedt y M. von Davier (eds.) *Issues and methodologies in large-scale assessments*. Princeton, NJ: IEA-ETS Research Institute, pp. 57-84.

Loeb, S. y Page, M. E. (2000). "Examining the link between teacher wages and student outcomes: the importance of alternative labor market opportunities and non-pecuniary variation". *Review of Economics and Statistics*, vol. 82(3), pp. 393-408.

López, E., Navarro, E. y Ordóñez, X. G. (2009). "Estudio de variables determinantes de eficiencia a través de los modelos jerárquicos lineales en la evaluación PISA 2006: el caso de España". *Archivos Analíticos de Políticas Educativas*, vol. 17(16), pp. 1-24.

Lupton, R. (2004). "Schools in disadvantaged areas: Recognising context and raising quality". *CASE Paper* 76. London: Centre for Analysis of Social Exclusion (CASE).

Lupton, R. (2005). "Social justice and school improvement: improving the quality of schooling in the poorest neighbourhoods". *British Educational Research Journal*, vol. 31(5), pp. 589-604.

Machin, S., McNally, S. y Silva, S. (2007). "New technology in schools: is there a payoff?" *Economic journal*, vol. 117(522), pp. 1145-1167.

Machin, S., McNally, S. y Costas, M. (2010). "Resources and standards in urban schools". *Journal of Human Capital*, Vol. 4(4), pp. 365-393.

Mancebón, M. J. y Muñiz, M. A. (2003). "Aspectos clave de la evaluación de la eficiencia productiva en la educación secundaria". *Papeles de Economía Española*, vol. 95, pp. 162-187.

McMahon W. (1999). *Education and development: Measuring the social benefits*. Oxford: Oxford University Press.

Mealli, F. y Rampichini, C. (2012). Evaluating the effects of university grants by using regression discontinuity designs. *Journal of the Royal Statistical Society*, vol. 175(3), pp. 775-798.

Mediavilla, M. (2013). "Heterogeneidad en el impacto de la política de becas en la escolaridad secundaria postobligatoria en España: un análisis por subgrupos poblacionales". *Estudios de Economía*, vol. 40(1), pp. 97-120.

Metzler, J. y Woessmann, L. (2010). "The impact of teacher subject knowledge on student achievement: evidence from within-teacher within-student variation". *IZA Discussion Paper Series* 4999. Bonn: The Institute for the Study of Labor.

Mora, T., Escardíbul, J. O. y Espasa, M. (2010). "The effects of regional educational policies on school failure rates in Spain". *Revista de Economía Aplicada*, vol. XVIII(54), pp. 79-106.

Mundel, D. S. (2008). "What do we know about the impact of grants on college students?" In: S. Baum, M. McPherson y P. Steele (eds.) *The Effectiveness of Student Aid Policies: What the Research Tell Us*. Lumina Foundation for Education, pp. 9-37.

Murnane, R. J. y Steele, J. L. (2007). "What is the problem? The challenge of providing effective teachers for all children". *The Future of Children*, vol. 17(1), pp. 15-43.

Nusche, D. (2009). "What works in migrant education? A review of evidence and policy options". OECD *Education Working Papers* 22. Paris: OECD.

OECD (2010). *PISA 2009 Results: What Makes a School Successful? Resources, Policies and Practices (Volume IV)*. Paris: OECD.

OECD (2011). *PISA 2009 Results: Students on Line: Digital Technologies and Performance (Volume VI)*. Paris: OECD.

OECD (2012). *PISA in Focus* 13. Paris: OECD.

Ono, H. (2007). "Does examination hell pay off? A cost-benefit analysis of "ronin" and college education in Japan". *Economics of Education Review*, vol. 26, pp. 271-284.

Papke, L. (2005). "The Effects of Spending on Test Pass Rates: Evidence from Michigan". *Journal of Public Economics*, vol. 89(5-6), pp. 821-839.

Patrinos H. A. (2002). *A review of demand-side financing initiatives in education*. Washington: The World Bank.

Paulsen, M. B. y St. John, E. P. (2002). "Social class and college costs: Examining the financial nexus between college choice and persistence". *Journal of Higher Education*, vol. 73, pp. 189-236.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, D.C.: Economic Policy Institute.

Rivkin, S., Hanushek, E. y Kain, J. (2005). "Teachers, schools, and academic achievement". *Econometrica*, vol. 73, pp. 417-458.

Rockoff, J. (2004). "The impact of individual teachers on student achievement: Evidence from panel data". *American Economic Review*, vol. 94(2), pp. 247-252.

Rouse, C. E. y Krueger, A. B. (2004). "Putting computerized instruction to the test: a randomized evaluation of a 'scientifically-based' reading program". *Economics of Education Review*, vol. 23(4), pp. 323-338.

Roy, J. (2011). "Impact of school finance reform on resource equalization and academic performance: Evidence from Michigan". *Education Finance and Policy*, vol. 6(2), pp. 137-167.

Santín, D. y Sicilia , G. (en prensa) "Evaluar para mejorar: hacia el seguimiento y la evaluación sistemática de las políticas educativas". *Reflexiones sobre el sistema educativo español*. Madrid. Fundación Europea Sociedad y Educación.

Schloter, M., Schwerdt, G. y Woessmann, L. (2011). "Econometric methods for causal evaluation of education policies and practices: A non-technical guide". *Education Economics* Vol. 19(2), pp. 109-137.

Schwerdt, G. y Wuppermann, A. C. (2011). "Is traditional teaching really all that bad? A within-student between-subject approach". *Economics of Education Review*, vol. 30, pp. 365-379.

Scott-Clayton, J. (2012). "On money and motivation. A quasi-experimental analysis of financial incentives for college achievement". *Journal of Human Resources*, vol. 46(3), pp. 614-646.

Sequeira, T. y Robalo, P. (2008). "Schooling quality in a cross section of countries: A replication exercise and additional results". *Economics Bulletin*, vol. 9(9), pp. 1-7.

Sims, D. P. (2009). "Crowding Peter to educate Paul: Lessons from a class size reduction externality". *Economics of Education Review*, 28, pp. 465-473.

Slavin, R. (2010). "Can financial incentives enhance educational outcomes?: Evidence from international experiments". *Educational Research Review*, vol. 5(1), pp. 68-80.

Spencer, M. B., Noll, E. y Cassidy, E. (2005). "Monetary incentives in support of academic achievement: Results of a randomized field trial involving high-achievement, low-resource, ethnically diverse urban adolescences". *Evaluation Review*, vol. 29, pp. 199-222.

St. John, E. P., Paulsen, M. B. y Carter, D. F. (2005). Diversity, college costs, and postsecondary opportunity: An examination of the financial nexus between college choice and persistence for African Americans and Whites. *Journal of Higher Education*, vol. 76, pp. 545-569.

Stater, M. (2009). "The impact of financial aid on college GPA at three flagship public institutions". *American Educational Research Journal*, vol. 46(3), pp. 782-815.

Valdenaire, M. (2006). "Do younger pupils need smaller classes?" Londres: Centre for Economic Performance, London School of Economics.

Vignoles, A., Levacic, R., Walker, J., Machin, S. y Reynolds, D. (2000). "The relationship between resource allocation and pupil attainment: A review". *Working Paper* CEEDP 2. London: Centre for the Economics of Education. London School of Economics and Political Science.

Woessmann, L. (2003). "Schooling resources, educational institutions, and student performance: The international evidence". *Oxford Bulletin of Economics and Statistics*, vol. 65(2), pp. 117-170.

Woessmann, L. (2011). "Cross-country evidence on teacher performance pay". *Economics of Education Review*, vol. 30, pp. 404-418.

Woessmann, L. y West, M. (2006). "Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS". *European Economic Review*, vol. 50(3), pp. 695-736.

Zimmer, R. W. y Toma, E. (2000). "Peer effects in private and public schools across countries". *Journal of Policy Analysis and Management*, vol. 19(1), pp. 75-92.

Zinovyeva, N., Felgueroso, F. y Vázquez, P. (2008). "Immigration and students' achievement in Spain". *Documentos de trabajo* (FEDEA), núm. 37, pp. 1-38.

TERCERA PARTE

REFLEXIONES SOBRE EL DISEÑO Y LA EVALUACIÓN DE LAS POLÍTICAS EDUCATIVAS

IGUALDAD DE OPORTUNIDADES: INCLUSIÓN Y EXCLUSIÓN EDUCATIVA EN ESPAÑA

JORGE CALERO
J. ORIOL ESCARDÍBUL
Universidad de Barcelona e IEB

RESUMEN

En el texto se han seleccionado tres puntos de las trayectorias de las personas, con el objetivo general de conocer cuáles son los factores que conducen a la inclusión o a la exclusión del sistema de educación y formación en España. El primero de ellos es el final de la educación obligatoria (15 años), cuando las competencias de los jóvenes son evaluadas en PISA. El segundo, el periodo entre 18 y 24 años, en el que medimos el abandono prematuro de los estudios, que se sitúa en España en el 24,9%, muy por encima del 10% fijado como objetivo por la Unión Europea. Y, el tercero, la edad adulta entre los 25 y los 64 años. El bajo nivel de competencias de la población adulta es otra característica relevante de nuestro sistema de educación y formación: un 27,5% de la población potencialmente activa española se sitúa en los niveles más bajos de competencias, proporción que casi dobla la de otros países de la OCDE evaluados.

En los análisis correspondientes a los tres puntos resulta muy destacable la fuerte incidencia de las variables relativas al origen sociocultural y económico de las personas. Se ha explorado, también en los tres puntos, el efecto de variables como la condición de inmigrante (de primera y segunda generación), el haber cursado educación infantil, la titularidad del centro educativo, la composición social del centro educativo, los recursos humanos y materiales del mismo, y otras variables relativas al entorno social, territorial y de mercado de trabajo de las personas.

INTRODUCCIÓN

Las trayectorias de las personas en el sistema educativo son diversas en función de una serie de factores, en buena medida condicionados por los ámbitos socioculturales y económicos. Esta diversidad de trayectorias conduce a situaciones de inclusión o de exclusión, no sólo en términos puramente educativos sino, también, relativos al mercado de trabajo y, en general, a la participación social en su conjunto. De hecho, en las últimas décadas, hemos asistido a un incremento de la importancia del sistema educativo en la determinación de las trayectorias sociales.

Las trayectorias educativas presentan, además, un rasgo muy relevante: su carácter acumulativo o de retroalimentación. Así, las personas que adquieren un nivel educativo inicial elevado están en mejor disposición para seguir adquiriendo formación adicional durante toda su vida. Es por ello por lo que, cuando nos referimos a la inclusión y a la exclusión educativa nos estamos refiriendo a un fenómeno con hondas repercusiones en las trayectorias vitales de las personas y con una clara influencia sobre las desigualdades sociales. Adicionalmente, hasta qué punto el sistema educativo consigue ser inclusivo va a condicionar la capacidad de la sociedad para mejorar el *stock* de capital humano y la productividad potencial del sistema productivo.

En el presente texto nos planteamos tres objetivos. En primer lugar, describir las diferencias entre la población incluida y excluida de determinadas áreas del sistema educativo. En segundo lugar, analizar cuáles son los factores que conducen a la exclusión. Y, finalmente, analizar qué efecto tienen las desigualdades previas sobre las trayectorias de los individuos en diferentes puntos del sistema de educación y formación. Los tres objetivos se plantearán en referencia al caso español.

Hemos seleccionado tres puntos de las trayectorias personales para concretar en ellos los objetivos planteados. Primero, el final de la educación obligatoria (15 años), cuando las competencias de los jóvenes son evaluadas en PISA. El criterio que utilizamos en este punto para delimitar la exclusión es la adquisición de un nivel de competencias del nivel 1 o inferior (en 2012 un 23,6% de los jóvenes en España se encontraban en esta situación). Segundo, el periodo entre los 18 y 24 años; en este periodo podemos medir el abandono prematuro o falta de continuidad en los estudios, uno de los principales (si no el principal) motivos de exclusión educativa. El nivel de abandono prematuro era en España, en 2013, del 24,9%, muy superior al del resto de países europeos. Y, tercero, la edad adulta potencialmente activa entre los 25 y los 64 años. El criterio utilizado para establecer la situación de exclusión en este caso es la adquisición de competencias del nivel 1 o menor en PIAAC-2012. Un 27,5% de la población relevante española se encuentra en los niveles mencionados.

El estudio de la población excluida del sistema de educación y formación, en los tres puntos que hemos comentado, resulta de especial relevancia en el caso de

España. Y ello por dos motivos. En primer lugar, por la historia del sistema educativo español, que sólo en las tres últimas décadas ha superado el fuerte elitismo que le caracterizaba y que ha determinado un muy bajo nivel educativo medio de los actuales progenitores de los jóvenes en edad de cursar estudios. Y, en segundo lugar, debido a la necesidad de mejorar la capacidad de nuestro sistema productivo, con vistas a la superación de la crisis económica que comenzó en 2007, por medio de la reorientación de la producción hacia ámbitos de mayor valor añadido.

La estructura del texto es la siguiente. En el primer apartado nos centramos en la adquisición de competencias en el final de la educación obligatoria. En el segundo, abordamos el fenómeno del abandono prematuro. A continuación trataremos el nivel de competencias adquiridas por la población adulta. Finalmente, presentaremos las conclusiones del conjunto del texto.

EL BAJO RENDIMIENTO EN LAS EVALUACIONES DE PISA. CARACTERIZACIÓN Y FACTORES ASOCIADOS

PRESENTACIÓN

En este apartado se analizan los determinantes del nivel de competencias de los estudiantes en el ámbito de las matemáticas. En concreto, se pretende conocer los factores que favorecen que los alumnos se sitúen en niveles de baja competencia. De este modo, se caracterizan los estudiantes que obtienen menores resultados y que, en consecuencia, tienen una mayor probabilidad de no finalizar los estudios obligatorios (están en situación de riesgo de fracaso escolar).

Investigaciones recientes relacionan la adquisición de competencias con el rendimiento económico en el mercado de trabajo, tanto de un modo directo (a mayores competencias mejores condiciones laborales) como indirecto (un mayor nivel de competencias favorece que los alumnos progresen en el sistema educativo y, en consecuencia, que obtengan un mayor rendimiento económico a lo largo de su trayectoria laboral).

En el primer caso, diversos estudios muestran que los individuos que tienen un mayor rendimiento en pruebas de evaluación de competencias obtienen unos mayores ingresos en el mercado laboral. En estos análisis resulta habitual estimar una función de ingresos de tipo minceriano, que incluye como variables explicativas la cantidad de educación recibida y la experiencia laboral de los trabajadores (y otros factores que también pueden influir en los ingresos), a la que se añade una medida de las capacidades cognitivas individuales (véase una revisión en Hanushek y Woessmann, 2008). Este tipo de análisis utiliza tanto datos de pruebas específicas

de los países (en especial Estados Unidos, aunque también el Reino Unido) como provenientes de evaluaciones internacionales de competencias (especialmente de matemáticas y comprensión lectora), básicamente del *Programme for International Student Assessment* (PISA), aunque también del *International Mathematics and Science Study* (TIMSS) y *la International Adult Literacy Survey* (IALS). Al respecto, véase un reciente estudio, para diversos países a partir de la base de datos de PIAAC-2012, en Hanushek *et al.* (2015).

En el ámbito específico de los jóvenes, y con datos de competencias de PISA y TIMSS, la revisión de estudios de Lee y Newhouse (2013) muestra que, para diversos países, las cohortes con mayores puntuaciones tienen una menor probabilidad de estar desempleados (los estudios tienen en cuenta el PIB per cápita y el nivel de empleo juvenil pasado, entre otras variables de control). Asimismo, los países que experimentaron las mejoras más rápidas en las puntuaciones de las pruebas experimentaron una mayor reducción en las tasas de desempleo juvenil. En su estudio propio, Lee y Newhouse (2013) revelan que la obtención de resultados más elevados en PISA y TIMSS conduce a mayores tasas de formación continua, menores niveles de desempleo y a ocupaciones de mayor calidad (en términos salariales y no salariales).

Finalmente, la evidencia expuesta sobre el efecto positivo directo de la adquisición de competencias sobre el rendimiento laboral de los individuos (aproximación microeconómica) también se ha empezado a analizar a nivel macroeconómico. En este caso, diversos estudios relacionan los resultados en pruebas de competencias internacionales (como PISA y TIMSS) con el crecimiento económico de los países (véase una revisión en Hanushek y Woessmann, 2010, así como un análisis para los países en desarrollo en Hanushek, 2013).

En el segundo caso, con datos de la evaluación de PISA, Sun (2008) muestra la positiva relación existente entre tener un nivel alto de competencias y permanecer en el sistema escolar. Asimismo, como señala Hansen (2010), un mayor nivel de competencias aumenta la probabilidad de que los estudiantes que han abandonado los estudios vuelvan a continuar su formación. En consecuencia, puede concluirse que la adquisición de competencias incide positivamente en el mercado de trabajo también de manera indirecta, ya que aquellos que alcanzan mayores niveles educativos (y reciben formación continua) tienen mejores condiciones laborales (en tipo de ocupación o nivel salarial, por ejemplo).

El análisis que presentamos aquí, centrado, como hemos dicho, en el conjunto de estudiantes con competencias más bajas, utiliza los datos correspondientes a la evaluación de competencias de PISA que desarrolla la OCDE para alumnos de 15 años de edad, en la convocatoria de 2012. El estudio se desarrolla en el ámbito rela-

cionado con la competencia matemática por ser el área evaluada en profundidad en la prueba de ese año.

Un nivel bajo de competencias se define a partir de los niveles que establece la OCDE en la evaluación de PISA que, en la evaluación de matemáticas, se establece en no alcanzar el nivel 2 (en esta competencia se han establecido seis niveles de rendimiento, más un séptimo que agrupa al alumnado que no alcanza el nivel 1). De este modo, se asume que los alumnos que no alcanzan el nivel 2 de matemáticas (están por debajo de 420,1 puntos) carecen de las competencias suficientes para participar activamente en situaciones de la vida cotidiana o laboral (OECD, 2010). Obviamente, este umbral tiene un cierto componente arbitrario y, en todo caso, resulta discutible su adecuación a la diversidad de realidades educativas y laborales de los países de la OCDE. Sin embargo, aceptando su carácter convencional, lo consideramos un buen punto de partida instrumental para nuestro análisis.

En el conjunto de los países de la OCDE, el 23% de los alumnos de 15 años se encuentra en los niveles más bajos de rendimiento en matemáticas (niveles inferiores a 2). Dicho porcentaje asciende al 23,9% en el promedio de la Unión Europea (de 25 países) y 23,6% para el caso de España (INEE, 2013). No hay que olvidar que los *Objetivos de la Estrategia Educación y Formación* 2020 de la Unión Europea establecen que el porcentaje de alumnos de 15 años con un nivel de competencia en las escalas de PISA inferior a 2 se sitúe por debajo del 15% (Diario Oficial de la UE, 2009).

DATOS Y METODOLOGÍA

Como se ha señalado anteriormente, para desarrollar el estudio propuesto se utiliza la base de datos de la evaluación de PISA de 2012. Este programa evalúa desde el año 2000, con una periodicidad trianual, los resultados de los sistemas educativos de diversos países. En la edición de 2012 participaron 65 países (34 de los cuales pertenecen a la OCDE). En el caso español, la muestra de alumnos seleccionada es representativa a nivel estatal, así como para todas las Comunidades Autónomas excepto Canarias, Castilla-La Mancha y la Comunidad Valenciana. La muestra estatal comprende 25.313 alumnos en 902 escuelas.

La estructura de la base de datos de PISA es compleja, tanto por los métodos de selección del alumnado como de obtención de los resultados individuales. Con respecto a la selección, en primer lugar se escogen las escuelas participantes y, posteriormente, los estudiantes. Los alumnos proporcionan información relacionada con aspectos personales, familiares y de actitud hacia la escuela y el aprendizaje. La información relativa a los centros (características, recursos, modo de gestión y procesos de enseñanza-aprendizaje) la facilitan los directores de los mismos. En

cuanto a los resultados de los estudiantes, la evaluación de PISA no proporciona la puntuación directa de los alumnos en las pruebas, sino que se refiere a un rango de valores plausibles para cada estudiante, calculados en función de sus respuestas, la dificultad de las preguntas y un conjunto de variables condicionantes, relacionado con su entorno familiar y escolar. Estos resultados, a su vez, tienen ciertas calibraciones y adaptaciones (para mejorar la calidad de los mismos) de ámbito nacional e internacional, así como un tratamiento para la falta de respuesta (*missing values*). El rango de valores plausibles reduce el error de medición, así como el sesgo de inferencia debido a la medición de un inobservable (la capacidad del alumno) a través de un test con un reducido número de preguntas (véase OECD, 2010).

A partir de dichos resultados, se crean niveles de conocimiento para cada ámbito de análisis, como se ha indicado anteriormente. En nuestro estudio, dichos niveles se transforman en una variable de carácter dicotómico que toma valor "1" cuando los alumnos no alcanzan el nivel 2 de conocimiento y "0" a partir de dicho nivel. No alcanzar dicho nivel implica que los estudiantes son capaces de llevar a cabo con éxito las tareas del nivel 1, pero es poco probable que puedan completar las tareas de niveles superiores. En consecuencia, sólo pueden resolver cuestiones que requieren operaciones matemáticas directas relacionadas con contextos familiares. Dicho nivel de competencia matemática es un reflejo de escasas competencias adquiridas y riesgo de fracaso escolar y, como se ha señalado en el apartado anterior, dificulta participar activamente en situaciones de la vida cotidiana y actividades laborales habituales.

El sistema bietápico de selección de la muestra hace que no se cumpla con el principio de independencia de las variables, al ser los estudiantes de cada centro más similares (en cuanto a composición social por ejemplo) que si se escogieran aleatoriamente (Hox, 1995). Dada la estructura jerárquica de los datos, la metodología empírica aplicada en este trabajo está basada en técnicas multinivel. Asimismo, como se estudian los factores que inciden sobre la probabilidad de que un alumno no alcance el nivel 2 de competencias en la evaluación de PISA, la variable dependiente del modelo es dicotómica. En consecuencia, se utiliza un modelo logístico multinivel que se explicita en las ecuaciones (1) a (3).

$$\log\left[\frac{p_{ij}}{(l-p_{ij})}\right] = \beta_{oj} + \sum_{k=1}^{n} \beta_{1j} X_{kij} + \varepsilon_{ij} \qquad \text{Nivel 1 (estudiantes)} \qquad (1)$$

$$\beta_{oj} = \gamma_{00} + \sum_{l} \gamma_{01} Z_{lj} + \mu_{oj} \qquad \text{Nivel 2 (centros educativos)} \quad (2)$$

$$\beta_{1j} = \gamma_{10} \qquad \text{Nivel 2 (centros educativos)} \quad (3)$$

En el modelo, la probabilidad de que se cumpla el suceso (que el alumno tenga un bajo nivel de competencias) se define como p_{ij} =P(Y_{ij} =1), con Y_{ij} como una variable binaria para el estudiante "i" del centro "j". Dicha p_{ij} puede modelizarse mediante una función logística (1), con los parámetros β desarrollados en (2) y (3). X_{kij} es un conjunto de "k" variables independientes asociadas a los "i" alumnos, y Z_{lj} representa un conjunto de "l" variables a nivel de centro. Las tres ecuaciones anteriores pueden resumirse en (4), que se obtiene de reemplazar en la ecuación (1) los coeficientes β de las ecuaciones (2) y (3). De este modo, se distingue una serie de efectos fijos $\gamma_{oo}+\gamma_{1o}X_{kij}+\gamma_{01}Z_{lj}$ de una serie de efectos aleatorios ($\mu_{0j}+\varepsilon_{ij}$).

$$\log\left[\frac{p_{ij}}{(l-p_{ij})}\right]=\gamma_{oo}+\gamma_{1o}X_{kij}+\gamma_{01}Z_{lj}+\mu_{0j}+\varepsilon_{ij} \qquad (4)$$

Como muestra la ecuación (4), la estimación logística permite obtener el efecto de cada variable sobre la probabilidad relativa de que se cumpla un suceso (p_{ij}) respecto a que no se cumpla (1-p_{ij}). Por último, debido a la existencia de cinco valores plausibles, en el análisis empírico deben desarrollarse cinco estimaciones logit (una para cada uno) y obtener los coeficientes y errores estándar de los mismos como promedios de las cinco estimaciones realizadas.

El conjunto de parámetros (γ) se computa por el método de la máxima verosimilitud y, en su cálculo, se controlan posibles problemas de heteroscedasticidad y multicolinealidad. En la tabla de resultados, además del valor del coeficiente de la regresión, se adjunta el valor de exp(γ), donde "exp" equivale a la función exponencial. Este coeficiente muestra los cambios en la *odds-ratio*, es decir, en la probabilidad relativa (sin logaritmos) de una variación en las variables explicativas. En concreto, como muestra la ecuación (5), se considera el cambio en una unidad en les variables explicativas continuas y el cambio de categoría en las discretas:

$$\left[\frac{p_{ij}}{(l-p_{ij})}\right]=e^{\gamma_{00}+\gamma_{10}X_{kij}+\gamma_{01}Z_{lj}+\mu_{oj}+\varepsilon_{ij}} \qquad (5)$$

CARACTERIZACIÓN DE LOS ALUMNOS CON MENORES NIVELES DE COMPETENCIA MATEMÁTICA EN LA EVALUACIÓN DE PISA

Antes de presentar los resultados de la estimación del modelo, en el Cuadro 1 se exponen relaciones existentes entre la variable dependiente y las variables expli-

cativas incluidas en el modelo. En la tabla se muestra el porcentaje de individuos con un bajo nivel de competencias en función de los distintos valores que toman las variables independientes consideradas. En las variables discretas se considera cada categoría de dicho tipo de variables; en las variables continuas se tienen en consideración tres niveles de la misma (primer y último quintil, así como el valor promedio).

Debe advertirse que las relaciones entre variables explicativas y la dependiente que se presentan no pretenden sugerir una relación estadística significativa; ésta se presenta en el apartado de resultados mediante el método de regresión logística, en el que se controla el efecto de las variables más relevantes. Las relaciones que se observan en el Cuadro 1 son las siguientes:

- *Variables personales*. Se constata un incremento sustancial del porcentaje de alumnos con bajas competencias en aquellos que han repetido curso (especialmente si han repetido en primaria y secundaria, donde su nivel asciende a más del 75%), en los estudiantes que se ausentan de la escuela (se da el triple de individuos con un nivel bajo de competencias entre aquellos que más faltan con respecto a los que no se ausentan) y en los que no han cursado educación infantil más de un año (los alumnos con bajas competencias en este grupo suponen casi el doble con relación al grupo de los que sí la han cursado). La condición de inmigrante también se relaciona con el porcentaje de alumnos con un bajo nivel de competencias: éste es claramente inferior entre los nativos (20,7%), que entre los inmigrantes nacidos en España (35,6%) y, especialmente, entre aquellos nacidos en el extranjero (44,1%). No parecen existir diferencias importantes en el porcentaje de estudiantes con un nivel de competencia matemática inferior a 2 en función de la edad y el género.
- *Variables familiares*. Destaca la variabilidad del porcentaje de alumnos con un bajo nivel de competencia matemática en función del estatus socioeconómico y cultural de los padres. Esta variable se define teniendo en cuenta el nivel educativo y ocupación de los padres y de diversas variables relacionadas con el hogar, tales como la posesión de bienes materiales, culturales y de estudio en el hogar. En el quintil más bajo de esta variable dicho porcentaje es del 41,3%, en el quintil medio 21,1% y en el nivel más alto sólo del 7,5%. Apenas se constatan diferencias en función del idioma hablado en el hogar o si la familia es monoparental.
- *Características escolares*. El porcentaje de estudiantes con un nivel inferior a 2 en la evaluación de matemáticas de PISA en centros públicos es casi tres ve-

ces superior al existente en centros privados independientes. Una variable que muestra resultados muy dispares es el promedio de años de escolarización de los padres de los alumnos de los centros. Así, si la escuela tiene, en promedio, padres con estudios obligatorios, el 65,4% de los estudiantes están en los niveles inferiores de competencia matemática. Dicho porcentaje desciende al 9,3% si el centro tiene, en promedio, padres que han alcanzado los estudios universitarios. Otra característica relevante es el porcentaje de inmigrantes en la escuela. En este caso, cuando existe una concentración superior al 30% se eleva el porcentaje de estudiantes con bajo nivel de competencias a más de un tercio (casi el doble del porcentaje de individuos en dicho nivel presentes en escuelas sin presencia de inmigrantes). Las diferencias en el volumen de estudiantes con bajas competencias matemáticas no son importantes si se tienen en cuenta las variables referidas al número de alumnos en clase y en el centro escolar, así como al porcentaje de chicas en la escuela. Tampoco se vislumbra un posible efecto del nivel de competencia entre centros.

- *Variables de gestión escolar*. Se constatan diferencias importantes en cuanto al volumen de estudiantes con un bajo nivel de competencia matemática al considerar la ratio alumnos por profesor (a mayor ratio menor porcentaje), así como si se publicitan los resultados del centro (en caso afirmativo hay un mayor nivel de alumnos con bajas competencias). Las variables de autonomía muestran que se dan menores porcentajes de alumnos poco cualificados en centros con autonomía de gestión (en especial en el caso de la gestión curricular). No se observan diferencias importantes en el porcentaje de alumnos con un menor nivel de competencia matemática al considerar la dotación de recursos educativos materiales y de profesorado, así como si en los centros existen prácticas de agrupación de estudiantes o distintos métodos de evaluación.
- *Tecnologías de la Información y la Comunicación* (TIC). Iniciarse a una edad más tardía en el uso de las TIC se relaciona de manera importante con un mayor porcentaje de alumnos en los niveles más bajos de competencia matemática. Sin embargo, no puede establecerse una relación entre la disponibilidad de TIC en la escuela y dicho porcentaje.
- *Variables de localización*. Existe algún tipo de relación entre las variables de localización y el porcentaje de alumnado con bajas competencias matemáticas, tanto si se tiene en cuenta el tamaño del municipio de residencia de los estudiantes como si consideramos la Comunidad Autónoma. En el primer caso se observa un menor porcentaje en municipios de más de un millón de

habitantes (15,8%, la mitad que en municipios de hasta tres mil habitantes); en el segundo, dicho porcentaje es menor (por debajo del 20%) en las Comunidades de Asturias, Castilla y León, Madrid, Navarra y el País Vasco.

CUADRO 1. PORCENTAJE DE ALUMNOS EN NIVELES INFERIORES A 2 EN COMPETENCIA MATEMÁTICA

VARIABLES	NIVEL DE COMPETENCIA BAJO
VARIABLES PERSONALES	
EDAD (15 años)	24,2
(16 años)	22,9
MUJER (sí)	25,2
(no)	22,2
EDINFA (ha cursado educación infantil más de un 1 año)	20,3
(ha cursado ed. infantil menos de un 1 año o no ha cursado)	38,6
VARIABLES PERSONALES	
REPNO (no ha repetido curso)	9,2
REPPRIM (ha repetido primaria)	50,6
REPSEC (ha repetido secundaria)	40,2
REPRIMSEC (ha repetido primaria y secundaria)	75,5
ABSENT 0 (no ha faltado a clase en las últimas 2 semanas)	18,5
ABSENT 1 (faltó a clase 1-2 días)	32,6
ABSENT 2 (faltó más de dos días)	53,8
INMIG 0 (nativo)	20,7
INMIG 1 (inmigrante de primera generación: nacido en extranjero)	44,1
INMIG 2 (inmigrante de segunda generación: nacido en España)	35,6
VARIABLES FAMILIARES	
IDIOMA (idioma en casa igual al del test de PISA)	21,9
(idioma en casa distinto al del test)	27,9
MONOPARENTAL (sí)	25,1
(no)	21,8

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 1. PORCENTAJE DE ALUMNOS EN NIVELES INFERIORES A 2 EN COMPETENCIA MATEMÁTICA

VARIABLES	NIVEL DE COMPETENCIA BAJO
ESCS (índice de nivel socioeconómico y cultural): quintil más bajo	41,3
ESCS (índice de nivel socioeconómico y cultural): quintil medio	21,1
ESCS (índice de nivel socioeconómico y cultural): quintil más alto	7,5
VARIABLES DE CARACTERÍSTICAS ESCOLARES	
PÚBLICO (centro público)	28,7
CONCERT (centro privado concertado)	14,7
PRIVIND (centro privado independiente)	10,4
schlsize (número de alumnos en la escuela: quintil más bajo)	25,1
(número de alumnos en la escuela: quintil medio)	22,0
(número de alumnos en la escuela: quintil más alto)	20,0
ALUMCLAS (número de alumnos por clase: quintil más bajo)	25,0
VARIABLES DE CARACTERÍSTICAS ESCOLARES	
(número de alumnos por clase: quintil medio)	22,9
(número de alumnos por clase: quintil más alto)	23,4
PCGIRLS (% chicas en la escuela: quintil más bajo)	23,8
(% chicas en la escuela: quintil medio)	25,2
(% chicas en la escuela: quintil más alto)	22,8
AÑOSEDCEN (años medios escolarización padres: est. obligatorios)	65,4
Años medios escolarización padres: universidad	9,3
NONAT 0 (0% inmigrantes en la escuela)	21,0
NONAT 1 (% inmigrantes >0% y ≤10%)	19,7
NONAT 2 (% inmigrantes >10% y ≤20%)	27,2
NONAT 3 (% inmigrantes >20% y ≤30%)	28,9
NONAT 4 (% inmigrantes >30% y ≤40%)	37,0
NONAT 5 (% inmigrantes >40%)	43,3
COMPET 0 (no compite con ninguna escuela)	27,3

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 1. PORCENTAJE DE ALUMNOS EN NIVELES INFERIORES A 2 EN COMPETENCIA MATEMÁTICA

VARIABLES	NIVEL DE COMPETENCIA BAJO
VARIABLES DE CARACTERÍSTICAS ESCOLARES	
COMPET 1 (compite con una escuela)	31,0
COMPET 2 (compite con dos o más escuelas)	21,1
VARIABLES DE GESTIÓN ESCOLAR Y DEL APRENDIZAJE	
ALUMPROF (ratio alumnos/profesorado: quintil más bajo)	26,3
(ratio alumnos/profesorado: quintil medio)	24,6
(ratio alumnos/profesorado: quintil más alto)	15,2
SCMATEDU (índice de calidad recursos educativos: quintil más bajo)	25,4
(índice de calidad recursos educativos: quintil medio)	24,8
(índice de calidad recursos educativos: quintil más alto)	20,1
PROFESORADO (índice de falta de profesorado: quintil más bajo)	21,7
(índice de falta de profesorado: quintil medio)	23,8
(índice de falta de profesorado: quintil más alto)	23,7
AGRUP 0 (no agrupación de los estudiantes en clase)	19,3
AGRUP 1 (agrupación en algunas clases)	22,8
AGRUP 2 (agrupación en todas las clases)	25,2
PUBLIRE (se publicitan los resultados del centro: sí)	23,4
(se publicitan los resultados del centro: no)	32,8
RESPRES (índice autonomía gestión presupuestaria: quintil más bajo)	29,6
(índice de autonomía en la gestión presupuestaria: quintil medio)	28,5
(índice de autonomía en la gestión presupuestaria: quintil más alto)	11,8
RESPCUR (índice de autonomía gestión curricular: quintil más bajo)	29,9
(índice de autonomía en la gestión curricular: quintil medio)	23,0
(índice de autonomía en la gestión curricular: quintil más alto)	17,9
EVALU 1 (existe autoevaluación o evaluación interna: sí)	23,8
(existe autoevaluación o evaluación interna: no)	22,3

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 1. PORCENTAJE DE ALUMNOS EN NIVELES INFERIORES A 2 EN COMPETENCIA MATEMÁTICA

VARIABLES	NIVEL DE COMPETENCIA BAJO
VARIABLES DE GESTIÓN ESCOLAR Y DEL APRENDIZAJE	
EVALU 2 (en el centro existe evaluación externa: sí)	23,3
(en el centro existe evaluación externa: no)	25,3
VARIABLES TIC	
EDADTIC (edad de inicio en las TIC: 5 años)	17,9
(edad de inicio en las TIC: 11 años)	31,4
(edad de inicio en las TIC: 16 años)	35,9
TICESC (disponibilidad TIC en la escuela: quintil más bajo)	24,4
(disponibilidad TIC en la escuela: quintil medio)	18,2
(disponibilidad TIC en la escuela: quintil más alto)	30,1
VARIABLES DE LOCALIZACIÓN	
POBLACIÓN 0 (hasta 3.000 habitantes)	31,4
POBLACIÓN 1 (3.000-15.000 habitantes)	28,0
POBLACIÓN 2 (15.000-100.000 habitantes)	22,8
POBLACIÓN 3 (100.000-1.000.000 habitantes)	22,6
POBLACIÓN 4 (más de 1 millón de habitantes)	15,8
ANDALUCÍA	27,5
ARAGÓN	21,3
ASTURIAS	18,7
BALEARES	26,6
CANTABRIA	22,0
CASTILLA Y LEÓN	14,9
CATALUÑA	20,2
EXTREMADURA	33,0
GALICIA	22,0
LA RIOJA	20,0
MADRID	17,8

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 1. PORCENTAJE DE ALUMNOS EN NIVELES INFERIORES A 2 EN COMPETENCIA MATEMÁTICA

VARIABLES	NIVEL DE COMPETENCIA BAJO
VARIABLES DE LOCALIZACIÓN	
MURCIA	31,1
NAVARRA	13,7
PAÍS VASCO	15,6
OTRAS COMUNIDADES AUTÓNOMAS	29,1

RESULTADOS DE LA ESTIMACIÓN DEL MODELO

Los resultados de la estimación se presentan en el Cuadro 2. En la primera columna se muestra el coeficiente obtenido en la regresión logística, así como el error estándar del mismo. El grado de significatividad se muestra en asteriscos. En la segunda columna se expone el resultado de la *odds-ratio*, que indica el efecto que generan en la probabilidad de que los alumnos estén en un bajo nivel de competencias matemáticas (respecto a un nivel superior) los cambios en las variables explicativas. En una variable categórica, la *odds-ratio* muestra la variación en la probabilidad relativa del cambio de categoría (por ejemplo chico respecto a chica). En una variable continua, la *odds-ratio* señala el cambio en la probabilidad relativa de un aumento en una unidad del valor de dicha variable.

Las variables personales muestran los siguientes resultados. La edad no incide en la probabilidad analizada (un resultado esperable dado que todos los estudiantes están en torno a los 15 años). En cuanto al género, las alumnas tienen una mayor probabilidad relativa de estar en el grupo de competencias bajas en matemáticas que los alumnos (dicha probabilidad aumenta un 65,3%). Haber cursado educación infantil más de un año reduce la probabilidad relativa de tener un bajo nivel de competencias un 32,2% (la *odds-ratio* es 0,678). Las variables referidas a la repetición de curso tienen un fuerte impacto positivo sobre la probabilidad relativa de pertenecer al grupo de estudiantes con menores niveles de competencia matemática. Así, con respecto a no repetir curso, la repetición en primaria multiplica por 8,3 dicha probabilidad, en secundaria por 5,4 y la repetición en ambos niveles educativos la incrementa casi 27 veces. Así, la repetición aparece claramente relacionada con un muy bajo nivel de competencias, en especial, si se produce en los dos niveles educativos obligatorios. El absentismo escolar también aumenta la probabilidad de que los estudiantes se sitúen en niveles bajos de competencias. Como cabe esperar, a mayor

nivel de absentismo peor resultado en las evaluaciones: con respecto a no haber faltado a clase dos semanas antes de la prueba, la probabilidad relativa de pertenecer a los niveles inferiores aumenta un 43,4% si el alumno no ha asistido como máximo dos días y se multiplica por 2 si la ausencia es más prolongada. El último factor individual considerado es la inmigración. Los resultados muestran que la condición de inmigrante sólo incide, en la probabilidad relativa de estar entre aquellos con un nivel de competencia matemática inferior a 2, si el estudiante es inmigrante de primera generación (es decir, nacido en el extranjero). En este caso, dicha probabilidad se incrementa un 61,7%. Sin embargo, no existen diferencias estadísticamente significativas entre alumnos nativos y aquellos nacidos en España con padres extranjeros.

Con respecto a las variables familiares, no se constata una incidencia estadísticamente significativa de las variables relacionadas con el idioma hablado en el hogar. Sin embargo, sí incide claramente el estatus socioeconómico y cultural del hogar (que reduce la probabilidad de que un estudiante esté en el grupo de aquellos con un nivel de competencia matemática inferior a 2), así como si el alumno vive en una familia monoparental (en este caso dicha probabilidad se reduce en un 20,7% con respecto a otro tipo de familias).

En cuanto a las características del centro, sólo resultan estadísticamente significativos los factores relacionados con la condición social de los padres de los estudiantes, reflejada mediante los años promedio de escolarización y el porcentaje de inmigrantes en el centro escolar. De este modo, un año más de escolarización del conjunto de padres de los alumnos en la escuela reduce la probabilidad de que los estudiantes del centro obtengan resultados muy bajos en la evaluación de PISA un 8,5%. En cuanto a la inmigración, resulta perjudicial para la adquisición de competencias estar en un centro con una concentración de alumnos inmigrantes superior al 40% (la probabilidad de estar en los menores niveles de competencia se dobla con respecto al resto de niveles). Para el resto de características, no son estadísticamente significativas ni la titularidad del centro (que sea público o privado), ni el número de alumnos en la escuela o por clase, así como el porcentaje de chicas en el centro y si la escuela compite con otras por el alumnado. Tampoco resultan relevantes la mayor parte de variables relacionadas con la gestión del centro: ratio de alumnos por profesor, índices relacionados con la falta de profesorado o medios materiales, la política de agrupación de alumnos y el sistema de evaluación; con respecto a las variables relacionadas con la autonomía de gestión de las escuelas, no es significativa la variable relacionada con la gestión curricular pero sí aquella referida a la gestión presupuestaria (en este caso, la presencia de autonomía reduce ligeramente el porcentaje de alumnos con bajos niveles de competencia, si bien la variable es significativa sólo al 10%). Por último, sí resulta significativo que el centro publicite

los resultados: aumenta la probabilidad relativa de que los alumnos estén en los niveles más bajos de competencias un 34,1%.

El efecto de la titularidad de centro es un resultado habitual en este tipo de estudios (véase una revisión en Calero y Escardíbul, 2007). Así, la relación bivariante que observábamos anteriormente (los centros públicos tienen un porcentaje mayor de individuos con un bajo nivel de competencias) no está generada por una relación estadísticamente significativa entra ambas variables. Dicha relación se produce porque otras variables inciden sobre ambas. Nos referimos a variables como el origen socioeconómico de los estudiantes, diferente en los centros según su titularidad (como se constata con la significatividad de los años de escolarización promedio de los padres o el porcentaje de inmigrantes en las escuelas). Con respecto al resto de variables, también es habitual la no incidencia de variables relacionadas con los recursos (físicos y humanos), debido a las pocas diferencias existentes en cuanto a su dotación en los centros españoles (véase Calero y Escardíbul, 2014). Otras variables asociadas con la gestión pueden relacionarse con la titularidad de los centros y, en consecuencia, su significatividad es escasa, ya que los factores relevantes son de tipo socioeconómico como se ha señalado con anterioridad.

Con respecto a las TIC, es significativa la edad de inicio en su uso, de modo que cuanto antes se empieza, menor es la probabilidad de que el estudiante pertenezca al grupo de alumnos con menores niveles de competencia. Los resultados no son concluyentes por lo que respecta a la disponibilidad de recursos TIC en la escuela. En la tabla se observa un resultado contraintuitivo (debido quizás a la presencia de una causalidad inversa), de modo que se relacionan positivamente la presencia de más recursos y la probabilidad de que los alumnos obtengan bajas cualificaciones. Ahora bien, otras variables relacionadas con la dotación de TIC en los centros, como la ratio de ordenadores para la enseñanza por alumno, o con el tiempo de uso en las aulas, no tienen incidencia estadísticamente significativa[1].

Por último, se analiza la localización de los estudiantes, tanto teniendo en cuenta el tamaño del municipio como la Comunidad Autónoma de residencia. Con respecto al primero, comparado con residir en los municipios más pequeños (con menos de 3.000 habitantes), los alumnos que viven en ciudades con un mayor número de habitantes (salvo en aquellos con más de un millón de habitantes) ven reducida la probabilidad relativa de obtener las peores cualificaciones. La reducción está en torno al 33%-37%. En cuanto a las Comunidades Autónomas, se analiza el efecto de residir en una de las Comunidades con una muestra independiente en la evaluación de PISA comparado con vivir en las 3 sin muestra independiente

1. Los resultados están disponibles previa solicitud a los autores.

(Baleares, Canarias y la Comunidad Valenciana). En este ámbito, la residencia en Aragón, Castilla y León, la Comunidad de Madrid y Navarra reduce la probabilidad relativa de pertenecer a los niveles más bajos de adquisición de competencias matemáticas. El efecto se sitúa en una reducción del 39,6% en Aragón y entre 47% y 58% en el resto de Comunidades indicadas. Residir en La Rioja también reduce dicha probabilidad (un 32,8%) pero dicha variable es significativa sólo al 10%.

Con objeto de establecer la consistencia temporal de los resultados presentados, a continuación éstos son comparados con investigaciones realizadas con datos procedentes de evaluaciones anteriores de PISA, que también analizan los factores que determinan que los alumnos tengan un muy bajo nivel de competencias[2]. Los trabajos seleccionados son Calero *et al.* (2012), con datos de la evaluación de PISA de 2009, y Calero y Escardíbul (2009), referido a datos del año 2003. El primero analiza los determinantes del rendimiento en comprensión lectora mientras que el segundo examina la adquisición de competencias matemáticas.

Con respecto a las variables personales, se pueden establecer comparaciones con las variables referidas al género, la asistencia más de un año a educación infantil, la repetición de curso y la condición de inmigrante. Como cabe esperar, ser mujer incide del mismo modo en la evaluación de 2003 (que considera las matemáticas) que en nuestro estudio y de manera contraria en la evaluación de 2009 (que analiza la comprensión lectora). Así, las chicas tienen una mayor probabilidad relativa de pertenecer al grupo de estudiantes que alcanza unos peores resultados en matemáticas y una menor probabilidad en el caso de la lectura. Asimismo, el hecho de haber cursado educación infantil más de un año reduce la probabilidad relativa de pertenecer a los grupos que tienen una menor competencia en las dos materias evaluadas. Además, la repetición de curso es una variable siempre muy significativa en todas las evaluaciones; ésta está asociada con que los estudiantes obtengan unos bajos resultados. Finalmente, el mismo efecto genera la variable referida a la condición de inmigrante (de primera generación) de los estudiantes, en las evaluaciones donde esta variable se incluía en el estudio (2009 y 2012).

Las variables familiares no pueden compararse estrictamente ya que en estudios anteriores se han considerado, en lugar del índice de nivel socioeconómico y cultural de las familias, algunas de las variables que lo componen (como la ocupación de los padres o el nivel cultural del hogar). En análisis anteriores, estas variables tienen el mismo signo negativo que el índice considerado en el presente estudio.

---

2. Aunque las regresiones no incluyen exactamente las mismas variables, sí pueden establecerse comparaciones, con cierta cautela, entre los signos de las variables explicativas incluidas ya que, en los diferentes estudios considerados, gran parte de las variables son las mismas, se utiliza la misma metodología econométrica y se define del mismo modo la variable dependiente.

La mayor parte de variables escolares no resultan significativas ni en el presente estudio ni en los anteriores. Sin embargo, merece la pena señalar la incidencia de la variable relacionada con el porcentaje de inmigrantes en la escuela. En los análisis previos se considera el porcentaje de inmigrantes no en intervalos (como en el estudio de 2012) sino mediante la creación de una variable dicotómica a partir de un determinado umbral (10% en el estudio de 2003, y 20% en 2009). Los resultados muestran que dicho porcentaje actúa negativamente en 2003 pero no se observa incidencia alguna de dicha variable en la evaluación de 2009. Por tanto, parece que ha aumentado el umbral a partir del cual la concentración de inmigrantes en los centros escolares favorece la probabilidad de que los estudiantes obtengan unos peores resultados. Finalmente, a diferencia del presente estudio, en anteriores evaluaciones resulta positiva para la adquisición de competencias una mayor presencia de chicas en la escuela.

La forma en que se han incluido las variables TIC en los distintos estudios revisados ha sido bastante diferente, de modo que resulta difícil comparar entre los mismos. Sin embargo, puede concluirse que la variable índice que define una mayor disponibilidad de recursos TIC en la escuela reduce la probabilidad relativa de que los alumnos estén en los niveles 1 o inferior de competencia matemática (aunque no así la variable que considera el número de ordenadores dedicados a la enseñanza por alumno, cuyo efecto varía en distintas muestras). Finalmente, no pueden establecerse comparaciones en torno a las variables de localización, debido a la falta de inclusión de variables regionales en otros estudios y la no incorporación, o distinta clasificación, en el caso del tamaño del municipio.

CUADRO 2. FACTORES DETERMINANTES DE QUE EL ALUMNADO ESTÉ EN UN NIVEL BAJO DE COMPETENCIAS EN MATEMÁTICAS. REGRESIÓN LOGÍSTICA MULTINIVEL, ESTIMACIÓN DE EFECTOS FIJOS CON ERRORES ESTÁNDAR ROBUSTOS

VARIABLES	COEFICIENTE (ERROR ESTÁNDAR)	ODDS-RATIO
CONSTANTE	-3,627* (2,024)	0,027
VARIABLES PERSONALES		
EDAD	0,099 (0,117)	1,105
MUJER	0,502*** (0,071)	1,653

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 2. FACTORES DETERMINANTES DE QUE EL ALUMNADO ESTÉ EN UN NIVEL BAJO DE COMPETENCIAS EN MATEMÁTICAS. REGRESIÓN LOGÍSTICA MULTINIVEL, ESTIMACIÓN DE EFECTOS FIJOS CON ERRORES ESTÁNDAR ROBUSTOS

VARIABLES	COEFICIENTE (ERROR ESTÁNDAR)	ODDS-RATIO
VARIABLES PERSONALES		
EDINFA (ha cursado ed. infantil más de un 1 año)	-0,389*** (0,106)	0,678
REFERENCIA: NO HA REPETIDO CURSO		
REPPRIM (repetido primaria)	2,117*** (0,086)	8,309
REPSEC (repetido secundaria)	1,682*** (0,094)	5,375
REFERENCIA: NO HA FALTADO A CLASE EN ÚLTIMAS 2 SEMANAS		
ABSENT 1 (faltó a clase 1-2 días)	0,360*** (0,080)	1,434
ABSENT 2 (faltó más de dos días)	0,726*** (0,162)	2,066
REFERENCIA: NATIVO		
INMIG 1 (inmigrante de primera generación)	0,480*** (0,119)	1,617
INMIG 2 (inmigrante de segunda generación)	0,316 (0,234)	1,371
VARIABLES FAMILIARES		
IDIOMA (Idioma en casa igual al del test)	-0,113 (0,101)	0,893
MONOPARENTAL	-0,232** (0,115)	0,793
ESCS (índice socioeconómico y cultural)	-0,193*** (0,038)	0,824
VARIABLES DE CARACTERÍSTICAS ESCOLARES		
REFERENCIA: CENTRO PÚBLICO		
PRIVIND (Centro privado independiente)	-0,076 (0,199)	0,927
CONCERT (Centro privado concertado)	-0,243 (0,152)	0,785

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 2. FACTORES DETERMINANTES DE QUE EL ALUMNADO ESTÉ EN UN NIVEL BAJO DE COMPETENCIAS EN MATEMÁTICAS. REGRESIÓN LOGÍSTICA MULTINIVEL, ESTIMACIÓN DE EFECTOS FIJOS CON ERRORES ESTÁNDAR ROBUSTOS

VARIABLES	COEFICIENTE (ERROR ESTÁNDAR)	ODDS-RATIO
VARIABLES DE CARACTERÍSTICAS ESCOLARES		
SCHLSIZE (número de alumnos en la escuela)	0,000 (0,000)	1,000
ALUMCLAS (número de alumnos por clase)	0,005 (0,007)	1,005
PCGIRLS (% chicas en la escuela)	-0,317 (0,384)	0,728
AÑOSEDCEN (años medios de escolarización de los padres)	-0,089** (0,040)	0,915
REFERENCIA: 0% INMIGRANTES EN LA ESCUELA		
NONAT 1 (% inmigrantes >0% y ≤ 10%)	-0,056 (0,119)	0,946
NONAT 2 (% inmigrantes >10% y ≤ 20%)	0,118 (0,161)	1,126
NONAT 3 (% inmigrantes >20% y ≤ 30%)	0,043 (0,214)	1,044
NONAT 4 (% inmigrantes >30% y ≤ 40%)	0,315 (0,207)	1,370
NONAT 5 (% inmigrantes >40%)	0,798*** (0,280)	2,221
REFERENCIA: NO ESCUELAS EN COMPETENCIA		
COMPET 1 (una escuela compite)	0,205 (0,178)	1,228
COMPET 2 (dos o más escuelas compiten)	0,217 (0,166)	1,242
VARIABLES DE GESTIÓN ESCOLAR Y DEL APRENDIZAJE		
ALUMPROF (ratio alumnos/profesorado)	0,003 (0,003)	1,003
SCMATEDU (índice de calidad de recursos educativos)	-0,019 (0,050)	0,981

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 2. FACTORES DETERMINANTES DE QUE EL ALUMNADO ESTÉ EN UN NIVEL BAJO DE COMPETENCIAS EN MATEMÁTICAS. REGRESIÓN LOGÍSTICA MULTINIVEL, ESTIMACIÓN DE EFECTOS FIJOS CON ERRORES ESTÁNDAR ROBUSTOS

VARIABLES	COEFICIENTE (ERROR ESTÁNDAR)	ODDS-RATIO
VARIABLES DE GESTIÓN ESCOLAR Y DEL APRENDIZAJE		
PROFESORADO (índice de falta de profesorado)	-0,040 (0,064)	0,961
REFERENCIA: NO AGRUPACIÓN DE ALUMNADO EN CLASE		
AGRUP 1 (agrupación en algunas clases)	0,202 (0,203)	1,223
AGRUP 2 (agrupación en todas las clases)	0,248 (0,201)	1,281
PUBLIRE (se publicitan los resultados del centro)	0,293** (0,123)	1,341
AUTONOMÍA EN EL CENTRO EDUCATIVO EN:		
RESPRES (índice de autonomía de gestión presupuestaria)	-0,172* (0,090)	0,842
RESPCUR (índice de autonomía de gestión curricular)	0,064 (0,079)	1,066
SISTEMA DE EVALUACIÓN EN EL CENTRO EDUCATIVO:		
EVALU 1 (autoevaluación o evaluación interna)	0,280 (0,167)	1,323
EVALU 2 (evaluación externa)	-0,094 (0,119)	0,910
VARIABLES TIC		
EDAD TIC (edad de inicio en las TIC)	0,077*** (0,014)	1,080
TICESC (disponibilidad TIC en la escuela)	0,115*** (0,038)	1,122
VARIABLES DE LOCALIZACIÓN		
REFERENCIA: POBLACIÓN DE HASTA 3.000 HABITANTES		
POBLACIÓN 1 (3.000-15.000 habitantes)	-0,408** (0,206)	0,665
POBLACIÓN 2 (15.000-100.000 habitantes)	-0,462** (0,230)	0,630

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 2. FACTORES DETERMINANTES DE QUE EL ALUMNADO ESTÉ EN UN NIVEL BAJO DE COMPETENCIAS EN MATEMÁTICAS. REGRESIÓN LOGÍSTICA MULTINIVEL, ESTIMACIÓN DE EFECTOS FIJOS CON ERRORES ESTÁNDAR ROBUSTOS

VARIABLES	COEFICIENTE (ERROR ESTÁNDAR)	ODDS-RATIO
VARIABLES DE LOCALIZACIÓN		
POBLACIÓN 3 (100.000-1.000.000 habitantes)	-0,409* (0,245)	0,664
POBLACIÓN 4 (más de 1 millón de habitantes)	-0,479 (0,313)	0,619
REFERENCIA: OTRAS COMUNIDADES AUTÓNOMAS		
ANDALUCÍA	0,000 (0,203)	1,000
ARAGÓN	-0,504** (0,222)	0,604
ASTURIAS	-0,213 (0,213)	0,808
BALEARES	-0,267 (0,238)	0,766
CANTABRIA	0,013 (0,247)	1,013
CASTILLA Y LEÓN	-0,874*** (0,203)	0,417
CATALUÑA	-0,048 (0,224)	0,953
EXTREMADURA	0,177 (0,213)	1,194
GALICIA	0,043 (0,218)	1,044
LA RIOJA	-0,398* (0,212)	0,672
MADRID	-0,629*** (0,203)	0,533
MURCIA	-0,234 (0,212)	0,792
NAVARRA	-0,842*** (0,242)	0,431
PAÍS VASCO	-0,199 (0,208)	0,820

*** Significativa al 1%. ** Significativa al 5%. * Significativa al 10%.

EL ABANDONO PREMATURO COMO PRINCIPAL FACTOR DE EXCLUSIÓN EN EL SISTEMA EDUCATIVO ESPAÑOL

PRESENTACIÓN

El principal elemento de exclusión en el sistema educativo español lo constituye, a nuestro juicio, el abandono prematuro, la falta de continuidad de los estudios al finalizar la educación obligatoria a los dieciséis años. El acceso o no a la Educación Secundaria postobligatoria determina en muy buena medida las posibilidades de formación adicional, las oportunidades en el mercado de trabajo, y las trayectorias vitales en diversos ámbitos. Por lo que respecta a la formación adicional, las probabilidades de participar en actividades de formación a lo largo de la vida se ven muy incrementadas en función del nivel de formación inicial (Calero, 2012). Las oportunidades en el mercado de trabajo, como las probabilidades de evitar el desempleo y acceder a puestos de trabajo mejor remunerados dependen de los niveles educativos de partida y el abandono prematuro incide muy negativamente sobre tales probabilidades (Calero y Gil, 2014). Finalmente, el abandono prematuro impide el acceso a una serie de beneficios no monetarios que mejoran las condiciones de vida en ámbitos como la salud, el consumo o la participación (Fernández y Calero, 2014).

El nivel de abandono prematuro es en España considerablemente más alto que en la media de los países de la Unión Europea: en 2013, siguiendo la definición estándar de EUROSTAT[3], el abandono prematuro era en España del 24,9% (12% era el nivel medio europeo). En el inicio de la crisis económica (2008), sin embargo, el nivel de abandono en España y la diferencia con la media europea era considerablemente más alta: en ese año, el 31,9% de los jóvenes españoles no había continuado sus estudios después de los dieciséis años (14,9% era la media europea). Esta reducción del abandono prematuro está causada en buena medida por el incremento del desempleo juvenil durante la crisis, que convierte en más "asequible" la continuidad de los estudios debido a que los jóvenes no tienen acceso a puestos de trabajo ni a sus salarios. Estos últimos constituían, hasta 2007, un elevado coste de oportunidad para la inversión educativa, especialmente para los hombres, que accedían a puestos de trabajo de bajo valor añadido en los sectores de la construcción y los servicios. A pesar de la reducción descrita, el nivel de abandono prematuro se sitúa, en España, todavía considerablemente lejos del objetivo fijado por la Comisión Eu-

3. En la actualidad EUROSTAT define el abandono prematuro de la educación y la formación (anteriormente, abandono prematuro de la educación) como el porcentaje de personas con edades comprendidas entre 18 y 24 años que ha finalizado como máximo la Educación Secundaria obligatoria y no realiza actividades adicionales de educación o formación.

ropea en su Estrategia *Education and Training* 2020, que consiste en alcanzar el 10% de abandono en el año 2020.

En los siguientes dos apartados abordamos cuantitativamente la situación y evolución reciente del abandono prematuro en España, utilizando en primer lugar una aproximación puramente descriptiva para, posteriormente analizar sus factores determinantes mediante la estimación de modelos de regresión logística. En ambos casos empleamos los microdatos de EU-SILC en sus olas de 2007 y 2013.

UNA DESCRIPCIÓN DE LA INCIDENCIA DEL ABANDONO PREMATURO EN ESPAÑA

El Cuadro 3 proporciona una descripción de cómo incide el abandono prematuro sobre diferentes colectivos y cómo ha evolucionado esta incidencia en los últimos años, en concreto entre 2007 y 2013. Los dos años comparados son relevantes, en tanto que el primero corresponde a un momento previo a la crisis económica y el último, 2013, corresponde a los datos disponibles más recientes. Observamos, en primer lugar, que la incidencia del abandono es mucho mayor en los hombres que en las mujeres; los hombres sufren mayores niveles de fracaso escolar y tienen una mayor tendencia a salir al mercado de trabajo (cuando ello es posible). La caída del abandono durante los años de crisis, sin embargo, se ha concentrado en los hombres: desde 2007 a 2013 el abandono ha caído 5,7 puntos porcentuales en el caso de los hombres (1,3 únicamente en el caso de las mujeres). El efecto de variables relativas a la composición de la familia, como el número de hermanos y la convivencia en el hogar del padre y la madre del joven ("pareja intacta") es el esperado. A medida que crece el número de hermanos (con la salvedad del paso entre el hijo único y el que tiene un hermano) crece la incidencia del abandono prematuro. Los hijos de progenitores que conviven en el mismo hogar sufren una menor incidencia del abandono. Ambas pautas se han mantenido estables entre 2007 y 2013, si bien en 2013 encontramos una menor distancia entre las dos categorías definidas por "pareja intacta".

También en el Cuadro 3 podemos observar la incidencia del abandono en función de tres variables relacionadas con el origen social de la familia. En primer lugar, vemos cómo las familias en las que ambos progenitores son extranjeros (origen inmigrante) presentan un mayor nivel de abandono. En 2007 la diferencia entre estas familias y el resto era muy importante (11,5 puntos porcentuales); en 2013 la diferencia se ha reducido notablemente, hasta únicamente 1,7 puntos. Esta reducción tan sustancial está, en parte, provocada por el diferente tipo de persona a la que se refiere la variable en los dos años considerados: en 2007 la gran mayoría de jóvenes

con edades comprendidas entre 18 y 24 años con padres extranjeros había llegado a España hacía pocos años y, por tanto, había participado poco en el sistema educativo español. Seis años más tarde (2013) es más fácil encontrar en ese colectivo a jóvenes con trayectorias más largas en el sistema educativo español y, por tanto, con menor incidencia del abandono prematuro.

Por lo que respecta a la renta disponible del hogar, observamos la gran distancia que se establece entre quintiles. La incidencia del abandono en el quintil de renta más bajo es seis veces mayor que en el quintil de renta más alto. La reducción del abandono entre 2007 y 2013 ha beneficiado a todos los quintiles, pero ni mucho menos ha sido homogénea: resulta muy interesante observar cómo a medida que sube el nivel de renta (con excepción del quinto quintil) más intensa ha sido la reducción del abandono prematuro. El quintil primero ha visto reducido su abandono en un 16,9%, el quintil cuarto en un 52,3%. Una pauta similar la encontramos en la incidencia del abandono en función del nivel educativo de los padres. Observamos, tanto en 2007 como en 2013, que el nivel educativo de los progenitores (el más elevado de padre y madre) protege eficazmente del abandono prematuro. En 2013, por ejemplo, únicamente el 8,3% de los hijos de padres con educación superior resultaba afectado por el abandono, unas seis veces menos que los hijos de padres con Educación Primaria o inferior. Las dificultades para acceder a un puesto de trabajo han reducido el abandono prematuro de los hijos cuyos padres tienen cualquier nivel de estudios, salvo el nivel más bajo (primaria o inferior); en ese caso el abandono de hecho se incrementa hasta alcanzar el 48,5%. La reducción del abandono es proporcionalmente mayor a medida que sube el nivel de estudios de los padres, con la excepción del nivel de estudios superior.

También apreciamos la misma pauta en la variable "categoría socioprofesional de los padres". La incidencia del abandono es mayor en los grupos con categoría de "cuello azul" y, tanto en la categoría de "cuello azul" como en la de "cuello blanco" el abandono es mayor en los grupos no cualificados. La reducción del abandono entre 2007 y 2013 beneficia, en este caso, a todas las categorías; sin embargo, la reducción es proporcionalmente mayor en la categoría de "cuello blanco cualificado" (27,9%) e inferior en el resto de categorías (se sitúa en las tres entre 16 y 17%).

Finalmente, también en el Cuadro 3 podemos observar el cruce relativo a la Comunidad Autónoma de residencia. Se aprecia una gran diversidad tanto en los niveles de abandono prematuro como en su evolución en el periodo 2007-2013. Un grupo de Comunidades del sur de la península, además de las dos Comunidades insulares y las Ciudades Autónomas (Andalucía, Baleares, Canarias, Castilla-La Mancha, Extremadura, Murcia y Ceuta y Melilla) presentan valores de abandono por encima de la media. Otras Comunidades septentrionales como Aragón, Astu-

rias, Castilla y León y, de forma más pronunciada, Cantabria, Navarra y País Vasco, presentan valores muy reducidos. La caída de los valores de abandono entre 2007 y 2013 ha sido muy irregular: Comunidades como el País Vasco o Cantabria presentan reducciones por encima del 40%, mientras que otras Comunidades, como Asturias o Madrid (ambas, también, con valores por debajo de la media) han reducido el abandono únicamente en un 10,9% y un 7,3%, respectivamente. La dispersión de la incidencia del abandono entre Comunidades, en términos relativos, ha permanecido estable a lo largo del periodo.

CUADRO 3. INCIDENCIA DEL ABANDONO PREMATURO EN FUNCIÓN DE DIFERENTES VARIABLES. ESPAÑA, 2007 Y 2013. CIFRAS EN PORCENTAJE

	2007	2013
TOTAL	**28,9**	**24,9**
SEXO		
Hombre	34,8	29,1
Mujer	21,7	20,4
NÚMERO DE HERMANOS		
0	27,7	23,7
1	24,5	20,6
2	37,7	33,2
3 ó más	50,9	47,0
PAREJA INTACTA		
No	34,7	27,0
Sí	27,9	24,3
ORIGEN INMIGRANTE		
No	28,7	24,7
Sí	40,2	26,4
QUINTIL DE RENTA DISPONIBLE EQUIVALENTE		
1	43,8	36,4
2	33,0	27,0
3	27,3	20,7
4	19,5	9,3
5	6,9	6,1

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 3. INCIDENCIA DEL ABANDONO PREMATURO EN FUNCIÓN DE DIFERENTES VARIABLES. ESPAÑA, 2007 Y 2013. CIFRAS EN PORCENTAJE

	2007	2013
TOTAL	**28,9**	**24,9**
NIVEL DE ESTUDIOS MÁXIMOS DE LOS PADRES		
Primaria o inferior	43,4	48,5
Secundaria inferior	35,0	31,9
Secundaria superior	18,9	13,9
Superior	6,9	8,3
CATEGORÍA SOCIOPROFESIONAL DE LOS PADRES		
Cuello blanco cualificado	11,1	8,0
Cuello blanco no cualificado	23,4	19,4
Cuello azul cualificado	39,7	33,3
Cuello azul no cualificado	43,0	35,7
COMUNIDAD AUTÓNOMA DE RESIDENCIA		
Andalucía	38	28,4
Aragón	27	18,8
Asturias	21	18,7
Baleares	44	29,7
Canarias	37	28,3
Cantabria	27	11,8
Castilla y León	23	18,8
Castilla-La Mancha	35	26,6
Cataluña	32	24,4
Com. Valenciana	32	22,3
Extremadura	35	28,5
Galicia	22	20,4
Madrid	25	20,2
Murcia	38	26,9

[CONTINÚA EN PÁGINA SIGUIENTE]

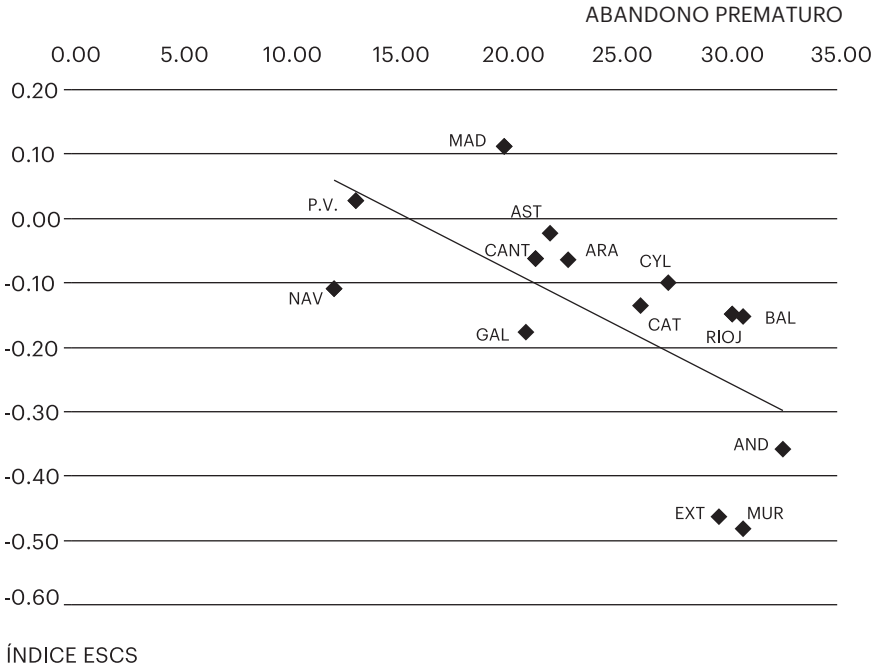
CUADRO 3. INCIDENCIA DEL ABANDONO PREMATURO EN FUNCIÓN DE DIFERENTES VARIABLES. ESPAÑA, 2007 Y 2013. CIFRAS EN PORCENTAJE

	2007	2013
TOTAL	**28,9**	**24,9**
COMUNIDAD AUTÓNOMA DE RESIDENCIA		
Navarra	17	12,8
País Vasco	15	8,8
La Rioja	30	21,3
Ceuta y Melilla	55	36,3

Fuente: Todos los cruces, salvo el correspondiente a la Comunidad Autónoma, elaboración propia a partir de microdatos de EU-SILC-2007 y EU-SILC-2013. Cruce por Comunidad Autónoma: Instituto de Evaluación (2009) e INEE (2014); se han utilizado en este caso datos censales con objeto de evitar las desviaciones que se producen al segmentar por Comunidad Autónoma la muestra de EU-SILC.

En buena medida, la diversidad de la incidencia del abandono prematuro entre Comunidades Autónomas es el resultado de una distribución diversa de la población en términos de sus recursos económicos y socio-culturales. El Gráfico 1 permite ilustrar tal afirmación: en él se sitúa a las Comunidades en función de su nivel de abandono prematuro y de la puntuación media de los estudiantes de la Comunidad Autónoma del índice ESCS (estatus económico y socio cultural) construido por la OCDE en la evaluación de PISA y que fija la media de 0 en el valor correspondiente a la media de la OCDE. Se puede observar la clara relación inversamente proporcional que se establece entre ambas variables. Adicionalmente, destacan en el gráfico los casos de algunas Comunidades distanciadas, por encima o por debajo, de la recta de regresión. Es el caso de Andalucía, Extremadura y Murcia, con valores de ESCS muy bajos pero que, además, presentan niveles de abandono prematuro inferiores a los que se predicen en función del valor de ESCS. En el otro extremo, se encuentra la Comunidad de Madrid que, con un valor de ESCS muy elevado, tiene un abandono prematuro superior al predecible.

GRÁFICO 1. RELACIÓN ENTRE EL ÍNDICE ESCS DE PISA Y EL ABANDONO ESCOLAR PREMATURO POR COMUNIDADES AUTÓNOMAS. 2011-12



Fuente: Elaboración propia a partir de datos de PISA-2012 (para índice ESCS) y de INEE (2013a) para datos de abandono prematuro.

MODELOS EXPLICATIVOS DE LA PROBABILIDAD DE INCURRIR EN ABANDONO PREMATURO

En este apartado presentamos los resultados de la estimación de un modelo, aplicado a los datos de EU-SILC de 2007 y 2013, con el que pretendemos analizar el efecto de los factores que determinan la probabilidad de incurrir en situación de abandono prematuro en España y su evolución en los últimos años. La estimación se ha efectuado mediante la técnica de la regresión logística, dado el carácter dicotómico (abandono/no abandono) de la variable explicada.

En los Cuadros 4 y 5 aparecen los resultados de la estimación, referida, respectivamente, a los años 2007 y 2013. Como veremos, los resultados son similares en los dos años, aunque se percibe una clara evolución que confirma lo observado en el análisis bivariante del apartado anterior la reducción del abandono prematuro ha beneficiado proporcionalmente más a los grupos sociales con más recursos económicos y socio-culturales.

CUADRO 4. DETERMINANTES DE LA PROBABILIDAD DE ABANDONO PREMATURO. RESULTADOS DE LA REGRESIÓN LOGÍSTICA. ESPAÑA, AÑO 2007

	COEF.	SIG.	ERROR ESTÁNDAR	EXP(B)
Edad	0,048	**	0,023	1,049
Sexo (1= hombre; 0= mujer)	0,881	***	0,095	2,412
Número de hermanos	0,364	***	0,054	1,439
Pareja intacta	-0,185		0,127	0,831
Origen inmigrante	0,379		0,282	1,461
QUINTIL DE RENTA DISPONIBLE EQUIVALENTE. CATEGORÍA DE REFERENCIA: QUINTIL 3				
Quintil 1	0,343	***	0,134	1,409
Quintil 2	0,080		0,142	1,083
Quintil 4	-0,064		0,164	0,938
Quintil 5	-0,794	***	0,263	0,452
NIVEL DE ESTUDIOS MÁXIMOS DE LOS PADRES. CATEGORÍA DE REFERENCIA: EDUCACIÓN PRIMARIA O INFERIOR				
Educación Secundaria inferior	-0,648	***	0,115	0,523
Educación Secundaria superior	-1,339	***	0,140	0,262
Educación superior	-2,233	***	0,196	0,107
CATEGORÍA SOCIOPROFESIONAL DE LOS PADRES. CATEGORÍA DE REFERENCIA: CUELLO BLANCO CUALIFICADO				
Cuello blanco no cualificado	0,063		0,134	1,065
Cuello azul cualificado	0,257	*	0,136	1,294
Cuello azul no cualificado	0,425	***	0,139	1,529
COMUNIDAD AUTÓNOMA DE RESIDENCIA. CATEGORÍA DE REFERENCIA: ANDALUCÍA				
Aragón	-0,289		0,282	0,749
Asturias	-0,297		0,269	0,743
Baleares	0,792	***	0,265	2,209
Canarias	-0,083		0,208	0,921
Cantabria	-0,554		0,356	0,575
Castilla y León	-0,574	***	0,231	0,564

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 4. DETERMINANTES DE LA PROBABILIDAD DE ABANDONO PREMATURO. RESULTADOS DE LA REGRESIÓN LOGÍSTICA. ESPAÑA, AÑO 2007

	COEF.	SIG.	ERROR ESTÁNDAR	EXP(B)
COMUNIDAD AUTÓNOMA DE RESIDENCIA. CATEGORÍA DE REFERENCIA: ANDALUCÍA				
Castilla - La Mancha	0,156		0,221	1,169
Cataluña	-0,009		0,193	0,991
Comunidad Valenciana	-0,065		0,204	0,937
Extremadura	-0,200		0,230	0,819
Galicia	-0,580	***	0,212	0,560
Madrid	-0,079		0,241	0,924
Murcia	0,581	***	0,202	1,788
Navarra	-0,445	*	0,331	0,641
País Vasco	-0,970	***	0,331	0,379
La Rioja	-0,374		0,291	0,688
Ceuta y Melilla	0,407		0,292	1,503

Nota: ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente.
Fuente: Elaboración propia a partir de microdatos de EU-SILC, base de datos de España, 2007.

CUADRO 5. DETERMINANTES DE LA PROBABILIDAD DE ABANDONO PREMATURO. RESULTADOS DE LA REGRESIÓN LOGÍSTICA. ESPAÑA, AÑO 2013

	COEF.	SIG.	ERROR ESTÁNDAR	EXP(B)
Edad	0,172	***	0,026	1,188
Sexo (1= hombre)	0,510	***	0,104	1,665
Número de hermanos	0,249	***	0,061	1,282
Pareja intacta	-0,058		0,133	0,944
Origen inmigrante	0,262		0,260	1,299
QUINTIL DE RENTA DISPONIBLE EQUIVALENTE. CATEGORÍA DE REFERENCIA: QUINTIL 3				
Quintil 1	0,342	**	0,152	1,408
Quintil 2	0,188		0,165	1,207
Quintil 4	-0,346		0,224	0,707
Quintil 5	-0,421		0,271	0,657

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 5. DETERMINANTES DE LA PROBABILIDAD DE ABANDONO PREMATURO. RESULTADOS DE LA REGRESIÓN LOGÍSTICA. ESPAÑA, AÑO 2013

	COEF.	SIG.	ERROR ESTÁNDAR	EXP(B)
NIVEL DE ESTUDIOS MÁXIMOS DE LOS PADRES. CATEGORÍA DE REFERENCIA: EDUCACIÓN PRIMARIA O INFERIOR				
Educación Secundaria inferior	-0,705	***	0,137	0,494
Educación Secundaria superior	-1,431	***	0,157	0,239
Educación superior	-1,773	***	0,178	0,170
CATEGORÍA SOCIOPROFESIONAL DE LOS PADRES. CATEGORÍA DE REFERENCIA: CUELLO BLANCO CUALIFICADO				
Cuello blanco no cualificado	0,021		0,150	1,021
Cuello azul cualificado	0,208		0,156	1,231
Cuello azul no cualificado	0,455	***	0,162	1,576
COMUNIDAD AUTÓNOMA DE RESIDENCIA. CATEGORÍA DE REFERENCIA: ANDALUCÍA				
Aragón	-0,252		0,309	0,777
Asturias	-0,105		0,305	0,901
Baleares	0,071		0,316	1,073
Canarias	0,367		0,240	1,443
Cantabria	-1,032	**	0,513	0,356
Castilla y León	-0,303		0,262	0,739
Castilla - La Mancha	0,340		0,231	1,404
Cataluña	-0,119		0,208	0,888
Comunidad Valenciana	-0,316		0,233	0,729
Extremadura	0,277		0,225	1,319
Galicia	0,426	*	0,244	1,531
Madrid	-0,227		0,224	0,797
Murcia	0,032		0,248	1,033
Navarra	0,096		0,339	1,101
País Vasco	-0,758	**	0,359	0,469
La Rioja	-0,301		0,366	1,351
Ceuta y Melilla	0,364		0,307	0,695

Nota: ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente.
Fuente: Elaboración propia a partir de microdatos de EU-SILC, base de datos de España, 2013.

Vemos, en primer lugar, que la edad (que oscila entre 18 y 24 años para el grupo analizado, en función de la definición de abandono prematuro), afecta positivamente a la probabilidad de estar en situación de abandono. Se trata de un resultado esperado, debido al mayor alejamiento del sistema educativo que se produce con el transcurso del tiempo; adicionalmente, en 2013, los jóvenes con menos edad dentro del grupo analizado han estado en contacto con un mercado de trabajo muy poco receptivo. También es un resultado esperado el que encontramos en relación con el sexo: el coeficiente positivo nos indica una mayor probabilidad de abandono para los hombres a igualdad del resto de variables. Este coeficiente se reduce entre 2007 y 2013, como era de esperar. También resulta significativo y positivo (incrementa la probabilidad de incurrir en abandono prematuro) el efecto del número de hermanos. Sin embargo, la variable "pareja intacta" no es significativa en la estimación del modelo.

Merece la pena destacar el efecto no significativo de la variable relativa al origen inmigrante (o no) del joven. Veíamos en el apartado anterior cómo los jóvenes de origen inmigrante sufrían, en 2007, una incidencia mucho mayor del abandono prematuro que el resto de jóvenes; la diferencia en la incidencia entre ambos grupos se había reducido considerablemente en 2013. Los resultados del modelo nos permiten afirmar que las diferencias entre los dos grupos se derivan de otros factores que concurren frecuentemente con la condición de inmigrante (recursos económicos y socio-culturales más bajos, principalmente), ya que el efecto "aislado" de la condición de inmigrante es nulo. Ello es así tanto en 2007 como en 2013. Se trata de un resultado que descarta la existencia de una discriminación específica, en el sistema educativo, vinculada con el origen inmigrante.

El nivel de renta disponible por los hogares sí que está vinculado con las probabilidades de abandono. Resulta destacable, especialmente, el efecto de estar situado en el quintil de renta más bajo, en donde las probabilidades de caer en situación de abandono son significativamente más elevadas. Esta situación de desventaja asociada a la renta para el primer quintil, además, se ha mantenido estable entre 2007 y 2013: la reducción del abandono prematuro beneficia sobre todo a los grupos de mayor renta y, en este caso, no ha llegado a beneficiar a los jóvenes con menor renta.

Por lo que respecta al efecto del nivel de estudios de los padres, tanto en 2007 como en 2013, a medida que crece el nivel educativo se reduce, de forma significativa, la probabilidad de caer en el abandono prematuro. En 2013 observamos, por ejemplo, un *odd-ratio* (Exp(B) en el Cuadro 5) de 0,17 en el caso del nivel de educación superior, lo que significa que la probabilidad de estar en situación de abandono prematuro es 5,88 veces menor (1/0,17) para el hijo de padres con educación superior que para el hijo de padres con Educación Primaria o inferior. El efecto de

la categoría socioprofesional de los padres es prácticamente constante entre 2007 y 2013. Especialmente en el caso de la categoría "cuello azul no cualificado" se observa un efecto significativo positivo sobre la probabilidad de caer en el abandono prematuro. En el conjunto de variables socioeconómicas, destaca el efecto de la educación de los progenitores por encima de la renta y la categoría laboral.

La crisis económica y el consiguiente incremento del desempleo juvenil han modificado sustancialmente el efecto de la Comunidad Autónoma de residencia sobre el abandono prematuro[4]. En 2007 encontrábamos seis Comunidades que provocaban un efecto significativo, positivo o negativo, sobre el nivel de abandono: positivo en el caso de Baleares y Murcia y negativo en el caso de Castilla y León, Galicia, Navarra y el País Vasco. En 2013, sin embargo, únicamente tres Comunidades provocan un efecto significativo: positivo para Galicia y negativo para Cantabria y el País Vasco. Esta última Comunidad es la única positivamente significativa en ambos análisis. El mercado de trabajo, más o menos atractivo para los jóvenes, resultaba un factor significativo, antes de la crisis económica, a la hora de diferenciar los niveles de abandono prematuro entre Comunidades Autónomas. Durante la crisis, en 2013, las diferencias vinculadas específicamente a la Comunidad de residencia se han reducido, en tanto que de forma considerablemente homogénea los jóvenes entre 18 y 24 años tienen enormes dificultades para participar en el mercado de trabajo con independencia de la Comunidad Autónoma en la que residan.

LAS COMPETENCIAS DE LA POBLACIÓN ADULTA COMO FACTOR DE INCLUSIÓN Y EXCLUSIÓN EDUCATIVA

PRESENTACIÓN

En la secuencia de tres pasos que estamos realizando en este texto, la última que corresponde abordar en este apartado está relacionada con las competencias que ha adquirido la población adulta y con la identificación de los factores que diferencian a un grupo de población "excluida". La definimos como "excluida" en tanto que el nivel de competencias adquiridas a lo largo de la vida no va a permitir, a este grupo de personas, comprender y actuar adecuadamente en situaciones sociales y laborales complejas y va a relegarles a puestos de trabajo muy poco cualificados o, según el momento del ciclo económico, al desempleo. El nivel de competencias de un

4. Introducimos en el modelo únicamente *dummies* relativas a la Comunidad Autónoma. En modelos alternativos en los que se han introducido variables relativas a la incidencia por Comunidades del desempleo juvenil y al porcentaje de ocupación en la construcción (véase Casquero y Navarro, 2010 e IVIE, 2013) no se ha mejorado la capacidad predictiva, resultando ambas variables no significativas.

adulto es el resultado de la interacción entre al menos cuatro factores: las competencia adquiridas en el sistema educativo (educación formal), la formación recibida a lo largo de la vida de adulto, las competencias adquiridas en el puesto de trabajo y, finalmente, las competencias adquiridas durante otros tipos de interacción social. Hasta cierto punto, por consiguiente, la exclusión a la que nos vamos a referir en este apartado está determinada por exclusiones que han tenido lugar previamente, durante la educación formal, aunque en algunos casos de resiliencia la persona puede superarla, mediante la acción conjunta de los otros tres factores.

En 2013 se publicaron los datos correspondientes a la primera edición del programa PIAAC (Programa para la Evaluación Internacional de las Competencias de Adultos) de la OCDE, en la que España tomó parte junto con otros 24 países de dicha organización. En este programa se evalúa el nivel de competencias de matemáticas, lectura y "resolución de problemas en entornos digitales" de la población adulta potencialmente activa, con edades comprendidas entre 16 y 64 años. La muestra de adultos utilizada en el caso español fue de 6.055 personas. La publicación de estos datos, que utilizaremos en este apartado, ha permitido profundizar en el conocimiento de un ámbito, las competencias de la población adulta, que conocíamos sólo de forma superficial hasta el momento.

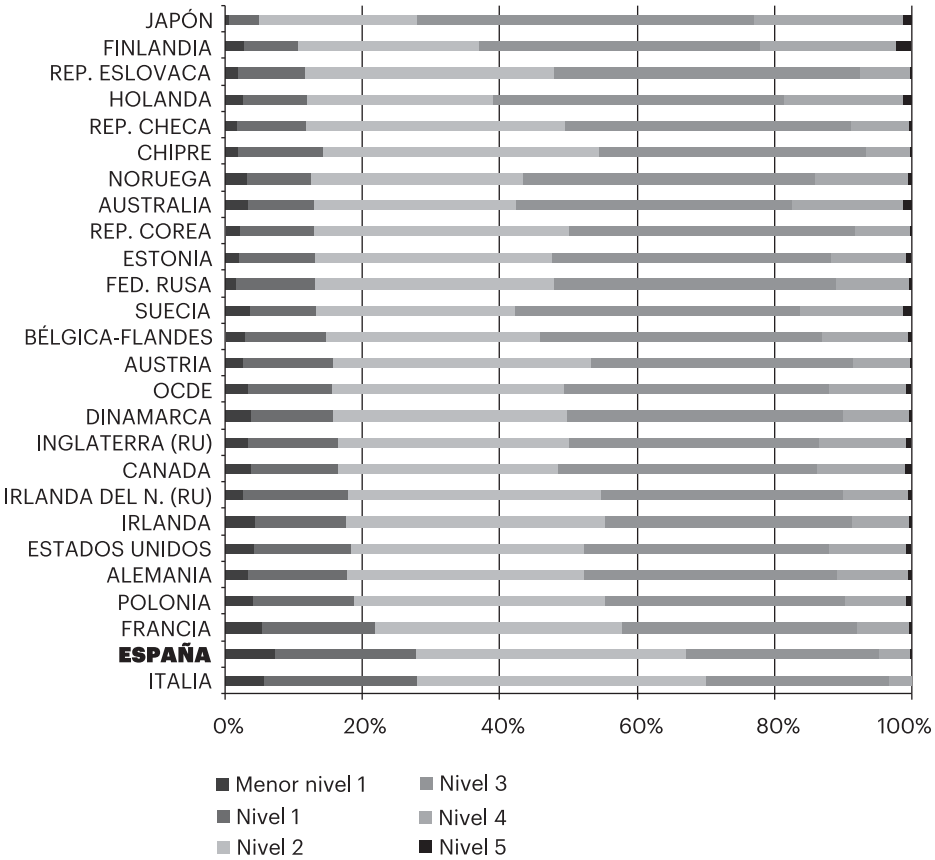
En los dos siguientes apartados abordamos la descripción de la población con niveles más bajos de competencias y estimamos un modelo de regresión logística orientado a predecir la probabilidad de haber adquirido competencias muy bajas.

UNA DESCRIPCIÓN DE LA INCIDENCIA DE LAS COMPETENCIAS BAJAS EN LA POBLACIÓN ADULTA ESPAÑOLA

Nos centraremos aquí, específicamente, en la competencia de lectura, la competencia que ha recibido una mayor atención en la primera edición de PIAAC-2012. En el Gráfico 2 aparece, para todos los países participantes en la evaluación, la distribución de la población en función de los distintos niveles de competencia, entre menor de 1 y 5. Vamos a prestar una especial atención a los dos niveles inferiores, menor de 1 y 1, que la OCDE identifica como aquellos en los que la población adulta tiene más dificultades para integrarse en el mercado de trabajo y, de forma más general, en la sociedad. Este grupo supone el 15,5% de la población adulta en el conjunto de los países de la OCDE evaluados y, en España, el 27,5%, un nivel sólo equivalente al de Italia.

La identificación de una proporción tan elevada de población adulta española con bajas competencias de lectura nos lleva a una primera reflexión que quisiéramos apuntar aquí: la proporción de alumnos de 15 años que PISA identifica

GRÁFICO 2. PIAAC-2012. DISTRIBUCIÓN DE LOS NIVELES DE COMPETENCIA DE LECTURA DE LA POBLACIÓN ADULTA. PAÍSES DE LA OCDE PARTICIPANTES EN PIAAC-2012



Fuente: OECD (2013).

como de bajas competencias es considerablemente menor y muy similar a la media de la Unión Europea y de la OCDE. El sistema educativo español consigue ese resultado a pesar del handicap que supone el muy reducido nivel de competencias de las generaciones anteriores a las que actualmente están estudiando. Esta constatación, sin llevarnos a aceptar como satisfactorios los resultados de PISA, sirve para relativizar las críticas a menudo indiscriminadas que recibe nuestro sistema educativo.

El Cuadro 6 muestra cómo incide, con diferente intensidad, la situación de competencias bajas de lectura para diversos grupos poblacionales. Los grupos están configurados en función de diferentes variables relevantes, que luego participarán (véase apartado siguiente) en nuestro modelo explicativo. Estas variables están vinculadas a características personales (que incluyen variables relativas a la

condición de inmigrante), características relativas al origen familiar (entorno socioeconómico), características vinculadas a la participación en el mercado de trabajo y, finalmente, a la participación en la formación.

CUADRO 6. INCIDENCIA DE LA ADQUISICIÓN DE COMPETENCIAS BAJAS EN FUNCIÓN DE DIFERENTES VARIABLES. ESPAÑA, 2012. PORCENTAJE Y CANTIDAD TOTAL DE PERSONAS EN CADA GRUPO

	% COMPETENCIAS BAJAS	TOTAL PERSONAS
TOTAL	27,50	24.892.664
EDAD		
16-24 años	20,88	1.746.595
25-34 años	19,90	6.157.103
35-44 años	22,26	7.002.492
45-54 años	29,60	5.813.064
55-64 años	47,37	4.173.411
SEXO		
Mujer	27,45	11.369.785
Hombre	27,55	13.522.879
No inmigrante	24,84	21.217.614
Inmigrante de primera generación (nacido en el extranjero)	46,34	3.177.789
Inmigrante de segunda generación (nacido en España de padres extranjeros)	11,87	69.290
AÑOS DE RESIDENCIA EN EL PAÍS (POBLACIÓN INMIGRANTE)		
0-5 años	41,76	758.673
6-10 años	47,29	1.337.272
11-15 años	56,47	651.124
Más de 15 años	31,84	819.846
TIENE COMO LENGUA MATERNA UNA LENGUA OFICIAL DEL PAÍS		
No	49,45	1.561.860
Sí	26,04	23.330.804

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 6. INCIDENCIA DE LA ADQUISICIÓN DE COMPETENCIAS BAJAS EN FUNCIÓN DE DIFERENTES VARIABLES. ESPAÑA, 2012. PORCENTAJE Y CANTIDAD TOTAL DE PERSONAS EN CADA GRUPO

	% COMPETENCIAS BAJAS	TOTAL PERSONAS
NIVEL DE ESTUDIOS MÁXIMOS		
Educación Primaria o inferior	59,12	4.536.135
Educación Secundaria inferior	37,15	6.174.819
Educación Secundaria superior profesional	28,39	626.946
NIVEL DE ESTUDIOS MÁXIMOS		
Educación Secundaria superior académica (Bachillerato)	18,82	4.754.439
Educación post-secundaria no superior	22,78	454.357
Educación superior no universitaria	15,26	2.197.647
Educación superior universitaria	5,68	6.138.474
NIVEL DE ESTUDIOS DE LOS PADRES		
Educación Secundaria inferior como máximo	32,70	17.382.575
Educación Secundaria superior	14,36	3.567.418
Educación superior	10,93	3.193.980
AÑOS DE EXPERIENCIA LABORAL		
0-5 años	23,89	4.509.612
6-10 años	23,81	4.266.941
11-20 años	24,55	7.075.144
21-30 años	25,42	4.853.764
31 y más años	42,80	4.101.135
SECTOR DE ACTIVIDAD		
Primario	46,45	1.379.893
Industria	27,31	2.966.536
Construcción	39,86	2.390.138
Servicios	24,47	18.156.098
NÚMERO DE HORAS DE PARTICIPACIÓN EN EDUCACIÓN NO FORMAL (AÑO ANTERIOR)		
No ha participado	36,11	13.183.668
1-20 horas	21,99	2.774.555

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO 6. INCIDENCIA DE LA ADQUISICIÓN DE COMPETENCIAS BAJAS EN FUNCIÓN DE DIFERENTES VARIABLES. ESPAÑA, 2012. PORCENTAJE Y CANTIDAD TOTAL DE PERSONAS EN CADA GRUPO

	% COMPETENCIAS BAJAS	TOTAL PERSONAS
NÚMERO DE HORAS DE PARTICIPACIÓN EN EDUCACIÓN NO FORMAL (AÑO ANTERIOR)		
21-50 horas	18,00	2.488.003
51-100 horas	15,43	1.722.858
Más de 100 horas	16,13	4.465.162

Fuente: Elaboración propia a partir de microdatos de PIAAC-2012.

Se puede observar, en primer lugar, cómo la población con más de 45 años contiene una gran proporción de personas con competencias bajas, proporción que alcanza casi el 50% en la población de más de 55 años. Esta asociación entre nivel de competencias y edad está parcialmente explicada al menos por dos factores: en primer lugar, la caída de competencias debido a un posible deterioro de las capacidades psicológicas de las personas y, en segundo lugar, a una adquisición de competencias iniciales, durante la juventud, que en su momento fuera de menor nivel que la disfrutada por las generaciones más jóvenes. El peso proporcional de cada uno de estos dos factores en cada país es una cuestión abierta por el momento a la discusión (OECD, 2013).

Vemos también en el Cuadro cómo no existen diferencias, a nivel agregado, entre la incidencia de las competencias bajas entre hombres y mujeres (cuestión que podremos matizar a la vista de los resultados del apartado siguiente). Por el contrario, se establecen muy profundas diferencias en función del origen (inmigrante o no) de la población: los inmigrantes de primera generación, nacidos en el extranjero, presentan una incidencia que es casi el doble de la que sufre la población no inmigrante. La incidencia en la población inmigrante de segunda generación (nacida en España de padres extranjeros) es muy reducida (11,87%). Este dato podría explicarse por la juventud de la población a la que se refiere, teniendo en cuenta que la llegada de población inmigrante es un fenómeno reciente; sin embargo, debe ser considerado con precaución debido al reducido tamaño de la muestra considerada. Para la población inmigrante de primera generación, los años de residencia en el país establecen diferencias importantes en cuanto a la incidencia de las bajas cualificaciones: los llegados más recientemente y los que llevan más de 15 años en el país presentan una incidencia comparativamente baja; sin embargo, los que llegaron a España hace entre 11 y 15 años presentan una incidencia máxima del 56,47%. También se establecen diferencias importantes en función de la condición

de tener como lengua materna una lengua oficial del país. Casi la mitad de los que no la tienen pertenecen al grupo de competencias bajas en la evaluación de PIAAC.

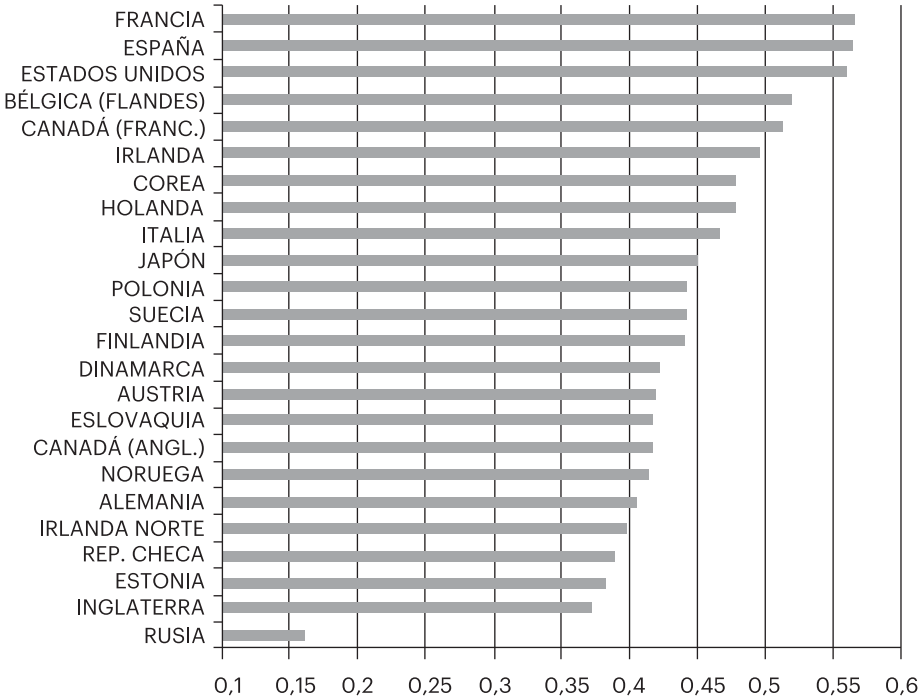
El nivel de estudios máximos alcanzado configura, como era de esperar, grupos de adultos con competencias muy diferenciadas. En el grupo de estudios más bajos (que en España recoge a 4,5 millones de personas) la incidencia de las bajas competencias alcanza a casi el 60% de las personas. Si agrupamos los dos grupos con estudios más bajos (primaria o inferior y secundaria inferior) estamos considerando a 10,7 millones de personas, o el 43% de la población considerada entre 16 y 64 años; en este grupo el 46,45% de las personas tiene un nivel de competencias de 1 o menor de 1. La incidencia disminuye a medida que sube el nivel educativo, situándose en el 5,68% en la población con educación universitaria (cifra que, sin embargo, sigue siendo sorprendentemente elevada). Los niveles educativos profesionales (tanto en Educación Secundaria como en superior) están asociados a un mayor porcentaje de personas con niveles educativos muy reducidos, en comparación con los niveles equivalentes académicos (Bachillerato, en Educación Secundaria, y Universidad en Educación Superior).

La correlación entre nivel educativo y nivel de competencias de la población adulta es elevada en todos los países que participan en PIAAC-2012 con la excepción de Rusia, como puede observarse en el Gráfico 3. En el gráfico se comprueba cómo España está entre los países donde la correlación es más elevada, lo que señala a que otros factores (como la formación continua o las competencias adquiridas en el puesto de trabajo) tienen un efecto relativamente menor a la hora de configurar las competencias de los adultos.

El nivel educativo de los padres permite distinguir, también, entre grupos con una presencia muy desigual de las competencias bajas. Éstas son tres veces más probables de identificar en una persona cuyos padres tenían el nivel educativo más bajo que en una persona cuyos padres tenían educación superior. La educación de los padres se proyecta sobre el nivel educativo de la persona y éste, a su vez, sobre el nivel de competencias. Veremos en el apartado siguiente, sin embargo, que existe una relación directa entre el nivel educativo de los padres y las competencias de los adultos.

Por lo que respecta a las variables relativas a la participación en el mercado de trabajo, observamos que la experiencia laboral no configura grupos diferenciados en cuanto a la incidencia de las bajas competencias, con la excepción del grupo con más experiencia (31 y más años). También veremos en el apartado siguiente cómo esta mayor incidencia en el grupo con más experiencia laboral está en realidad asociada con la mayor edad de las personas en ese grupo. El sector de actividad permite distinguir claramente entre grupos en función de la incidencia de las competencias

GRÁFICO 3. NIVELES DE CORRELACIÓN ENTRE LOS NIVELES DE COMPETENCIA DE LECTURA Y EL NIVEL EDUCATIVO MEDIDO EN AÑOS. POBLACIÓN ADULTA. PAÍSES DE LA OCDE PARTICIPANTES EN PIAAC-2012



Fuente: Elaboración propia a partir de microdatos de PIAAC-2012.

bajas. Éstas se concentran en el sector primario (muy especialmente) y en el de la construcción.

Finalmente, la participación (o no) en la educación no formal (formación continua y ocupacional) establece grupos muy bien delimitados en cuanto a las competencias. El grupo de personas que no participa en la formación presenta una incidencia de las bajas competencias muy superior a la media. La simple participación en actividades de formación ya se vincula a una incidencia por debajo de la media de las competencias bajas, que desciende hasta el 15%-16% para los adultos que participan en actividades de formación durante más de 50 horas al año.

MODELO EXPLICATIVO DE LA PROBABILIDAD, PARA LA POBLACIÓN ADULTA, DE SITUARSE EN LOS NIVELES MÁS REDUCIDOS DE COMPETENCIAS

En este apartado estimaremos un modelo explicativo de la probabilidad de situarse en los niveles 1 y menor de 1 de competencias de lectura para la población adulta española. Como sucedía en los apartados anteriores, referidos a las competencias de PISA y al abandono prematuro, la variable dependiente es dicotómica (situarse o no en los niveles más bajos de competencias), por lo que optamos por aplicar una regresión logística. Las variables independientes se refieren a características personales, familiares (nivel de estudios de los padres), relativas a la participación del individuo en el mercado de trabajo y, finalmente, una variable referida a la participación en la formación. En la muestra empleada se ha seleccionado a la población con algún tipo de experiencia laboral, con objeto de conocer el efecto de diversas variables relativas al puesto de trabajo y a la trayectoria laboral. Tal decisión reduce la población analizada con respecto a la que describíamos en el Cuadro 6, quedando la muestra sometida al análisis en 4.582 casos.

Los primeros resultados destacables, en el Cuadro 7, pertenecen al ámbito de las características personales. La edad presenta el resultado positivo esperado (la probabilidad de encontrarse en el grupo de competencias más bajas crece con la edad). El sexo, sin embargo, presenta unas diferencias entre hombres y mujeres que no aparecían en la descripción bivariante del Cuadro 6: los hombres tienen una probabilidad significativamente menor que las mujeres, a igualdad del resto de variables, de estar en el grupo de menores competencias de lectura.

Resulta también interesante apreciar cómo la condición de inmigrante de primera generación se asocia a un incremento de la probabilidad de situarse en el grupo de competencias más bajas, lo que significa que las diferencias que observábamos en la descripción bivariante son el resultado, específicamente, de la condición de inmigrante y no de otras variables que actúan simultáneamente. Los años de residencia en el país y tener como lengua materna una lengua oficial del país son significativas y tienen el signo esperado: en ambos casos se asocian a reducciones de la probabilidad de estar en el grupo de población "excluida" en términos de sus competencias.

Tomando como referencia el nivel de Educación Secundaria inferior (la actual ESO, por ejemplo), el nivel inferior (Educación Primaria o inferior) se asocia a una mayor incidencia de las competencias bajas, y todos los niveles superiores (con la muy notable excepción de la Educación Secundaria superior profesional –la antigua Formación Profesional o los actuales Ciclos Formativos de Grado Medio–) se asocian a menores incidencias.

CUADRO 7. DETERMINANTES DE LA PROBABILIDAD DE SITUARSE EN LOS NIVELES MÁS BAJOS DE COMPETENCIA DE LECTURA. ESPAÑA. PIAAC-2012

	COEFICIENTE (ERROR ESTÁNDAR)	SIG.	ODDS-RATIO
Edad	0,049 (0,009)	***	1,050
Hombre	-0,319 (0,087)	***	0,727
Inmigrante de primera generación (nacido en el extranjero)	0,679 (0,253)	***	1,972
Inmigrante de segunda generación (nacido en España de padres extranjeros)	0,022 (0,585)		1,022
Años de residencia en el país	-0,022 (0,008)	***	0,979
Tiene como lengua materna una lengua oficial del país	-0,346 (0,179)	*	0,707
NIVEL DE ESTUDIOS MÁXIMOS. CATEGORÍA DE REFERENCIA: EDUCACIÓN SECUNDARIA INFERIOR			
Educación Primaria o inferior	0,824 (0,097)	***	2,279
Educación Secundaria superior profesional	-0,287 (0,232)		0,750
Educación Secundaria superior académica (Bachillerato)	-0,967 (0,116)	***	0,380
Educación superior no universitaria	-1,068 (0,153)	***	0,344
Educación superior universitaria	-2,236 (0,153)	***	0,107
NIVEL DE ESTUDIOS DE LOS PADRES. CATEGORÍA DE REFERENCIA: EDUCACIÓN SECUNDARIA INFERIOR COMO MÁXIMO			
Educación Secundaria superior	-0,466 (0,135)	***	0,627
Educación superior	-0,387 (0,158)	**	0,679
Años de experiencia laboral	0,002 (0,005)		1,002
SECTOR DE ACTIVIDAD. CATEGORÍA DE REFERENCIA: SERVICIOS			
Primario	0,292 (0,148)	***	1,340
Industria	-0,045 (0,119)		0,956
Construcción	0,287 (0,127)	**	1,333
Número de horas de participación en educación no formal (año anterior)	0,000 (0,000)		1,000
Constante	-1,305 (0,260)	***	0,271

Fuente: Elaboración propia a partir de microdatos de PIAAC-2012.

Un nivel de estudios elevado de los padres está asociado a menores probabilidades de estar situado en el grupo de competencias más bajas de PIAAC-12. El efecto se produce a partir de la Educación Secundaria superior (de hecho, el efecto es ligeramente mayor para este nivel que para la educación superior). Conviene tener en cuenta que este efecto es el "aislado" con respecto al nivel educativo del propio individuo, que ya está en buena medida afectado por el nivel educativo de los padres. Podríamos hablar, por tanto, de un "residuo" de desigualdad intergeneracional, que sigue afectando al individuo incluso después de haber concluido su educación formal.

En el ámbito de variables relativas a la participación en el mercado de trabajo, merece la pena constatar el efecto no significativo de los años de experiencia laboral. Se trata de un resultado interesante, que está señalando a cómo la experiencia laboral no constituye un recurso suficientemente potente, en España, para la adquisición de competencias que permitan salir del grupo de competencias más bajas. El sector de actividad, por el contrario, se asocia en dos casos (sector primario y construcción) a la adquisición de un nivel bajo de competencias. Finalmente, el número de horas de participación en educación no formal, durante el año anterior, no introduce ningún efecto significativo sobre la adquisición de competencias de los adultos en España.

Los resultados que hemos comentado identifican una pauta de poca flexibilidad en la adquisición de competencias de la población adulta en España. Éstas están en buena medida establecidas al final de la educación formal. Las personas que, en función de su nivel educativo (que depende a su vez del nivel educativo de sus padres, aunque éste sigue actuando con posterioridad) y su origen migratorio caen en el grupo de competencia más bajas tienen pocas probabilidades de salir de él por medio de su participación en el mercado de trabajo y de la participación en la educación no formal.

CONCLUSIONES

Nos planteábamos, al principio de este trabajo, los siguientes tres objetivos. En primer lugar, describir las diferencias entre la población incluida y excluida de determinadas áreas del sistema educativo. En segundo lugar, analizar cuáles son los factores que conducen a la exclusión. Y, finalmente, analizar qué efecto tienen las desigualdades previas sobre las trayectorias de los individuos en diferentes puntos del sistema de educación y formación. Creemos que nuestra revisión de tres puntos concretos de las trayectorias de educación y formación en España ha permitido alcanzar tales objetivos. Destacaremos en estas conclusiones algunos de los elementos más relevantes que hemos encontrado en el proceso.

El primero de los puntos de las trayectorias de educación y formación que hemos abordado ha sido la adquisición de competencias al final de la educación obligatoria, estableciendo como límite arbitrario de exclusión la divisoria entre el nivel 1 y el 2 de competencias de la evaluación de PISA. En España, para la competencia de matemáticas, un 23,6% de alumnos queda por debajo de ese límite, proporción similar a la que encontramos para la media europea (23,9%) y de la OCDE (23%), aunque distante del objetivo de la Unión Europea para el año 2020 (15%).

Algunos de los resultados que quisiéramos destacar en cuanto a los factores que inciden sobre la probabilidad de situarse en el grupo de competencias más bajas son los siguientes. Como resultado más destacable, aunque muy previsible, el estatus sociocultural y económico del hogar incide fuertemente sobre tal probabilidad. Asimismo, que el alumno haya cursado educación infantil más de un año se asocia a reducciones significativas de la probabilidad. Resulta también destacable que sólo la condición de inmigrante de primera generación altera significativamente la probabilidad de situarse en el grupo con menos competencias. Los inmigrantes de segunda generación (nacidos en España de padres extranjeros) no presentan, por el contrario, diferencias significativas con respecto a los alumnos nacidos en España de padres españoles; este resultado ofrece una buena señal en cuanto al funcionamiento del sistema educativo a la hora de integrar al alumnado inmigrante.

Por lo que respecta a las variables que actúan en el nivel del centro, destacaremos el efecto no significativo de la titularidad del centro y la incidencia, ésta sí significativa, del nivel de escolarización media de los padres y madres de los estudiantes del centro. El efecto de la concentración de inmigrantes en las escuelas, sin embargo, se circunscribe únicamente a los centros con más del 40% de alumnos inmigrantes. En éstos la probabilidad relativa de estar en los niveles de competencias más bajos es el doble. No encontramos, como es habitual en este tipo de estudios, diferencias significativas de probabilidad vinculadas con el efecto de los recursos (físicos y humanos) del centro. Finalmente, la residencia en algunas Comunidades Autónomas (como Aragón, Castilla y León, Madrid y Navarra) se asocia a probabilidades más bajas de situarse en el grupo de competencias más bajas.

El segundo punto que hemos abordado, correspondiente a las edades donde se mide el abandono prematuro, es a nuestro juicio el más determinante a la hora de establecer las trayectorias educativas y laborales y, por tanto, aquel en el que se concentra el mayor riesgo de exclusión educativa y formativa. El abandono prematuro es un rasgo distintivo del sistema educativo español, en tanto que se sitúa en un valor del 24,9%, más del doble del valor medio europeo. El incremento del desempleo juvenil en el periodo 2007-2013 ha provocado una importante reducción del abandono prematuro; sin embargo, nuestros resultados indican que esta reducción

ha beneficiado proporcionalmente más a los grupos sociales con más recursos económicos y culturales.

Hemos identificado, en nuestro análisis, cómo las variables relativas al origen sociocultural y económico de los jóvenes inciden fuertemente sobre la probabilidad de abandono prematuro. Entre estas variables destaca el efecto de la educación de los progenitores por encima del efecto de la renta y la categoría laboral. Resulta también destacable el efecto no significativo sobre la probabilidad de abandono de la variable relativa al origen inmigrante del joven. La incidencia mayor del abandono prematuro sobre los jóvenes de origen inmigrante se explica, pues, en función de los menores recursos socioculturales y económicos que van asociados frecuentemente a la condición de inmigrante; se descarta, de este modo, la existencia de una discriminación específica vinculada con la nacionalidad.

El efecto de la Comunidad Autónoma de residencia sobre la probabilidad de abandono prematuro se ha modificado sustancialmente durante los años de crisis. Mientras que en 2007 encontrábamos seis Comunidades con un efecto significativo, positivo o negativo, sobre el nivel de abandono, en 2013 encontramos únicamente tres. Antes de la crisis económica las diferencias en los mercados de trabajo de las Comunidades Autónomas provocaba diferencias en el abandono prematuro de los jóvenes. La crisis ha homogeneizado los mercados de trabajo (incrementándose fuertemente, en general, el desempleo juvenil) y ha provocado que estos sean menos relevantes a la hora de determinar el abandono prematuro.

En la tercera y última zona de nuestro análisis, correspondiente a las competencias de la población adulta, nos hemos centrado en la población con los niveles más bajos (1 y menor de 1) en la evaluación de PIAAC de 2012. Se trata de un nivel de competencias que dificulta la participación en el mercado de trabajo y, en general, en la vida social, constituyendo, por tanto, un elemento de potencial exclusión importante. Un 27,5% de la población española entre 25 y 64 años se sitúa en ese nivel, en contraposición a un 15,5% de la población equivalente en los países de la OCDE evaluados. Este dato proporciona una información sintética muy relevante sobre los déficits históricos del sistema de educación y formación en España. Su comparación con los resultados de PISA, considerablemente mejores, permite relativizar las críticas que recibe nuestro sistema educativo, en tanto que los resultados obtenidos por los jóvenes de 15 años se consiguen a pesar del muy reducido nivel de competencias de las generaciones anteriores.

De forma general, los resultados de nuestro análisis sobre la población adulta con menores competencias identifican una pauta de reducida flexibilidad en cuanto a las diferentes vías de adquisición de competencias. Éstas quedan en buena medida establecidas al finalizar la educación formal, de tal manera que las personas

que, en función principalmente de su origen sociocultural, adquieren inicialmente pocas competencias, tienen pocas probabilidades de mejorar su situación por otros medios. En efecto, ni la experiencia laboral ni la participación en la educación no formal son factores que permitan salir del grupo de competencias más bajas.

Un último resultado que desearíamos subrayar es el efecto del nivel educativo de los padres y madres (de las personas ya adultas) sobre la probabilidad de estar situado en el grupo de competencias más bajas de PIAAC-12. Se trata de un "residuo" de desigualdad intergeneracional que produce efectos después de haber concluido la educación formal (cuyo nivel ya depende en buena medida del nivel educativo de los padres).

BIBLIOGRAFÍA

Calero, Jorge (2012). "El acceso a la formación permanente: efectos de la crisis económica". *Revista de Ciencias y Humanidades de la Fundación Ramón Areces*, 7, pp. 73-78.

Calero, Jorge (2008). "What happens after compulsory education? Problems of continuity and possible policies in the case of Spain". *The Social Science Journal*, vol. 45, n. 3, pp. 440-456.

Calero, Jorge y Álvaro Choi (2013a). *Equidad y educación*. Madrid: Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI).

Calero, Jorge y Álvaro Choi (2013b) "Educación y políticas educativas en el estado de las autonomías", en VV.AA. *Informe Comunidades Autónomas* 2012. Barcelona: Instituto de Derecho Público, pp. 121-150.

Calero, Jorge y J. Oriol Escardíbul (2014). *Recursos escolares y resultados de la educación*. Madrid: Fundación Europea Sociedad y Educación.

Calero, Jorge y J. Oriol Escardíbul (2013). "El rendimiento del alumnado de origen inmigrante en PISA-2012". *PISA-2012 Programa para la Evaluación Internacional de los Alumnos*. Informe Español. Volumen II. Análisis secundario. Madrid: Instituto Nacional de Evaluación Educativa, pp. 4-31.

Calero, Jorge y J. Oriol Escardíbul (2009). "Los que se quedan atrás. Un estudio comparativo de la pobreza educativa y sus causas en tres países europeos". V. Navarro (dir.): *La situación social en España* III. Madrid. Biblioteca Nueva.

Calero, Jorge y J. Oriol Escardíbul (2007). "Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003". *Hacienda Pública Española*, Vol. 183(4), pp. 33-66.

Calero, J., J. Oriol Escardíbul y Álvaro Choi (2012). "El fracaso escolar en la Europa mediterránea a través de PISA-2009. Radiografía de una realidad latente". *Revista Española de Educación Comparada*, Vol.19, pp. 69-103.

Calero, Jorge y María Gil (2014). "Too much to pay: an estimation through microsimulation techniques of the monetary costs of early school leaving in Spain". *Journal of Simulation*, vol. 8, n. 4, pp. 314-328.

Calero, Jorge, María Gil y Marcos Fernández (2011). *Los costes del abandono escolar prematuro. Una aproximación a las pérdidas monetarias y no monetarias causadas por el abandono prematuro en España*. Madrid: Ministerio de Educación.

Casquero Tomás, Antonio y María Lucía Navarro (2010). "Determinantes del abandono escolar temprano en España: un análisis por género". *Revista de Educación*, número extraordinario 2010, pp. 191-223.

Diario Oficial de la Unión Europea (2009). "Conclusiones del Consejo de 12 de mayo de 2009 sobre un marco estratégico para la cooperación europea en el ámbito de la educación y la formación". 28 de mayo, pp. 119/02-119/10.

European Commission (2013). *Education and Training Monitor 2013*. Strasbourg: European Commission.

Fernández, Marcos y Jorge Calero (2014). "Los costes no monetarios del abandono educativo prematuro: una estimación en términos de años de buena salud". *Educación XX1*, vol. 17, n. 2, pp. 241-263.

Hansen, Jorgen (2010). "How does academic ability affect educational and labour market pathways in Canada". *OECD Education Working Papers*, No. 30. Paris: OECD Publishing. http://dx.doi.org/10.1787/5kmldw10hl6l-en.

Hanushek, Erik A. (2013). "Economic growth in developing countries:The role of human capital". *Economics of Education Review*, vol.37, pp. 204-212.

Hanushek, Erik A. y Ludger Woessmann (2010). "Education and economic growth". In: P. Peterson, E. Baker y B. McGaw (Eds.), *International Encyclopedia of Education*. Volume 2, pp. 245-252. Oxford: Elsevier.

Hanushek, Erik A. y Ludger Woessmann (2008). "The role of cognitive skills in economic development". *Journal of Economic Literature*, vol. 46(3), pp. 607-668.

Hanushek, Erik A.; G. Schwerdt; S. Wiederhold y L. Woessmann (2015) "Returns to skills around the world: Evidence from PIAAC". *European Economic Review*, vol.73, pp. 103-130.

Hox, Joop (1995). *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties.

INEE (2014). *Sistema estatal de indicadores de la educación*. Edición 2014. Madrid: INEE.

INEE (2013). PISA 2012. *Programa para la evaluación internacional de los alumnos. Informe español. Volumen I: resultados y contexto*. Madrid: Instituto Nacional de Evaluación Educativa.

Instituto de Evaluación (2009). *Sistema estatal de indicadores de la educación. Edición 2009*. Madrid: Instituto de Evaluación.

IVIE (2013). *El abandono educativo temprano. Análisis del caso español*. Valencia: IVIE.

Lee, Jean N. y David Newhouse (2013). "Cognitive skills and youth labor market outcomes". *World Development Report 2013*. Washington, D.C.: The World Bank.

Levin, Henry M., Ilja Cornelisz y Barbara Hanisch-Cerda (2013). "Does educational privatization promote social justice?" *Oxford Review of Education*, vol. 389, n. 4, pp. 514-532.

Mancebón, María Jesús, Jorge Calero, Álvaro Choi y Domingo P. Ximénez-de-Embún (2012). "The efficiency of public and publicly-subsidized high schools in Spain. Evidence from PISA-2006", *Journal of the Operational Research Society*, vol. 63, pp. 1516-1533.

OECD (2013). *OECD Skills Outlook 2013 First Results from the Survey of Adult Skills*. París: OCDE.

OECD (2010). PISA 2009 Results: What Students Know and Can Do. Student Performance in Reading, Mathematics and Science. Volume I. París: OECD.

Sun, Man (2008). *A dynamic analysis of education in Canada: Evidence from the YITS*. MA paper. Department of Economics, Concordia University.

LA EVALUACIÓN DOCENTE BASADA EN EL RESULTADO COMO VÍA DE MEJORA DEL SISTEMA EDUCATIVO

PAU BALART
ANTONIO CABRALES
Universidad Carlos III y UCL Economics

RESUMEN

La investigación en Economía de la Educación ofrece dos resultados referentes al profesorado con importantes implicaciones para la gestión educativa. En primer lugar, el papel del profesor resulta esencial para el éxito del proceso educativo. En segundo lugar, la eficacia docente no se explica por las características observables del profesor, como su nivel de estudios o la experiencia más allá del primer año de ejercicio.

A la vista de estos resultados, una buena gestión educativa requerirá identificar correctamente aquellas personas que puedan llevar a cabo la actividad docente de manera más eficaz. La imposibilidad de identificar aquellos profesores con unas mejores aptitudes docentes a partir de sus características observables (*inputs*) recomienda la sustitución de métodos de evaluación y compensación del profesorado, basados en características como su antigüedad o asistencia a cursos, por métodos más avanzados, basados en los resultados alcanzados por sus alumnos (*output*).

Este capítulo proporciona un ejemplo práctico a partir de datos de la Universidad Carlos III de Madrid que ilustra el modo de llevar a cabo una evaluación docente basada en el resultado de los alumnos. Se discuten algunas dificultades que pueden aparecer en la evaluación y se relaciona su aplicación con otros métodos como las encuestas de satisfacción docente. El estudio permite ver cómo, en caso de disponer de los datos necesarios, se puede llevar a cabo un sistema de evaluación docente más avanzado que permita una mejora del sistema educativo español.

INTRODUCCIÓN

Una correcta gestión educativa requiere entender qué elementos intervienen en el proceso educativo y en qué medida cada uno de ellos contribuye a la eficacia de dicho proceso. El tamaño de las aulas, los recursos materiales o el entorno familiar son algunos de los aspectos que han sido largamente estudiados. Pero, de entre todos estos elementos, hay uno que ha sido objeto de especial atención: el papel del profesorado. Dado que la educación es una actividad intensiva en capital humano, el profesorado constituye su principal factor productivo y el mayor elemento presupuestario de los centros educativos. Por lo tanto, dicha atención está más que justificada. Al mismo tiempo, tanto padres y madres de alumnos como los propios estudiantes otorgan una relevancia especial al papel del profesor (Hanushek, 1986). Por lo tanto, desde un punto de vista de la gestión educativa, será clave responder a la siguiente pregunta: ¿qué grado de importancia tiene el papel del profesor dentro del proceso educativo?

La investigación en Economía de la Educación ha estudiado esta cuestión. Los avances en la disciplina han hecho cambiar notablemente la respuesta que se ha dado a esta pregunta. A pesar de que los primeros resultados en este campo parecían indicar que la actividad docente tiene un papel más bien reducido en la formación del alumno, véase entre otros Coleman *et al.* (1966), otros estudios más recientes, representados principalmente por los trabajos de Erik Hanushek y sus coautores, consideran al profesor como un elemento esencial. La notable diferencia en las conclusiones de estos estudios radica en el enfoque utilizado a la hora de evaluar la importancia del profesorado. Mientras los primeros llevaban a cabo una evaluación basada en los *inputs*, como el nivel de estudios, las acreditaciones u otras características observables del profesor, la literatura más reciente ha utilizado un enfoque basado en el *output* o resultado del proceso educativo, es decir, en medidas del nivel de conocimientos adquiridos por los estudiantes. Este segundo enfoque entiende como efectividad docente o calidad de la docencia la capacidad de un profesor para conseguir que sus estudiantes aprendan.

Una revisión de los trabajos de Hanushek y sus coautores revela dos conclusiones claras. En primer lugar, la docencia tiene un papel más importante del que se había pensado tradicionalmente. Por lo tanto, será importante asignar el capital humano más adecuado para realizar la labor educativa y llevar a cabo una evaluación del profesorado que permita crear los incentivos necesarios para asegurar la efectividad docente. En segundo lugar, los *inputs* y características observables del profesor no son indicativos del grado de eficacia de la labor educativa. Esta conclusión complica el establecimiento de criterios estandarizados para la contratación de buenos docentes y, al mismo tiempo, desaconseja la utilización de medidas ba-

sadas en el *input* como los sexenios, el nivel de estudios o las credenciales a la hora de evaluar y compensar al profesorado.

Este capítulo responde a la pregunta sobre la importancia del profesor a partir de una revisión de la literatura. En segundo lugar, ilustra, con datos obtenidos en la Universidad Carlos III de Madrid, cómo las técnicas usadas por esta literatura permiten llevar a cabo métodos de evaluación del profesorado basados en el resultado.

LA IMPORTANCIA DE LA CALIDAD DOCENTE EN EL PROCESO EDUCATIVO

LAS CARACTERÍSTICAS OBSERVABLES DEL PROFESOR Y SU RELACIÓN CON LA CALIDAD DOCENTE

Los primeros estudios orientados a evaluar la importancia del profesorado en el proceso educativo partían de la base de que una labor educativa eficaz debía poder explicarse a partir de características observables de los profesores o *inputs*. El principal trabajo dentro de este campo fue el Informe Coleman (Coleman at al., 1966), encargado por el gobierno de Estados Unidos. Este trabajo y los numerosos estudios que lo siguieron, se basaban en lo que se conoce como modelo de la función de producción educativa. Dicho modelo consiste en representar el rendimiento académico de los estudiantes como una función de diversos *inputs* observables. Siguiendo a Hanushek y Rivkin (2006), formalmente se puede describir de la siguiente manera:

$$R_g = f(X_g, G_g, P_g)$$

donde, omitiendo un subíndice para cada estudiante, R_g es el resultado académico de cada estudiante en el curso g (generalmente sus notas en tests estandarizados de conocimiento, aunque también puede ser el nivel de estudios alcanzado o incluso su renta una vez finalizado el proceso educativo), y X_g, G_g, y P_g son vectores que contienen valores observables de *inputs* educativos relacionados con la familia y el entorno (por ejemplo, nivel socioeconómico, tipo de unidad familiar o la zona de residencia), el grupo, clase y efectos de pares (por ejemplo, tamaño del grupo o la composición por género) y los profesores (como por ejemplo, su experiencia, titulación o acreditaciones) respectivamente.

Uno de los principales problemas de esta metodología radica en la escasa posibilidad de interpretar como causales la mayoría de las relaciones que puedan existir entre los *inputs* educativos y el rendimiento académico cuando la asignación entre estos no es aleatoria (Rothstein, 2010). Si, por ejemplo, aquellos profesores

con mayor experiencia son capaces de elegir escuelas con mejores estudiantes, esto provocaría un sesgo de simultaneidad a la hora de hacer una interpretación causal de la importancia de la experiencia docente. Por esta razón, en muchos casos, el análisis será llevado a cabo dentro de una misma escuela.

Encontramos otro de los problemas en la posible omisión de variables y su consecuente sesgo en las estimaciones. Entre ellas destaca la habilidad de los estudiantes. Por esta razón, una de las extensiones más utilizadas en esta literatura consiste en utilizar calificaciones anteriores de los estudiantes para controlarlos por su habilidad. De esta forma la función de educación anterior se extendería para incluir el resultado del curso anterior, R_{g-1}.

$$R_g = f(X_g, G_g, P_g, R_{g-1})$$

Generalmente los estudios de este tipo estiman especificaciones lineales de la función anterior:

$$R_g = \mu_0 + \mu_1 R_{g-1} + \mu_2 X_g + \mu_3 G_g + \mu_4 P_g + \varepsilon$$

Donde, μ_0, μ_1, μ_2, μ_3 y μ_4 son vectores de parámetros desconocidos y ε un término de error aleatorio. En la literatura se pueden encontrar distintas especificaciones respecto al parámetro μ_1. Es habitual encontrar especificaciones donde se supone igual a 1 y se utiliza como variable dependiente la diferencia en el resultado entre un curso y el siguiente[1].

Una de las principales conclusiones a las que se llega, después de revisar los numerosos estudios basados en la metodología de la función de producción educativa, es que las características del profesorado tienen un papel muy discreto en la efectividad docente. Véase Hanushek (1997 y 2003) para un exhaustivo resumen estadístico de las conclusiones de estos estudios. Algunos de los *inputs* analizados por la literatura han sido el nivel educativo del profesor, la experiencia docente y las acreditaciones.

La investigación en este campo no ha encontrado evidencia robusta acerca de la existencia de efectos del nivel educativo del profesorado sobre la efectividad docente. Hanushek (1997) ofrece un resumen estadístico donde se aprecia, como en la mayoría de estudios, que la formación del profesor aparece como no significativa e incluso como negativamente asociada con la eficacia docente en más de

1. Este enfoque es especialmente adecuado en aquellos casos en que se dispone de exámenes anuales que hacen una evaluación acumulativa del conocimiento de los estudiantes. Véase por ejemplo Sass *et al.* (2014).

la mitad de los casos donde resulta significativa. Incluso altos niveles educativos, como másteres, no garantizan un mejor resultado de la labor educativa.

Por otro lado, la experiencia adquirida ejerciendo la docencia parece tener un cierto impacto en la eficacia educativa aunque de forma no lineal (Murnane y Philip, 1981). En concreto, estudios más recientes encuentran que la experiencia tan sólo tiene un impacto sobre la calidad docente durante el primer año, véase Chingos y Peterson (2011), Hanushek *et al.* (2005) y Rivkin *et al.* (2005). Es decir, al parecer, la efectividad docente requiere de una serie de habilidades que se adquieren ejerciendo la docencia. Sin embargo su adquisición es rápida y se lleva a cabo, sobre todo, durante el primer año de ejercicio.

Por último, la literatura ofrece resultados diversos a la hora de valorar el papel de las acreditaciones sobre la efectividad docente. Kane *et al.* (2008) encuentran que las certificaciones tienen un papel discreto sobre la efectividad docente, mientras que Clotfelter *et al.* (2007), Goldhaber y Brewer (2000) y Darling-Hammond *et al.* (2001) encuentran que las certificaciones en matemáticas de los profesores mejoran el resultado de los alumnos en pruebas estandarizadas. Sin embargo Jepsen y Rivkin (2002) encuentran que estos efectos no son importantes una vez controlados los efectos no lineales de la experiencia docente comentados en el párrafo anterior.

Una mala interpretación de algunos de estos resultados, sobre todo del Informe Coleman, supuso la generalización de una visión que minimizaba la relevancia del profesor en el éxito educativo. De esta forma se centró la atención en otros aspectos como el estatus socioeconómico familiar o el tamaño del aula, dejando de lado el papel del profesor. Sin embargo, que las características observables del profesor no permitan explicar diferencias en el resultado educativo, en ningún caso significa que no haya diferencias notables en su efectividad docente. En primer lugar, porque generalmente las bases de datos acostumbran a ofrecen una cantidad reducida de variables observables del profesor. En segundo lugar, y de forma más importante, el proceso educativo se lleva a cabo en un aula, donde solo estudiantes y profesor son testigos del proceso formativo. Por lo tanto, pueden existir variables no observables que expliquen en gran parte las diferencias en la calidad docente.

LOS EFECTOS FIJOS Y EL VALOR AÑADIDO DEL PROFESOR: EL PAPEL DEL PROFESOR IMPORTA

Evitando una mala interpretación de los resultados comentados en el apartado anterior, otro grupo de trabajos muestra cómo, a pesar de que las características observables de los profesores no tienen un impacto relevante sobre el resultado educativo de los estudiantes, la influencia del profesor es esencial.

Estos trabajos se basan en la inclusión de efectos fijos que recogen las diferencias en las calificaciones de los estudiantes según su profesor. Formalmente, este enfoque consiste en añadir a las funciones de producción educativas variables ficticias que recojan el efecto de cada profesor:

$$R_g = \mu_0 + \mu_1 R_{g-1} + \mu_2 X_g + \mu_3 G_g + \mu_4 P_g + \pi_j + \varepsilon$$

donde π_j captura el efecto fijo del profesor, por lo tanto, mide el valor añadido de cada uno de ellos[2]. Un buen profesor será aquel que genere un mayor valor añadido en sus alumnos, es decir, un mayor valor de π_j indicará una mayor calidad docente, en términos de rendimiento o de conocimientos adquiridos.

Como detallan Hanushek y Rivkin (2006), este enfoque ofrece varias ventajas. En primer lugar, no está limitado por la necesidad de restringir el efecto del profesor a unas características concretas, y por lo tanto, queda menos sujeto a la disponibilidad de datos de dichas características. En segundo lugar, no requiere de un conocimiento previo sobre la forma funcional en la que actúan las distintas características. En tercer lugar, ofrece un marco para medir la calidad docente que permitirá la comparación y evaluación de políticas educativas.

Numerosos estudios como los de Aaronson, *et al.* (2003), Hanushek (1971, 1992), Hanushek *et al.* (2005), Harbison y Hanushek (1992), Rivkin *et al.* (2005) o Rockoff (2004), entre otros, han utilizado esta metodología para analizar la importancia del profesorado. Estos estudios llegan a las mismas conclusiones que ya habíamos comentado acerca del escaso efecto de las características observables o *inputs* de los profesores[3]. En cambio, los efectos fijos asociados a los profesores captan gran parte de la variación en el resultado de los alumnos. De esta forma, esta literatura encuentra que las diferencias en el valor añadido individual de los profesores son sustanciales. En uno de sus estudios pioneros, Hanushek (1992) mostraba cómo, en Estados Unidos, la diferencia en el aprendizaje entre una docencia de calidad y una de mala calidad llegaba a ser equivalente al conocimiento adquirido durante todo un curso académico. Hanushek y Rivkin (2010) ofrecen una revisión de esta literatura mostrando resultados más modestos, aunque todavía importantes. La diferencia entre tener un profesor situado en el percentil 25 de la distribución de la efectividad docente y tenerlo en el percentil 75, es de alrededor de 0,2 desviaciones estándar en el resultado en matemáticas de un solo año.

---

2. También se pueden evaluar grupos de profesores, como veremos en el apartado "Un ejemplo: la evaluación docente basada en resultados".
3. Una excepción la encontramos en el caso de la Educación Superior con respecto al mérito investigador de los profesores. Rodríguez y Rubio (2013) encuentran que el mérito en investigación medido en publicaciones está relacionado con una mayor efectividad docente.

El enfoque del valor añadido del profesor requiere de la disponibilidad de datos longitudinales tanto a nivel de alumno como de profesor. Por un lado, los datos longitudinales a nivel de alumno permitirán tener un registro de los resultados previos (R_{g-1}) y controlar así la influencia de factores como la habilidad y los conocimientos previos del alumno. Por otro lado, la disponibilidad de datos longitudinales de cada profesor permitirá comprobar la persistencia de los efectos fijos o valor añadido del profesor, evitando el ruido introducido por las características de los grupos impartidos en un curso concreto.

La principal dificultad, a la hora de estimar e interpretar el valor añadido del profesor, será la posible existencia de variables omitidas como podría ser la auto-selección de estudiantes en clases o escuelas. Rothsetein (2010) muestra cómo el sesgo introducido por este problema puede producir resultados erróneos. Si tanto los buenos profesores como los buenos estudiantes son atraídos hacia las mismas escuelas o grupos, se sobrestimará el valor añadido del profesor. Por esta razón, muchos de los estudios se han basado en la estimación del valor añadido del profesor a nivel de escuela, documentando que gran parte de las diferencias se dan dentro de un mismo centro (véase Aaronson *et al.* (2007), Jacob y Lefgren (2008) o Nye *et al.* (2004) entre otros). La posible formación no aleatoria de grupos dentro de la escuela también ha sido estudiada en distintos trabajos. Hanushek *et al.* (2005) y Rivkin *et al.* (2005) encuentran diferencias importantes en el valor añadido del profesor una vez considerada la posible formación no aleatoria de grupos dentro de una misma escuela, aunque su valor es sensiblemente inferior al de los estudios que no controlan esta posibilidad. Kane and Staiger (2008) estiman efectos algo mayores y más cercanos a los encontrados en estudios que no controlan por la influencia de los problemas de selección dentro de los centros. Hanushek y Rivkin (2010a) distinguen entre escuelas susceptibles de estar configurando los grupos de manera aleatoria de aquellas que lo hacen de manera no aleatoria, para el distrito de Texas en Estados Unidos, observando que también existen notables diferencias de valor añadido dentro de aquellas escuelas que configuran los grupos de forma aleatoria. Finalmente, Rockoff (2004) trata el problema de selección en escuelas y en formación de grupos a partir de la introducción de efectos fijos tanto a nivel profesor como estudiante, detectando importantes diferencias en la efectividad docente de los profesores.

En resumen, aunque las conclusiones sobre el tamaño del efecto del profesorado varían dependiendo de las especificaciones y del tratamiento que se dé a los distintos problemas de selección, éstos muestran dos conclusiones claras: en primer lugar, la calidad de la docencia explica notables diferencias en el éxito del proceso educativo, y en segundo lugar, tal y como se apreciaba en el enfoque sin

efectos fijos, estas diferencias en el resultado no se pueden explicar a partir de las características observables de los profesores.

¿CÓMO MEDIR LA CALIDAD DE LA DOCENCIA?

Dada la relevancia de la labor educativa en los resultados educativos de los estudiantes surge una pregunta esencial para la gestión educativa: ¿cómo detectar un buen profesor o profesora?

Tradicionalmente, se han utilizado medidas basadas en los *inputs* del profesor como puede ser su nivel de estudios, su participación en cursos y seminarios o su experiencia para valorar la labor educativa. Otros métodos habituales, sobre todo en instituciones de Educación Superior, se basan en la realización de encuestas sobre la satisfacción docente. Sin embargo, la imposibilidad de vincular la efectividad docente a las características observables del profesor, explicada en el apartado anterior anterior, aconseja la utilización de medidas de evaluación docente basadas en el valor añadido del profesor, es decir, en el *output* del proceso educativo.

Una evaluación educativa basada en los *inputs* del profesor corresponde con la que observamos de forma generalizada en el sistema educativo español (por ejemplo, haciendo depender la remuneración de los profesores de su nivel de estudios o de la acumulación de sexenios). Sin embargo, existen razones de peso para pensar que esta no es la mejor manera de evaluar la labor educativa. En primer lugar, como hemos visto en el apartado anterior, las características observables de los profesores en las que tradicionalmente se han basado tanto la evaluación como la selección del profesorado no están asociadas a unos mejores resultados educativos. En segundo lugar, incluso en el caso que los *inputs* educativos tuvieran poder explicativo sobre la efectividad docente, una evaluación basada en características como el nivel de estudios u otros méritos previos no proporcionará ningún incentivo a la mejora de la efectividad docente. Por lo tanto, este tipo de mecanismos no parecen los ideales a la hora de incentivar la mejora docente.

Otro método común de evaluación del profesorado, sobre todo en las instituciones de Educación Superior, son las encuestas de satisfacción docente. A diferencia del método anterior, éste presenta la ventaja de basarse en *outputs* de la actividad educativa en lugar de en los *inputs*. Sin embargo, no existe evidencia sobre la relación entre las valoraciones que los estudiantes hacen de la labor educativa de un profesor y su efectividad a la hora de transmitir conocimientos. Cabe la posibilidad de que los alumnos no estén basando su evaluación necesariamente en la capacidad educativa del profesor, sino en otros aspectos presentes en la interacción profesor-alumno como podrían ser su simpatía, amabilidad,

agrado o dificultad de la asignatura e, incluso, la laxitud a la hora de calificar a los alumnos. Será necesario profundizar en la relación entre los resultados docentes reales y las encuestas para validar la idoneidad de este último como método de evaluación docente. La literatura en este campo es todavía incipiente, sin embargo, Watchel (1998) muestra la poca confianza de muchos departamentos en este tipo de evaluación. A partir de datos de la Universidad Carlos III de Madrid, en Rodríguez y Rubio (2013) y en el apartado de este documento "Un ejemplo: la evaluación docente basada en resultados" también se cuestiona la validez de las encuestas de satisfacción.

El enfoque económico del valor añadido del profesor, explicado en el apartado anterior, podrá ser utilizado también para evaluar la actividad docente. Sustituyendo la evaluación educativa tradicional basada en los *inputs* por una evaluación educativa basada en el *output*, se evitaría juzgar la labor educativa mediante características para las cuales no existe ninguna evidencia que explique su intervención directa en la efectividad docente.

Como hemos comentado en el apartado anterior, llevar a cabo una evaluación de la docencia basada en el valor añadido del profesor requerirá de datos longitudinales que permitan hacer un seguimiento del resultado académico de los estudiantes y sus profesores. Desafortunadamente, en la actualidad, este tipo de datos se encuentran rara vez disponibles en el sistema educativo español. Sin embargo, tal y como se ilustra a continuación gracias a datos cedidos por la Universidad Carlos III de Madrid, la disponibilidad de este tipo de información permitiría llevar a cabo una evaluación docente basada en el valor añadido del profesorado.

UN EJEMPLO: LA EVALUACIÓN DOCENTE BASADA EN RESULTADOS

Tal y como se desprende de los modelos del valor añadido, la labor del profesor resulta clave en el proceso educativo. Por lo tanto, una buena gestión educativa requerirá identificar aquellas personas que puedan llevar a cabo la actividad docente de manera más eficaz. La imposibilidad de identificar aquellos profesores con unas mejores aptitudes docentes a partir de sus características observables o *inputs* recomienda el uso de métodos de evaluación basados en el resultado a partir del valor añadido del profesor. Por esta razón, será interesante contemplar la posibilidad de emplear los modelos del valor añadido como herramienta para la evaluación y posterior compensación o promoción del profesorado.

La falta de datos educativos es el principal obstáculo a la hora de llevar a cabo este tipo de análisis en el sistema educativo español. Sin embargo, encontramos algunas excepciones. En este caso la posibilidad de disponer de datos longitudinales

facilitados por la Universidad Carlos III de Madrid permitirá ilustrar el funcionamiento de una evaluación basada en el resultado.

El trabajo reciente de Rodríguez y Rubio (2013) utiliza los mismos datos que el presente estudio. A diferencia de la propuesta aquí presentada, su trabajo se centra en analizar si existe relación entre la calidad docente de los profesores y sus méritos investigadores medidos por la cantidad y calidad de sus investigaciones. Estos autores encuentran que los méritos investigadores están positivamente asociados a una mejor efectividad docente. Aunque para realizar su análisis también llevan a cabo un estudio del valor añadido del profesor, existen notables disparidades con las especificaciones utilizadas aquí[4]. A diferencia del trabajo de Rodríguez y Rubio (2013), el presente estudio se centrará en ilustrar el procedimiento y discutir las posibles dificultades que se desprenden de la aplicación de dichos métodos. Por esta razón, se estudiarán distintas especificaciones para medir el valor añadido y comparar sus resultados.

DATOS

La Universidad Carlos III de Madrid se fundó en el año 1989. La comunidad universitaria está formada por más de 18.000 alumnos y dividida en tres campus diferentes situados en Colmenarejo, Getafe y Leganés.

El curso académico consta de dos cuatrimestres lectivos y las asignaturas son de tres horas semanales. La mitad de las horas se dedican a clases de teoría y se imparten en grupos magistrales. La otra hora y media restante se dedica a clases prácticas y de resolución de problemas en grupos reducidos. Para la elaboración de este estudio únicamente se dispone de los datos referentes a los profesores de los grupos reducidos. Los estudiantes de la Universidad Carlos III de Madrid disponen de cuatro convocatorias para aprobar una asignatura. En este trabajo se ha tenido en cuenta únicamente la primera convocatoria de cada uno de los estudiantes.

La organización de los grupos en la Universidad Carlos III de Madrid presenta un elemento que resulta especialmente propicio a la hora de plantear un análisis del valor añadido del profesor. El criterio para asignar los estudiantes en grupos en su primer año de grado es alfabético. Como muestran Bagues, Berniell y Cabrales (2012) con estos mismos datos, no existe correlación entre el orden de los apellidos y las caracte-

4. Una primera diferencia es que el presente estudio utiliza tanto un enfoque del valor añadido individual del profesor como del grupo de profesores. En segundo lugar, existe también una importante diferencia técnica en la manera de estimar el valor añadido de los profesores. Mientras que Rodríguez y Rubio (2013) estiman el valor añadido de cada profesor a partir de la constante de una regresión que incluye únicamente los estudiantes del profesor en cuestión, en este trabajo la regresiones incluyen tanto los estudiantes del profesor como de los demás profesores usando la diferencia para el cálculo del valor añadido.

rísticas observables de los estudiantes, por lo tanto habrá una formación cuasi-aleatoria de grupos. Esto elimina uno de los principales problemas a la hora de llevar a cabo el cálculo del valor añadido del profesor que se ha comentado en la anterior revisión de la literatura: la posible selección endógena de los grupos. Al tratarse de una asignación alfabética se elimina la posibilidad de estar clasificando a los estudiantes según sus resultados evitando así posibles sesgos en las estimaciones del valor añadido.

Para la realización del presente estudio se dispone de datos individualizados que hacen un seguimiento de los estudiantes de grado en Administración de Empresas (ADE), Derecho (DER), Economía (ECO) y Finanzas y Contabilidad (FYCO), así como del doble grado de Derecho y Administración de Empresas para los cursos académicos comprendidos entre 2008 y 2012. Estos grados son impartidos en los campus de Getafe y de Colmenarejo (C), excepto en el caso de los grados de Economía y FYCO que solo se imparten en Getafe. Los datos incluyen información detallada sobre características previas del estudiante como la zona de residencia, género, grado que cursa, nota y vía de acceso; información detallada del expediente del estudiante como notas en las asignaturas (medidas en escala de cero a diez), número de convocatorias e información del grupo en que se cursa la asignatura como el profesor, el tamaño y la composición del grupo, y las evaluaciones docentes que recibe el profesor.

Uno de los primeros pasos necesarios para calcular el valor añadido del profesorado será elegir qué variable usar para medir la eficacia educativa de los profesores. Mucha de la literatura anteriormente mencionada dispone de datos sobre exámenes estandarizados que se utilizan a la hora de valorar el conocimiento adquirido. Las diferencias experimentadas por los estudiantes en este tipo de exámenes constituyen una buena medida del valor añadido, ya que tanto su planteamiento como corrección siguen un proceso homogeneizado. En el ejemplo aquí analizado no se dispone de exámenes estandarizados. A menudo pueden existir diferencias tanto en el tipo de exámenes como en los criterios de corrección de los profesores. Este hecho desaconseja el uso de la nota en cada una de las asignaturas como variable dependiente. Además, emplear la nota de la propia asignatura a la hora de instaurar un programa de evaluación docente crearía un obvio problema de incentivos perversos. Los profesores tendrían incentivos para sobrevalorar las calificaciones de sus estudiantes. El *output* sobre el cual se evalúe la labor docente deberá estar fuera del control del profesor. Por esta razón, el presente estudio tendrá en cuenta las notas de los estudiantes en asignaturas posteriores a la de los profesores evaluados a la hora de calcular el valor añadido del profesor.

Una de las principales dificultades que plantean los datos de la Universidad Carlos III de Madrid usados en este estudio es el posible solapamiento del efecto de

dos o más profesores. Es decir, dado que los estudiantes se asignan a los distintos profesores por orden alfabético, los grupos compartirán un mismo profesor en varias de las asignaturas, lo cual dificulta distinguir entre el efecto de los distintos profesores. El hecho de contar con datos de varios grupos y que los mismos profesores impartan más de un grupo de cada una de las asignaturas minimiza el impacto de este posible solapamiento. Aun así, para eliminar completamente este problema llevaremos a cabo un primer ejemplo de evaluación docente basado en el grupo de profesores. De este modo, se evita confundir los resultados entre profesores a partir de evaluar al conjunto de profesores, en lugar de a un solo profesor. De manera similar, a veces interesará evaluar la calidad docente de una escuela en lugar de la de un solo profesor.

EJEMPLO 1. CÁLCULO DEL VALOR AÑADIDO DE UN GRUPO DE PROFESORES

En este primer ejemplo, se calculará la efectividad docente de los profesores de primer curso sobre los resultados en las asignaturas de segundo curso. Esto supone un total de 3.971 estudiantes y de 169 grupos reducidos. Una vez desagregado el número total de estudiantes para cada una de las asignaturas consideradas se dispone de un total de 30.299 observaciones.

Los datos disponibles ofrecen 30 combinaciones distintas de profesores en el grado de ADE en Getafe, 14 en el grado de ADE en Colmenarejo, 31 en el grado de Derecho en Getafe, 8 de Derecho en Colmenarejo, 16 en el doble grado de ADE y Derecho en Getafe, 14 en el doble grado de ADE y Derecho en Colmenarejo, 36 en el grado de Economía y 20 en el grado de FYCO. Dado que los grupos reducidos se forman alfabéticamente, estos serán fijos. Por lo tanto, se podrá identificar cada uno de los grupos reducidos de primer curso como un grupo de profesores distinto. Dado que las combinaciones de profesores no coinciden entre un año y otro, en este caso no podremos usar los distintos años disponibles para evaluar la estabilidad del valor añadido ni controlar por cohortes.

Como se ha mencionado en la revisión de la literatura anterior, con el fin de calcular el valor añadido de cada profesor deberemos estimar un modelo de efectos fijos. Una manera habitual de estimar los efectos fijos será introducir una variable ficticia para cada uno de los grupos de profesores (excepto uno que será tomado como referencia). De esta forma en el Modelo 1 estimaremos:

$$\text{Modelo 1:} \quad R_{ij} = \mu_0 + \mu_1 N_i + \mu_2 X_i + \mu_3 E_i + \mu_4 M_j + \mu_5 S_i + \pi I_i + \varepsilon_{ij}$$

donde R_{ij} será la nota del estudiante i en la asignatura j de segundo curso; N_i corresponderá a la nota de acceso del estudiante i; X_i informará de otras características

personales del estudiante (sexo, zona geográfica, vía de entrada, si está matriculado a tiempo completo y si ya está en posesión de un título universitario anterior); E_i será un vector de cinco variables ficticias que capturarán las distintas combinaciones de grado y campus posibles y donde el grupo de referencia será el grado en Economía en Getafe; M_j será un vector de dos variables con las características del grupo al que pertenece el estudiante *i*: idioma en el que se imparte la asignatura y número de estudiantes[5]; S_{ij} será un vector con una variable ficticia para cada una de las asignaturas de segundo curso y, finalmente, I_i será un vector compuesto por 161 variables ficticias, una para cada grupo de profesores exceptuando un grupo de referencia para cada combinación de grado y campus. Por lo tanto, el vector de coeficientes π capturará las diferencias en el valor añadido de cada grupo de profesores respecto al grupo de referencia de cada grado.

Con el fin de controlar posibles características no observables de los estudiantes se estimará el modelo mediante efectos aleatorios. Esta técnica permitirá controlar por las características no observables, siempre y cuando no estén correlacionadas con el resto de variables explicativas introducidas en el modelo. Rockoff[6] (2004) combina efectos fijos de estudiante y profesor para controlar cualquier posible fuente de heterogeneidad entre estudiantes. Sin embargo, en nuestro caso, cada estudiante tendrá asociado un solo grupo de profesores lo que impide combinar efectos fijos para estudiante y profesores.

Estimar el valor añadido mediante el Modelo 1 requerirá tomar un grupo de profesores como referencia dentro de cada grado. De este modo, la interpretación de los coeficientes del vector π será respecto al grupo que se tome como referencia. Valores altos de π_g indicarán una mayor efectividad docente del grupo de profesores *g* con respecto al grupo de referencia dentro de su grado y campus. Se estimará el modelo dos veces, una tomando como referencia al grupo de profesores de mayor valor añadido y otra tomando al de menor valor añadido.

Tal y como vemos en la Tabla 1, podemos encontrar varios grupos de profesores con valores significativamente distintos a los mejores y a los peores grupos de profesores dentro de cada grado. Podemos ver que en los grados de Economía y Derecho existe un mayor solapamiento y aun así se pueden distinguir grupos de profesores de primer curso que tienen efectos significativos tanto en términos de un mayor como de un menor resultado en las asignaturas de segundo curso.

---

5. Otra variable que puede afectar significativamente al resultado académico es el horario del curso. Aunque las estimaciones aquí presentadas no incluyen el horario de la asignatura como variable explicativa, los resultados no se ven afectados al incluirlos.
6. Una de las reglas habituales a la hora de elegir entre un modelo convencional de mínimos cuadrados y uno con efectos aleatorios es el test de Breusch and Pagan que, en este caso, recomienda el uso de efectos aleatorios.

TABLA 1. VALOR AÑADIDO DEL GRUPO DE PROFESORES.
(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

ADE				ADE(C)				DER				DER(C)			
MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA	
17	0,137	1	-0,177	7	0,373	1	-0,238	3	0,441	2	-0,319	5	0,502	1	-0,157
	(0,484)		(0,320)		(0,427)		(0,385)		(0,443)		(0,357)		(0,343)		(0,338)
18	0,498	2	-0,233	8	0,589	2	max	4	0,531	5	-0,254	6	min	2	-0,120
	(0,430)		(0,306)		(0,425)	3	-0,0985		(0,425)		(0,334)				(0,341)
19	0,567	3	-0,438	9	0,645		(0,310)	6	0,701	8	-0,376			3	-0,508
	(0,426)		(0,310)		(0,467)	5	-0,296		(0,428)		(0,334)				(0,354)
20	min	4	-0,431	12	0,675		(0,333)	7	0,548	9	-0,431			4	-0,189
21	0,673		(0,320)		(0,424)	6	-0,167		(0,425)		(0,351)				(0,358)
	(0,461)	6	-0,325	13	min		(0,340)	10	0,656	11	-0,570			5	-0,566
			(0,319)	14	0,712				(0,441)		(0,348)				(0,356)
		7	max		(0,457)			12	0,690	12	-0,533			7	-0,0772
		8	-0,231						(0,444)		(0,358)				(0,372)
			(0,319)					14	0,522	13	-0,291			8	max
		12	-0,380						(0,454)		(0,369)				
			(0,323)					18	0,131	15	-0,243				
		14	-0,523						(0,478)		(0,379)				
			(0,327)					19	0,550	16	-0,227				
		15	-0,338						(0,440)		(0,388)				
			(0,342)					20	0,569	17	-0,0957				
		24	-0,528						(0,455)		(0,365)				
			(0,350)					21	0,274	24	-0,360				
		27	-0,00737						(0,449)		(0,357)				
			(0,383)					22	0,514	25	-0,160				
		29	-0,110						(0,419)		(0,374)				
			(0,443)					23	min	27	-0,230				
								24	0,722		(0,346)				
									(0,441)	30	-0,158				
								26	0,472		(0,415)				
									(0,486)	31	max				
								29	0,262						
									(0,434)						

[CONTINÚA EN PÁGINA SIGUIENTE]

TABLA 1. VALOR AÑADIDO DEL GRUPO DE PROFESORES.
(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

DER-ADE				DER-ADE(C)				ECO				FYCO			
MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA	
5	0,441	1	-0,371	1	0,554	3	-0,381	7	0,693	3	-0,0455	6	0,664	10	-0,0484
	(0,343)		(0,286)		(0,353)		(0,304)		(0,460)		(0,331)		(0,435)		(0,472)
8	min	2	max	2	0,221	4	0,0821	13	0,684	4	-0,306	7	0,584	17	max
9	0,301	3	-0,369		(0,384)		(0,291)		(0,482)		(0,324)		(0,412)	18	-0,436
	(0,326)		(0,322)	12	0,324	6	max	18	0,472	5	-0,243	8	0,523		(0,450)
11	0,527	4	-0,349		(0,366)	8	-0,512		(0,470)		(0,330)		(0,410)	19	-0,457
	(0,366)		(0,290)	13	min		(0,371)	22	0,503	6	-0,525	11	min		(0,415)
12	0,510	6	-0,358	14	0,168	9	-0,287		(0,508)		(0,350)		(0,444)		
	(0,345)		(0,307)		(0,347)		(0,311)	24	min	10	max	13	0,516		
14	0,215					10	-0,292	25	0,156	11	-0,130		(0,428)		
	(0,322)						(0,317)		(0,511)		(0,378)				
15	0,428							26	0,287	12	-0,285				
	(0,343)								(0,468)		(0,357)				
16	0,504							27	0,253	14	-0,169				
	(0,372)								(0,483)		(0,381)				
								28	0,629	15	-0,365				
									(0,463)		(0,355)				
								30	0,481	16	-0,155				
									(0,538)		(0,349)				
								31	0,580	17	-0,193				
									(0,494)		(0,368)				
								35	0,400	19	-0,129				
									(0,500)		(0,382)				
								36	0,185	20	-0,259				
									(0,565)		(0,322)				
										23	-0,292				
											(0,384)				
										34	-0,507				
											(0,376)				

La Tabla 1 divide los grupos de profesores en cada uno de los grados y campus en dos sub-columnas. La primera sub-columna muestra aquellos grupos de profesores con valores añadidos que no son significativamente distintos del grupo de profesores con un menor valor añadido dentro de un mismo grado y campus, mientras que la sub-columna de la derecha muestra aquellos que no son significativamente distintos del grupo de profesores con un mayor valor añadido dentro de un mismo grado y campus. Por ejemplo, al fijarse en el caso del doble grado en ADE y Derecho en Colmenarejo (DER-ADE(C)), se puede ver cómo los profesores del grupo identificado con el número 13 son los que obtienen un menor valor añadido al compararlos con el resto de grupos de profesores del mismo grado y campus, mientras que los grupos de profesores identificados con los números 1, 2, 12 y 14 no obtienen un efecto significativamente distinto al del grupo 13. En cambio, el resto de grupos de profesores del doble grado en Derecho y Administración y Dirección de Empresas cursado en Colmenarejo generan resultados significativamente mayores en sus estudiantes.

Alternativamente se podrá estimar un modelo más sencillo en términos de interpretación. Este consistirá en estimar separadamente una regresión para cada uno de los grupos de profesores considerados. Llamaremos a esta especificación Modelo 2:

$$R_{ij} = \mu^r_0 + \mu^r_1 N_i + \mu^r_2 X_i + \mu^r_3 E_i + \mu^r_4 M_j + \mu^r_5 S_i + \mu^r_6 A_{ir} + \pi^r I^r_i + \varepsilon^r_{ij} \text{ con } r=1, \ldots, 169$$

Una de las ventajas de este modelo reside en que, al no incluir una variable ficticia para cada uno de los grupos, permitirá controlar por la habilidad media del grupo, concretamente incluye la variable A^r_i. A diferencia del modelo anterior, ahora I^r_i será un escalar que tomará el valor 1 cuando el estudiante pertenezca al grupo *r*. El resto de variables explicativas serán las mismas que se habían incluido en el Modelo 1, con el único cambio de que ahora el superíndice *r* nos informará del grupo para el cual se está estimando el efecto fijo. Como antes, *r*=1,.., 169 denotará cada uno de los grupos de profesores para los cuales queremos evaluar el valor añadido. Por lo tanto, se realizarán un total de 30 regresiones para los profesores en el grado de ADE en Getafe, 14 en el grado de ADE en Colmenarejo, 31 para el grado de Derecho en Getafe, 8 para Derecho en Colmenarejo, 16 para el doble grado de ADE y Derecho en Getafe, 14 para el doble grado de ADE y Derecho en Colmenarejo, 36 para el grado de Economía y 20 para el grado de FYCO.

Nótese que mediante este procedimiento los coeficientes de las variables explicativas podrán variar en cada una de las regresiones. Idealmente, estos coeficientes deberían ser fijos como en el Modelo 1. Sin embargo, la variación de estos es mínima y, por lo tanto, este hecho no representará un problema para el modelo.

En cambio, la estimación del coeficiente del valor añadido, π^r variará notablemente. Este modelo permitirá una interpretación mucho más intuitiva de la evaluación de los profesores. En este caso, valores positivos de π^r estarán asociados con valores añadidos del profesor por encima de la media dentro del mismo grado y campus.

En la Tabla 2 se presentan de nuevo aquellos grupos de profesores que habían sido identificados con unos mejores y peores resultados en el modelo anterior. Sin embargo, ahora solo se incluyen las estimaciones de aquellos coeficientes significativamente distintos de cero. Se puede observar cómo, estableciendo un criterio de identificación basado únicamente en un valor añadido significativamente mayor o menor a la media, se reducen notablemente los grupos de profesores que aparecen como destacados. Esto sucede especialmente en aquellos grados donde había una mayor cantidad de grupos de profesores con resultados que no eran significativamente distintos de los mejores y peores resultados, como ocurría en el grado en Derecho en el campus de Getafe o con el grado en Economía. Al tomar como ejemplo el grado de Derecho en Getafe se puede observar cómo de los dieciséis grupos de profesores que producían peores resultados, tan solo el grupo de profesores identificado con el número 29 está asociado con un resultado en segundo curso significativamente menor a la media[7]. En concreto, se observa cómo los estudiantes que cursaron las asignaturas de primer curso con el grupo de profesores 29, obtuvieron de media 0.726 puntos menos (con puntuaciones medidas en escala del 0 al 10) en las calificaciones de segundo curso que el resto de alumnos que cursaron Derecho en el campus de Getafe. De manera parecida, los estudiantes que cursaron la asignatura de primero de Derecho en Getafe con el grupo de profesores 5 y 24 obtuvieron un resultado en segundo curso más de cuatro décimas por encima de la media de sus compañeros.

7. Como se puede observar, el grupo de profesores con unos resultados significativamente inferiores a la media de los demás grupos no siempre coincide con aquel que tenía un menor coeficiente en la estimación del Modelo 1. Esto se puede deber tanto a diferencias en el error estándar del coeficiente del grupo de profesores en cada una de las distintas regresiones como a los cambios inducidos al introducir una nueva variable explicativa que controle la habilidad media del grupo.

TABLA 2. VALOR AÑADIDO DEL GRUPO DE PROFESORES.
(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

ADE				ADE(C)				DER				DER(C)			
MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA	
17	-0,984	1	0,434	7		1		3		2		5		1	
	(0,331)		(0,234)	8		2	0,667	4		5	0,430	6	-0,889	2	
18	-0,713	2	0,395	9			(0,233)	6			(0,210)		(0,254)	3	
	(0,248)		(0,214)	12		3	0,590	7		8				4	
19	-0,754	3		13	-0,903		(0,240)	10		9				5	
	(0,240)	4			(0,321)	5	0,452	12		11				7	
20	-1,074	6		14			(0,270)	14		12					
	(0,352)	7	0,626			6	0,523	18		13					
21	-0,504		(0,239)				(0,278)	19		15					
	(0,299)	8	0,584					20		16					
			(0,229)					21		17					
		12						22		24	0,425				
		14						23			(0,245)				
		15						24		25					
		24						26		27					
		27						29	-0,726	30					
		29							(0,225)						
										31					

[CONTINÚA EN PÁGINA SIGUIENTE]

TABLA 2. VALOR AÑADIDO DEL GRUPO DE PROFESORES.
(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

DER-ADE				DER-ADE(C)				ECO				FYCO			
MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA		MENOR VA		MAYOR VA	
5		1	0,360	1		3		7		3	0,609	6		10	0,947
8	-0,579		(0,217)	2	-0,463	4	0,633	13	-0,479		(0,201)	7			(0,344)
9	(0,252)	2	0,722		(0,281)		(0,194)		(0,265)	4		8		17	0,931
11			(0,217)	12	-0,583	6	0,475	18		5	0,593	11	-0,913		(0,348)
12		3			(0,260)		(0,244)	22	-0,514		(0,209)		(0,290)	18	
14	-0,666	4		13	-0,933	8			(0,270)	6		13		19	
	(0,209)	6	0,421		(0,266)	9		24	-0,807	10	0,610				
15			(0,238)	14	-0,663	10			(0,398)		(0,240)				
16					(0,235)			25		11					
								26	-0,405	12					
									(0,242)						
								27	-0,453	14					
									(0,270)						
								28		15					
								30		16					
								31	-0,668	17					
									(0,246)						
								35		19					
								36		20	0,411				
											(0,248)				
										23					
										34					

Como se ha mencionado anteriormente, al evaluar el valor añadido del grupo de profesores se evita el problema de tener que distinguir entre el efecto de los distintos profesores. Por ejemplo, la asignatura de Teoría Microeconómica, que se imparte en segundo en todos los grados incluidos en este estudio excepto en el de Derecho, requiere tanto de conocimientos de teoría económica de la asignatura Microeconomía como de conocimientos que se imparten en la asignatura de Matemáticas II. Al disponer tan solo de datos de cuatro cursos distintos podría no haber suficientes combinaciones entre profesores de Matemáticas II y Microeconomía, lo cual podría llevar a confundir el efecto de un profesor con el otro. El enfoque basado en el grupo de profesores reconoce este problema y evalúa directamente los grupos de profesores. Sin embargo, el enfoque basado en el grupo de profesores presenta otro problema. Al calcular el valor añadido del grupo de profesores, podríamos estar capturando diferencias que emanen de los propios estudiantes en lugar de ser atribuibles a la labor del profesor. Aunque la asignación alfabética de los alumnos pueda considerarse como aleatoria, al estar tratando con grupos relativamente pequeños, de un máximo de cuarenta estudiantes, no se puede descartar que algunos de ellos estén formados por grupos de distinta habilidad. Este problema se solventa parcialmente al incluir en el Modelo 2 un control de la habilidad media del grupo basada en las notas de selectividad. No obstante, si existen otros motivos que provoquen diferencias en el rendimiento del grupo de estudiantes como pueda ser el ambiente en clase o la cooperación entre ellos, este enfoque no estará evaluando solamente la labor del profesor sino también el rendimiento del grupo de estudiantes.

Este problema no estará presente en el siguiente ejemplo de cálculo del valor añadido individual del profesor, ya que cada uno de los profesores imparte docencia en varios grupos. Además, en este caso, será posible comprobar la estabilidad de las estimaciones en distintos cursos académicos lo cual dotará de mayor robustez las evaluaciones.

EJEMPLO 2. CÁLCULO DEL VALOR AÑADIDO DEL PROFESOR

Este segundo ejemplo ilustrará cómo se podría llevar a cabo una evaluación individual de los profesores. En este caso se evaluará a los profesores de la asignatura de "Matemáticas para la Economía I" impartida en el primer cuatrimestre en los grados de Economía y Administración de Empresas, así como del doble grado en Derecho y Administración de Empresas (impartidos en Getafe y en Colmenarejo). Esto supone un total de 4.170 estudiantes.

TABLA 3. VALOR AÑADIDO DEL PROFESOR DE MATEMÁTICAS SOBRE LAS NOTAS DE LAS ASIGNATURAS DEL SEGUNDO CUATRIMESTRE (ESTADÍSTICA I, INTRODUCCIÓN A LA CONTABILIDAD, MATEMÁTICAS II Y MICROECONOMÍA)

(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

	EFECTOS ALEATORIOS (VA1)		MCO (VA2)		PRINCIPIOS DE ECONOMÍA (VA3)		FUNDAMENTOS ADMON. EMP. (VA4)		DERECHO CIVIL (VA5)	
PROFESOR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR
1	0,262**	(0,128)	0,264***	(0,082)	0,408***	(0,105)	0,169	(0,126)	0,297**	(0,129)
2	-0,0120	(0,168)	-0,00261	(0,106)	-0,191	(0,139)	-0,0473	(0,125)	-0,0749	(0,122)
3	-0,299*	(0,157)	-0,296***	(0,104)	-0,230*	(0,135)	-0,305**	(0,119)	-0,424***	(0,118)
4	-0,444**	(0,213)	-0,363***	(0,140)	-0,487***	(0,149)	-0,531***	(0,153)	-0,444***	(0,151)
5	-0,0828	(0,108)	-0,102	(0,070)	-0,0420	(0,085)	0,0789	(0,129)	0,157	(0,115)
6	-0,0938	(0,164)	-0,0746	(0,104)	-0,0225	(0,126)	-0,457***	(0,153)	-0,116	(0,135)
7	0,113	(0,104)	0,105	(0,069)	0,104	(0,084)	0,227**	(0,092)	0,123	(0,077)
8	0,107	(0,190)	0,149	(0,126)	0,156	(0,149)	-0,0886	(0,202)	-0,276	(0,200)
9	0,110	(0,119)	0,0979	(0,082)	0,0526	(0,097)	0,198*	(0,106)	0,206**	(0,105)
10	-0,0331	(0,137)	-0,0175	(0,087)	-0,0526	(0,096)	(,)	(,)	(,)	(,)
11	0,00899	(0,114)	0,00676	(0,077)	-0,0548	(0,126)	0,123	(0,098)	0,0836	(0,099)
12	0,137	(0,133)	0,151*	(0,084)	0,0631	(0,098)	-0,0731	(0,186)	0,151	(0,112)
13	-0,0476	(0,143)	-0,0694	(0,091)	-0,199**	(0,097)	-0,377	(0,252)	-0,527**	(0,238)
14	-0,373***	(0,143)	-0,369***	(0,090)	-0,548***	(0,105)	-0,198	(0,129)	-0,397**	(0,162)
15	0,112	(0,132)	0,125	(0,085)	0,527***	(0,134)	-0,0674	(0,132)	-0,0674	(0,119)
16	0,133	(0,110)	0,158**	(0,073)	0,319***	(0,090)	0,263***	(0,099)	-0,0109	(0,094)
17	-0,405*	(0,237)	-0,415***	(0,149)	-0,0986	(0,172)	(,)	(,)	(,)	(,)
18	0,230*	(0,134)	0,193**	(0,085)	0,158	(0,112)	0,150	(0,130)	0,182	(0,116)
19	-0,489**	(0,191)	-0,504***	(0,123)	-0,652***	(0,132)	-0,531***	(0,135)	-0,654***	(0,139)
20	-0,117	(0,109)	-0,121*	(0,069)	-0,368***	(0,088)	-0,178*	(0,097)	-0,146	(0,107)
21	0,0329	(0,091)	0,0253	(0,058)	0,360***	(0,076)	0,334**	(0,139)	0,305**	(0,121)
22	-0,308***	(0,109)	-0,313***	(0,072)	-0,470***	(0,092)	-0,124	(0,100)	-0,280***	(0,086)
23	-0,127	(0,118)	-0,152**	(0,077)	-0,168**	(0,082)	-0,0204	(0,143)	0,00941	(0,121)
24	0,351	(0,246)	0,377**	(0,155)	0,391**	(0,166)	0,387	(0,286)	0,286	(0,287)
25	0,128	(0,316)	0,119	(0,223)	-0,120	(0,231)	0,00552	(0,235)	0,00627	(0,249)

[CONTINÚA EN PÁGINA SIGUIENTE]

TABLA 3. VALOR AÑADIDO DEL PROFESOR DE MATEMÁTICAS SOBRE LAS NOTAS DE LAS ASIGNATURAS DEL SEGUNDO CUATRIMESTRE (ESTADÍSTICA I, INTRODUCCIÓN A LA CONTABILIDAD, MATEMÁTICAS II Y MICROECONOMÍA)

(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

	EFECTOS ALEATORIOS (VA1)		MCO (VA2)		PRINCIPIOS DE ECONOMÍA (VA3)		FUNDAMENTOS ADMON. EMP. (VA4)		DERECHO CIVIL (VA5)	
PROFESOR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR	VA	ERROR ESTÁNDAR
26	0,280*	(0,166)	0,328***	(0,111)	0,303**	(0,133)	0,0315	(0,160)	-0,00667	(0,159)
27	0,247**	(0,106)	0,246***	(0,070)	0,200*	(0,104)	0,202	(0,127)	0,476***	(0,132)
28	-0,0137	(0,178)	-0,0212	(0,113)	0,808	(0,581)	0,0108	(0,169)	-0,0409	(0,152)
29	0,182	(0,158)	0,207*	(0,106)	0,384***	(0,135)	0,000736	(0,227)	0,592	(0,974)
30	0,0314	(0,219)	-0,00778	(0,157)	0,424**	(0,189)	-0,0898	(0,195)	0,319	(0,207)
31	0,126	(0,153)	0,136	(0,103)	-0,140	(0,141)	-0,00418	(0,152)	0,0835	(0,141)
32	-0,0656	(0,190)	-0,0192	(0,125)	-0,105	(0,138)	-0,0188	(0,131)	-0,0212	(0,134)
33	-0,227	(0,158)	-0,233**	(0,102)	-0,112	(0,117)	-0,213	(0,217)	-0,278	(0,219)
34	-0,229	(0,180)	-0,272**	(0,121)	-0,130	(0,143)	-0,0483	(0,135)	-0,244*	(0,135)
35	0,0996	(0,106)	0,133*	(0,069)	0,212**	(0,084)	0,200**	(0,082)	0,227***	(0,076)
36	-0,106	(0,275)	-0,145	(0,174)	(,)	(,)	-0,283	(0,186)	-0,0757	(0,198)

Una vez más se basará la evaluación en una medida del *output* fuera del control de los profesores que se pretende evaluar, en este caso las notas en el segundo cuatrimestre en las asignaturas comunes para los grados considerados: Estadística I, Introducción a la Contabilidad, Matemáticas II y Microeconomía. Considerar únicamente el resultado en las asignaturas del segundo cuatrimestre permitirá mitigar los posibles efectos derivados de la labor de otros profesores ya que los estudiantes tan solo habrán recibido un cuatrimestre de clases en la universidad. Una vez combinado el número de estudiantes y sus calificaciones en las distintas asignaturas del segundo cuatrimestre quedan un total de 14.910 observaciones.

Se dispone de datos para un total de 36 profesores de Matemáticas I. Se estimará el valor añadido de cada uno de ellos a partir de la metodología del Modelo 2 comentada en el apartado anterior, con la única diferencia de que, en este caso, se añadirá un grupo de variables ficticias que controlen los posibles efectos de cohortes.

Como podemos ver en la primera columna de la Tabla 3, se pueden encontrar profesores de Matemáticas I cuyos estudiantes obtienen resultados significativa-

mente superiores y otros con resultados significativamente inferiores a la media, dentro de su grado y campus. Esta diferencia puede ser de algo más de tres décimas por encima de la media en el caso del profesor con mejores resultados (identificado con el número 27) y de hasta casi medio punto menos en el caso del profesor con peores resultados (identificado con el número 14). La desviación estándar del valor añadido del profesor es de 0,215. Por lo tanto, una desviación estándar en el valor añadido del profesor genera un aumento medio de 0,215 puntos adicionales en la nota de las asignaturas posteriores consideradas. La segunda columna muestra que las estimaciones son parecidas si estimamos el modelo sin efectos aleatorios individuales.

Tal y como se ha comentado anteriormente, una de las principales dificultades al evaluar individualmente a los profesores en función de su *output* consistirá en distinguir correctamente el efecto del profesor de matemáticas del de otras disciplinas relacionadas. Para comprobar que estos efectos no constituyen un problema, las estimaciones de las columnas restantes introducen efectos fijos para controlar la posible influencia de los demás profesores del primer cuatrimestre (únicamente para aquellas asignaturas que son comunes para todos los estudiantes incluidos en la muestra). De este modo en las columnas restantes se controla por la posible influencia de los profesores de "Principios de Economía", "Fundamentos de administración de empresas" e "Introducción al Derecho Civil Patrimonial". Se puede observar que, en algunos casos, profesores que tenían un valor añadido significativamente distinto a la media dejan de serlo después de tener en cuenta estos efectos, este es el caso para los profesores identificados con los números 18 y 26. En cambio hay varios profesores cuyo valor añadido sigue siendo significativamente distinto a la media en todas las especificaciones, como es el caso de los profesores 3, 4 y 19. En la Tabla 3 se han resaltado aquellos profesores cuyos resultados difieren significativamente de la media en la mayoría de las especificaciones. Se observa que los profesores 1, 14, 17, 22, 27 y 35 tienen valores añadidos significativamente distintos de la media en todas las especificaciones excepto en una, por lo tanto, no podríamos descartar completamente que las estimaciones estén capturando únicamente el valor añadido del profesor de matemáticas.

A la luz de los resultados de la Tabla 3, la implantación de un sistema de evaluación de la docencia sugeriría recompensar a los profesores 1, 27 y 35. Al mismo tiempo se debería prestar atención al modo de mejorar las posibles deficiencias en la labor docente de los profesores 3, 4, 14, 17, 19 y 22. Aunque el valor añadido de estos profesores es significativamente inferior a la media en la mayoría de las especificaciones, se deberá tener cierta precaución a la hora de tomar medidas ya que, como hemos visto, algunos de estos profesores podrían estar capturando el efecto de otras

asignaturas del primer cuatrimestre (profesores 14, 17 y 22). Aun así, este enfoque permitiría centrar la atención y los esfuerzos de evaluación y mejora en aquellos docentes para los cuales podría ser más necesario incrementar su efectividad docente.

Las regresiones de la Tabla 3 evalúan a los profesores basándose en los resultados de todos sus alumnos durante los cursos entre 2008 y 2012. Sin embargo, sería interesante llevar a cabo un seguimiento anual de los resultados. Esto proporcionaría dos ventajas. En primer lugar, reduciría las posibles diferencias en las medidas de *output*. Por ejemplo, podría ocurrir que la dificultad del examen de una asignatura del segundo cuatrimestre variara de un año a otro[8]. Evaluar a los profesores año a año permite controlar este tipo de efectos. En segundo lugar, mediante la comparación de la evaluación de los profesores en distintos cursos académicos, se podrá verificar la estabilidad de sus resultados, y hacer un seguimiento más preciso de su labor educativa. En este aspecto, los datos disponibles presentan ciertas limitaciones. Como se observa en la Tabla 4, tan solo 3 de los 36 profesores evaluados impartieron clase de Matemáticas I ininterrumpidamente entre los cursos 2008 y 2012. Ninguno de ellos está entre los que anteriormente habíamos identificado con efectos significativamente distintos de la media al considerar todos los cursos académicos de forma agregada. Por ejemplo, se puede ver cómo el profesor identificado con el número 21, para el que se dispone de datos para todos los años, obtiene resultados significativamente por encima de la media en los años 2008 y 2011 y resultados levemente y de manera no significativa por debajo de la media durante los otros tres cursos evaluados.

Fijando la atención en aquellos profesores que obtenían resultados significativamente distintos de la media sin separar por años académicos (Tabla 3), se observa cómo en el caso de los profesores 4, 29 y 33 únicamente se dispone de datos para un solo curso académico, lo que impediría comprobar la estabilidad temporal de sus evaluaciones. A pesar de no disponer de datos para todos los años académicos se puede observar cómo, por ejemplo, el profesor identificado con el número 27 tiene valor añadido por encima de la media (con la única excepción del año 2011, cuando presenta una evaluación levemente y de manera no significativa por debajo de la media). Aunque el signo positivo del valor añadido se cumple para todos los años excepto 2011, se puede apreciar cómo las desviaciones no son significativas en cada uno de los años. Una de las consecuencias de evaluar separadamente los distintos cursos académicos es que se reduce notablemente el número de obser-

8. En los resultados expuestos en la Tabla 3 este efecto se controlaba mediante la introducción de las variables ficticias de cohorte.

vaciones, lo que aumenta los errores estándar de las estimaciones y, en consecuencia, disminuyen los profesores con evaluaciones significativamente distintas de la media. Algo similar ocurre con los profesores 14, 19 y 22 pero con resultados por debajo de la media.

TABLA 4. VALOR AÑADIDO DEL PROFESOR DE MATEMÁTICAS SOBRE LAS NOTAS DE LAS ASIGNATURAS DEL SEGUNDO CUATRIMESTRE (ESTADÍSTICA I, INTRODUCCIÓN A LA CONTABILIDAD, MATEMÁTICAS II Y MICROECONOMÍA). DESGLOSADO POR AÑOS.
(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

	2008		2009		2010	
PROFESOR	**VA**	**ERROR ESTÁNDAR**	**VA**	**ERROR ESTÁNDAR**	**VA**	**ERROR ESTÁNDAR**
1	-0,032	(0,213)	-0,112	(0,251)	0,303	(0,481)
2	-0,172	(0,257)	-1,907	(1,377)	(,)	(,)
3	-0,281	(0,249)	0,282	(1,382)	-0,614*	(0,364)
4	(,)	(,)	(,)	(,)	-0,656**	(0,247)
5	-0,531***	(0,198)	-0,342	(0,212)	0,346	(0,312)
6	(,)	(,)	(,)	(,)	0,137	(0,332)
7	0,307*	(0,174)	-3,686**	(1,525)	0,601*	(0,312)
8	(,)	(,)	(,)	(,)	(,)	(,)
9	0,090	(0,266)	-0,245	(0,235)	0,717**	(0,303)
10	(,)	(,)	0,103	(0,249)	-0,0926	(0,314)
11	-0,111	(0,189)	-0,136	(0,297)	(,)	(,)
12	(,)	(,)	-0,028	(0,271)	-0,472	(0,320)
13	(,)	(,)	0,904	(1,285)	-0,257	(0,284)
14	(,)	(,)	0,125	(0,254)	-0,300	(0,440)
15	0,524**	(0,252)	0,227	(0,206)	(,)	(,)
16	-0,121	(0,169)	0,620**	(0,288)	0,458	(0,291)
17	(,)	(,)	(,)	(,)	(,)	(,)
18	0,383**	(0,182)	0,148	(0,206)	(,)	(,)
19	(,)	(,)	(,)	(,)	(,)	(,)
20	0,596*	(0,323)	-0,229	(0,203)	(,)	(,)
21	0,615***	(0,219)	-0,146	(0,250)	-0,162	(0,277)
22	-1,082***	(0,172)	0,005	(0,198)	(,)	(,)
23	-0,358**	(0,173)	-0,091	(0,245)	(,)	(,)
24	(,)	(,)	(,)	(,)	0,416	(0,354)

[CONTINÚA EN PÁGINA SIGUIENTE]

TABLA 4. VALOR AÑADIDO DEL PROFESOR DE MATEMÁTICAS SOBRE LAS NOTAS DE LAS ASIGNATURAS DEL SEGUNDO CUATRIMESTRE (ESTADÍSTICA I, INTRODUCCIÓN A LA CONTABILIDAD, MATEMÁTICAS II Y MICROECONOMÍA). DESGLOSADO POR AÑOS.
(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

	2008		2009		2010	
PROFESOR	**VA**	**ERROR ESTÁNDAR**	**VA**	**ERROR ESTÁNDAR**	**VA**	**ERROR ESTÁNDAR**
25	(,)	(,)	-0,619	(1,312)	0,198	(0,368)
26	(,)	(,)	(,)	(,)	(,)	(,)
27	0,319	(0,202)	0,380	(0,232)	0,100	(0,283)
28	(,)	(,)	(,)	(,)	(,)	(,)
29	0,143	(0,166)	(,)	(,)	(,)	(,)
30	0,454*	(0,243)	(,)	(,)	(,)	(,)
31	0,341	(0,229)	-0,333	(0,233)	1,325	(1,137)
32	(,)	(,)	-0,429	(0,322)	0,0992	(0,285)
33	(,)	(,)	(,)	(,)	-0,0364	(0,232)
34	(,)	(,)	0,241	(0,261)	(,)	(,)
35	-0,978	(0,126)	0,291	(0,204)	-0,197	(0,192)
36	(,)	(,)	(,)	(,)	(,)	(,)

	2011		2012	
PROFESOR	**VA**	**ERROR ESTÁNDAR**	**VA**	**ERROR ESTÁNDAR**
1	0,664*	(0,383)	(,)	(,)
2	(,)	(,)	0,548**	(0,271)
3	-1,126	(1,351)	-0,356	(0,302)
4	(,)	(,)	(,)	(,)
5	0,262	(0,276)	(,)	(,)
6	0,337	(0,262)	(,)	(,)
7	-0,525**	(0,254)	-0,229	(0,229)
8	(,)	(,)	-0,049	(0,225)
9	-0,320	(0,256)	(,)	(,)
10	0,085	(0,347)	-0,040	(0,328)
11	0,222	(0,260)	-0,102	(0,316)
12	0,047	(0,282)	0,139	(0,299)
13	-0,004	(0,335)	0,040	(0,328)
14	-0,390	(0,274)	(,)	(,)

[CONTINÚA EN PÁGINA SIGUIENTE]

TABLA 4. VALOR AÑADIDO DEL PROFESOR DE MATEMÁTICAS SOBRE LAS NOTAS DE LAS ASIGNATURAS DEL SEGUNDO CUATRIMESTRE (ESTADÍSTICA I, INTRODUCCIÓN A LA CONTABILIDAD, MATEMÁTICAS II Y MICROECONOMÍA). DESGLOSADO POR AÑOS.
(Errores estándar entre paréntesis. ***, ** y * denotan significatividad al 1%, 5% y 10% respectivamente)

	2011		2012	
PROFESOR	**VA**	**ERROR ESTÁNDAR**	**VA**	**ERROR ESTÁNDAR**
15	(,)	(,)	-0,151	(0,301)
16	0,212	(0,310)	(,)	(,)
17	(,)	(,)	-0,159	(0,298)
18	(,)	(,)	(,)	(,)
19	-0,706***	(0,244)	-0,094	(0,450)
20	-0,269	(0,281)	0,131	(0,213)
21	0,389*	(0,232)	-0,345	(0,310)
22	(,)	(,)	0,111	(0,232)
23	(,)	(,)	(,)	(,)
24	0,331	(0,438)	(,)	(,)
25	(,)	(,)	(,)	(,)
26	0,463	(0,367)	-0,006	(0,332)
27	-0,138	(0,270)	(,)	(,)
28	(,)	(,)	0,094	(0,237)
29	(,)	(,)	(,)	(,)
30	(,)	(,)	(,)	(,)
31	(,)	(,)	(,)	(,)
32	(,)	(,)	(,)	(,)
33	(,)	(,)	(,)	(,)
34	-0,811**	(0,320)	(,)	(,)
35	0,686***	(0,240)	(,)	(,)
36	(,)	(,)	0,206	(0,310)

Finalmente, podremos usar la correlación entre las puntuaciones recibidas por los profesores en las encuestas docentes y el valor añadido para verificar la bondad del método de las encuestas a la hora de determinar la efectividad docente de un profesor. La Tabla 5 indica la correlación entre el grado de satisfacción de los estudiantes y el valor añadido de cada profesor calculado según cada una de las especificaciones de la Tabla 3. Como podemos ver, la correlación entre ambas medidas es muy cercana a cero en todas las especificaciones. Por lo tanto el método de las encuestas no parece adecuado para llevar a cabo las evaluaciones docen-

tes. Este resultado es consistente con los encontrados en el trabajo de Braga *et al.* (2014), donde muestran que existe correlación positiva entre las notas que asigna el profesor y su valoración en las encuestas pero no entre el resultado en las encuestas y el valor añadido del profesor, medido como su impacto en las notas en cursos posteriores.

TABLA 5. CORRELACIÓN ENTRE CALIDAD DOCENTE Y PUNTUACIÓN EN LAS ENCUESTAS DE EVALUACIÓN DOCENTE

VALOR AÑADIDO DEL PROFESOR	PUNTUACIÓN EN LA ENCUESTA DE SATISFACCIÓN
VA1	0,0853
VA2	0,0902
VA3	0,0955
VA4	0,0554
VA5	0,1148

Este es un resultado que convendría estudiar con mayor profundidad. Cabe resaltar que en la Universidad Carlos III de Madrid la participación en estas encuestas es voluntaria y se lleva a cabo fuera de horas lectivas. Por un lado, esto supone una muy baja participación, además de un problema de autoselección; tan solo algunos estudiantes muy satisfechos o muy insatisfechos con la docencia podrían estar dispuestos a responder. Por lo tanto, los resultados de este estudio muestran poca relación entre el valor añadido de los profesores y sus evaluaciones en las encuestas.

CRÍTICAS Y DIFICULTADES DEL MÉTODO DEL VALOR AÑADIDO DEL PROFESOR

Como se ha comentado anteriormente, la importancia de la calidad docente junto a su no observabilidad aconsejan evaluar al profesorado mediante medidas basadas en sus resultados. Sin embargo, existen algunas críticas metodológicas y dificultades que se describen a continuación. Estas críticas deberán ser especialmente tenidas en cuenta ante la posibilidad de implementar las medidas de valor añadido como forma de evaluación y posterior compensación o promoción del profesorado.

Una de las principales críticas metodológicas gira alrededor de la fiabilidad de los tests. Como se ha descrito en la revisión de la literatura, la mayoría de estudios del valor añadido del profesor toman como medida del *output* la puntuación de los estudiantes en pruebas estandarizadas de conocimientos. Recientemente se ha cuestionado qué miden exactamente este tipo de pruebas, véase Jacob *et al.*

(2008), Kane and Staiger (2008), Rothstein (2010) o Segal (2008) entre otros. Sin embargo, la literatura también ha aportado evidencia sustancial sobre la relación entre la puntuación en pruebas estandarizadas con la renta futura de los estudiantes y el nivel de educación alcanzado, véase Murnane *et al.* (1995). Por otro lado, la no observabilidad directa del conocimiento alerta sobre la posible presencia de errores de medición en las pruebas estandarizadas. En este sentido, la mayoría de estudios sobre valor añadido usan estimadores del valor añadido que corrigen estos posibles errores de medición. Aunque tener en cuenta esta posibilidad reduce las diferencias en el valor añadido de los profesores, éstas siguen siendo importantes, véase Hanushek y Rivkin (2010). Por otro lado, como muestran McCaffrey *et al.* (2009), el seguimiento de los profesores durante varios cursos permite minimizar los errores de medición.

Otra crítica habitual al método del valor añadido es el posible sesgo causado por la omisión de variables. Como se ha explicado en en el apartado "La importancia de la calidad docente en el proceso educativo", la oportunidad de elegir escuela o la posible formación de grupos basados en el resultado académico puede suponer un problema a la hora de llevar a cabo una evaluación basada en el valor añadido. Por esta razón, la metodología del valor añadido parece especialmente adecuada para evaluar al profesorado a nivel de escuela. Aun así, existe la posibilidad que la formación de grupos no sea aleatoria dentro de un mismo centro. Como se ha ilustrado en los ejemplos obtenidos a partir de los datos de la Universidad Carlos III de Madrid, la formación aleatoria de grupos permite la evaluación de la labor docente mediante el valor añadido. A la hora de llevar a cabo estudios similares a nivel de primaria o secundaria será necesario saber qué centros llevan a cabo una formación aleatoria de sus grupos. Santín y Sicília analizan centros con asignación aleatoria de alumnos en el capítulo que coautorizan en este mismo proyecto de investigación.

El principal requerimiento a la hora de llevar a cabo una evaluación del profesorado basada en el enfoque del valor añadido será la existencia de datos longitudinales que hagan un seguimiento de estudiantes y profesores. En España no resulta habitual disponer de este tipo de datos en los niveles de Educación Primaria y secundaria. Facilitar y promover la disponibilidad de este tipo de datos será una medida clave y de escaso coste económico a la hora de mejorar la gestión educativa. En este sentido, la introducción de pruebas estandarizadas en la última reforma educativa debería ser vista como una oportunidad a partir de la cual empezar a publicar este tipo de datos que permitan facilitar la evaluación de resultados dentro del sistema educativo español. Un paso más allá, aunque ciertamente tendría un mayor coste, consistiría en extender las bases de datos con un seguimiento de la inserción laboral e indicadores de calidad de vida futuros de los estudiantes.

Finalmente, cabe destacar algunas de las limitaciones de estas medidas de evaluación. Obviamente, la efectividad docente sólo puede ser evaluada una vez se ha ejercido la labor educativa. Por lo tanto, medidas como el valor añadido del profesor no serán de utilidad a la hora de establecer qué criterios deben ser tenidos en cuenta para la contratación de nuevos profesores. Cabe recordar también que este tipo de medidas se basan habitualmente en la evaluación relativa de resultados. Por lo tanto, no permitirían diferenciar aquellas situaciones donde todos los profesores evaluados sean igualmente buenos de aquellas en las que todos sean igualmente malos. Por lo tanto, una evaluación del profesorado podría penalizar a los profesores de las escuelas con un personal docente por encima de la media.

CONCLUSIONES

La investigación en Economía de la Educación ofrece dos resultados referentes al profesorado con importantes implicaciones para la gestión educativa. En primer lugar, el papel del profesor resulta clave en el rendimiento de los estudiantes. En segundo lugar, la eficacia docente no se explica por las características observables del profesor como su nivel de estudios o la experiencia más allá del primer año de ejercicio.

A la vista de estos resultados, una buena gestión educativa requerirá identificar correctamente aquellas personas que puedan llevar a cabo la actividad docente de manera más eficaz. La imposibilidad de identificar aquellos profesores con unas mejores aptitudes docentes a partir de sus características observables (*inputs*) recomienda el uso de métodos de evaluación basados en sus resultados (*outputs*). Como se ha explicado e ilustrado en el presente trabajo, uno de estos métodos consiste en el cálculo del valor añadido del profesor. Este método evalúa la labor docente basada en el rendimiento futuro de los estudiantes. Por lo tanto, convendría sustituir los viejos métodos de evaluación largamente usados en el sistema educativo español, como los sexenios o la asistencia a cursos, por métodos más avanzados, basados en los resultados como los que se han ilustrado en el presente trabajo.

La falta de datos educativos impide llevar a cabo este tipo de evaluación en gran parte del sistema educativo español. La introducción de pruebas estandarizadas de la última reforma educativa debería ser vista como una oportunidad para mejorar la información y la disponibilidad de bases de datos. Esto permitiría conseguir una mayor rendición de cuentas y evaluación en la tarea educativa. Cabe destacar que centralizar y publicar este tipo de datos no supondría un coste elevado y permitiría sentar las bases para una mejora del sistema educativo.

A pesar de su eficacia en la evaluación docente, el método del valor añadido no es aplicable a la hora de decidir la contratación de profesores noveles. Esta limitación unida al escaso poder explicativo de las características observables del profesor, dificulta el establecimiento de unos criterios para la incorporación de nuevo profesorado. Sin embargo, teniendo en cuenta la importancia de la experiencia durante el primer año en la mejora de la efectividad docente, parece lógico intentar establecer un sistema de entrada de nuevos profesores que facilite por un lado la adquisición de experiencia por parte de los nuevos profesores y que al mismo tiempo limite la exposición de los estudiantes a los profesores que todavía se encuentran en su primer año de ejercicio.

Finalmente, aplicar el método del valor añadido a la evaluación del profesorado presenta ciertas limitaciones como las críticas metodológicas y de aplicación. En este sentido, Hanushek y Rivkin (2010b) proponen una mayor autonomía de los centros como una alternativa o complemento a la evaluación basada en el valor añadido del profesorado. Esta propuesta se basa en la evidencia documentada por Jacob and Lefgren (2008) sobre la capacidad de los directores de los centros para identificar cuáles son los profesores con mayores y menores niveles de efectividad docente.

Los malos resultados educativos de España en pruebas internacionales de conocimiento, como el *International Student Assessment Programme*, advierten sobre la necesidad de mejoras urgentes del sistema educativo. La investigación en Economía de la Educación aporta algunas claves para acometer este tipo de mejoras, una de las más claras es el papel del profesor como elemento clave en el proceso educativo. La creación de un sistema de selección, evaluación y provisión de incentivos al profesorado deberá contribuir a la mejora que necesita el sistema educativo español.

REFERENCIAS

Aaronson, D., Barrow, L. y Sander, W. (2007). "Teachers and student achievement in the Chicago public high schools". *Journal of Labor Economics*, 25(1), 95-135.

Bagues, M., Berniell, I. y Cabrales, A. (2012) "The production function of College Education: Evidence from a Natural experiment". *Working Paper*.

Braga, Michela, Marco Paccagnella y Michele Pellizzari. "Evaluating students' evaluations of professors". *Economics of Education Review* 41 (2014): 71-88.

Chingos, M. M. y Peterson, P. E. (2011). "It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness". *Economics of Education Review*, 30(3), 449-465.

Clotfelter, Charles T., Ladd, Helen F. y Vigdor, Jacob L. (2007). "Teacher credentials and student achievement: Longitudinal analysis with student fixed effects".

Economics of Education Review, *Elsevier*, vol. 26(6), pages 673-682, December.

Cohen, D.K. y Murnane, R.J. (1985). "The merits of merit pay". *Public Interest* 80 (Summer), 3-30.

Cohen, D.K. y Murnane, R.J. (1986). "Merit pay and the evaluation problem: Understanding why most merit pay plans fail and a few survive". *Harvard Educational Review* 56 (1), 1-17.

Cohn, E. (1996). "Methods of teacher remuneration: Merit pay and career ladders". In: Becker, W.E., Baumol, W.J. (Eds.), *Assessing Educational Practices: The Contribution of Economics*. MIT Press, *Cambridge*, MA, pp. 209-238.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D. y York, R.L. (1966). *Equality of Educational Opportunity*. U.S. Government Printing Office, Washington, DC.

Darling-Hammond, L., Berry, B. y Thoreson, A. (2001). "Does teacher certification matter? Evaluating the evidence". *Educational evaluation and policy analysis*, 23(1), 57-77.

Goldhaber, D. D. y Brewer, D. J. (2000). "Does teacher certification matter? High school teacher certification status and student achievement". *Educational evaluation and policy analysis*, 22(2), 129-145.

Hanushek, E. (1971). "Teacher characteristics and gains in student achievement: Estimation using micro data". *The American Economic Review*, 280-288.

Hanushek, E. A. (1992). "The trade-off between child quantity and quality". *Journal of political economy*, 84-117.

Hanushek, E. A. (1997). "Assessing the effects of school resources on student performance: An update". *Educational evaluation and policy analysis*, 19(2), 141-164.

Hanushek, E. A. (2003). "The Failure of Input Based Schooling Policies". *The economic journal*, 113(485), F64-F98.

Hanushek, E.A., Kain, J.F., O'Brien, D.M., Rivkin, S.G. (2005). "The market for teacher quality". NBER *Working Paper* 11154 (February).

Hanushek, E. A. y Rivkin, S. G. (2006). "Teacher quality". *Handbook of the Economics of Education*, 2, 1051-1078.

Hanushek, E. A. y Rivkin, S. G. (2010a). "Constrained Job Matching: Does Teacher Job Search Harm Disadvantaged Urban Schools?" NBER *Working Paper* No. 15816.

Hanushek, E. A. y Rivkin, S. G. (2010b). "Generalizations about using value-added measures of teacher quality". *The American Economic Review*, 267-271.

Jacob, Brian A. y Lars Lefgren. 2008. "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education". *Journal of Labor Economics*, 26(1): 101-36.

Jepsen, C. y Rivkin, S.G. (2002). "What is the trade-off between smaller classes and teacher quality?" *National Bureau of Economic Research*.

Kane, T. J., Rockoff, J. E. y Staiger, D. O. (2008). "What does certification tell us about teacher effectiveness? Evidence from New York City". *Economics of Education Review*, 27(6), 615-631.

Kane, T. J. y Staiger, D. O. (2008). "Estimating teacher impacts on student achievement: An experimental evaluation" (No. w14607). *National Bureau of Economic Research*.

McCaffrey, Daniel F., Tim R. Sass, J. R. Lockwood y Kata Mihaly. 2009. "The Intertemporal Variability of Teacher Effect Estimates". *Education Finance and Policy*, 4(4): 572-606.

Murnane, R.J. y Phillips, B.R. (1981). "Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance". *Economics of Education Review* 1 (4), 453-465.

Murnane, Richard J., John B. Willett y Frank Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination". *Review of Economics and Statistics*, 77(2): 251-66.

Nye, B., Konstantopoulos, S. y Hedges, L. V. (2004). "How large are teacher effects?" *Educational evaluation and policy analysis*, 26(3), 237-257.

Rivkin, S.G., Hanushek, E.A. y Kain, J.F. (2005). "Teachers, schools, and academic achievement". *Econometrica* 73 (2), 417-458.

Rockoff, J.E. (2004). "The impact of individual teachers on student achievement: Evidence from panel data". *American Economic Review* 94 (2), 247-252.

Rodríguez, R. y Rubio, G. (2013). "Teaching Quality and Academic Research". *Working Paper* UAM.

Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement". *Quarterly Journal of Economics*, 125(1): 175-214.

Santín, D., Sicilia, G. "El impacto de la educación infantil en los resultados de primaria: evidencia para España a partir de un experimento natural". En *Reflexiones sobre el sistema educativo español*. Fundación Europea Sociedad y Educación. 2015.

Sass, Tim R., Anastasia Semykina y Douglas N. Harris. "Value-added models and the measurement of teacher productivity". *Economics of Education Review*, 38 (2014): 9-23.

Wachtel, H. K. (1998). "Student evaluation of college teaching effectiveness: A brief review". *Assessment & Evaluation in Higher Education*, 23(2), 191-212.

CLAVES PARA MEJORAR LA EDUCACIÓN Y FORMACIÓN DE ADULTOS EN ESPAÑA EN LA POST-CRISIS

FLORENTINO FELGUEROSO
Fundación de Estudios de Economía Aplicada (FEDEA)

RESUMEN

Este capítulo tiene un doble objetivo. En primer lugar, trata de analizar los datos más recientes sobre competencias y aprendizaje de nuestros adultos desde dos perspectivas: por una parte, su evolución a lo largo de la última recesión y, por otra, en comparación con los países europeos que deberían servir de referencia, si deseamos mejorar nuestras tasas de participación en la educación y formación de adultos. Los resultados de este análisis permiten poner en evidencia nuestras principales carencias tanto en la educación y formación reglada como en la no reglada. En segundo lugar, trata de ofrecer una revisión de la evidencia empírica internacional sobre la eficacia de los incentivos públicos a la formación, prestando especial atención a los que se prevé incluir en la actual reforma de la formación para el empleo.

Las principales conclusiones de este estudio son que, en España, se produce una clara infra-inversión en el aprendizaje de las personas con mayor necesidad de mejorar sus competencias para su inserción en el mercado de trabajo. Al igual que en los países de nuestro entorno, convendría desarrollar un programa nacional de adquisición de competencias básicas, al margen de la educación inicial reglada, aunque con un reconocimiento oficial y una línea de financiación propia. Por otra parte, nuestro país difiere también de los demás en la escasa transferencia directa de gasto público en políticas activas de formación hacia los participantes en el aprendizaje permanente. La actual reforma de la formación para el empleo prevé incorporar instrumentos tales como los bonos de formación que cubrirían esta deficiencia. La evidencia internacional muestra que la eficacia de este tipo de incentivos debería ser probada ex-ante por medio de una evaluación experimental y que solo parecen funcionar cuando vienen acompañados de información acerca de la calidad de los centros de formación y de un buen asesoramiento sobre itinerarios de formación.

INTRODUCCIÓN

Ahora que España está saliendo de la recesión, es el momento de hacer balance con el objetivo de analizar si se ha aprovechado este largo período para mejorar las competencias de nuestra población adulta y reflexionar no solo sobre las deficiencias que se han de corregir sino también sobre el modo de hacerlo.

Este estudio consta de cuatro secciones. En la primera, se presenta una panorámica de nuestras principales deficiencias en la educación y en las competencias básicas de nuestros adultos, en comparación con los demás países de la UE. En la segunda, se analiza la evolución de la participación en el aprendizaje permanente a lo largo de la gran recesión. A continuación, se estudian las características del perfil de los participantes en los sistemas de educación y formación reglados y no reglados en España y en ocho países que deben servir de referencia para alcanzar mayores cuotas de participación en el aprendizaje de adultos. Para finalizar este apartado, se documentan en detalle las principales carencias de nuestro modelo de aprendizaje. En la sección tercera, hacemos una revisión de la evidencia empírica internacional sobre incentivos públicos a la formación, prestando especial atención a los que se prevé incluir en la actual reforma de la formación para el empleo. Por tanto, ponemos el foco en los instrumentos que se suelen utilizar para incentivar el aprendizaje de adultos y la evidencia de la que se dispone sobre su eficacia. Finalmente, en la última sección, se presentan las principales conclusiones y recomendaciones de este capítulo y se realiza una valoración del actual sistema de aprendizaje y de cómo se puede mejorar.

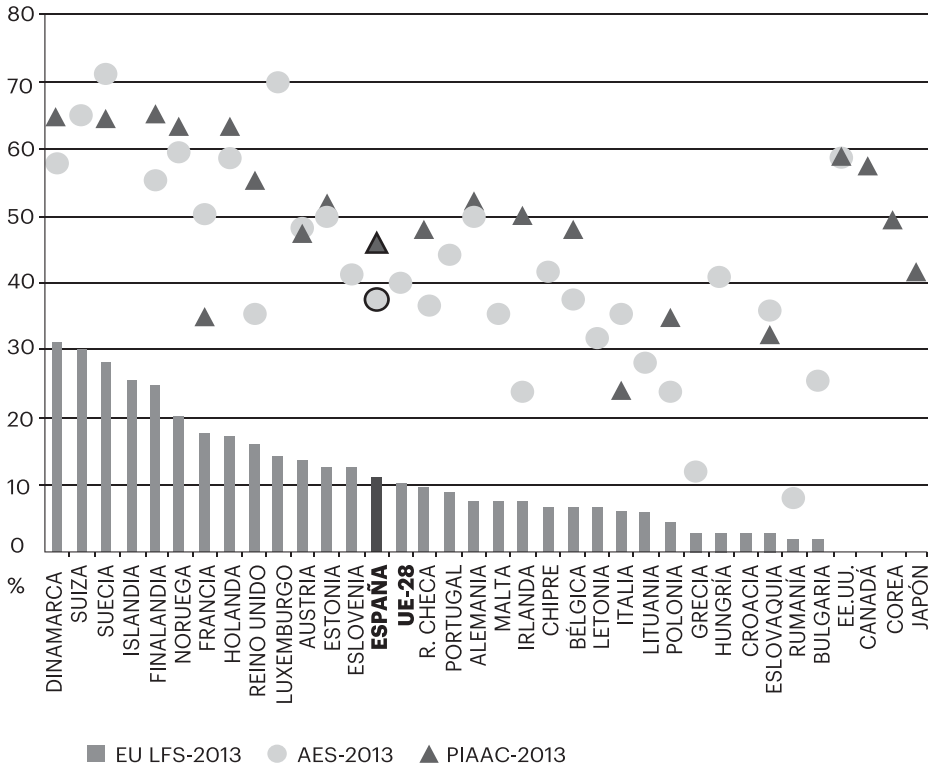
LOS RETOS DEL APRENDIZAJE DE ADULTOS EN ESPAÑA

La Comisión Europea lleva casi dos décadas situando el aprendizaje permanente como un factor clave del crecimiento económico en la Unión Europa. Lo incorporó entre los objetivos de la Agenda de Lisboa para el 2010, y los volvió a situar entre los objetivos estratégicos para el año 2020. En concreto, para esta fecha, todos los países miembros deberían alcanzar una tasa de participación de los adultos en la educación y formación del 15%[1]. En concreto, la tasa de aprendizaje de adultos se obtiene de las respuestas a las preguntas sobre participación en actividades de aprendizaje (formales o informales) en las cuatro semanas previas a la realización de las entrevistas. Los últimos resultados publicados por Eurostat son los de 2013

1. Los adultos se definen aquí como las personas de 25 a 64 años. Las fuentes de datos utilizadas son las Encuestas de Población Activa reunidas en la *European Labour Force Survey*, a las que nos referiremos en el texto por sus siglas EU-LFS.

y se muestran en el Gráfico 1. Como se puede observar, España está aún a unos 4 puntos del objetivo, aunque se sitúa en la media europea. Trataremos de analizar en este capítulo si seremos capaces de alcanzar esta meta del 15% y nos preguntaremos también por la manera de hacerlo.

GRÁFICO 1. POBLACIÓN ADULTA QUE PARTICIPA EN ACTIVIDADES EDUCATIVAS Y FORMATIVAS. (25-64 AÑOS)



Fuentes: EUROSTAT y OCDE.

No obstante, también realizaremos un análisis crítico del objetivo en sí ya que cabe dudar de que tal objetivo informe realmente acerca del aprendizaje permanente de un país. En primer lugar, porque como bien señaló Oosterbeek (2013), centrarse exclusivamente en los cursos realizados en las últimas cuatro semanas discrimina entre los países que hacen cursos más cortos o intensivos, de aquellos que los hacen de forma extensiva. De hecho, otras fuentes de datos utilizan un período más largo (comúnmente 12 meses). En un período tan corto como el fijado en la *European Labour Force Survey* (LFS), las actividades formativas de corta duración están sub-representadas. Por ejemplo, si nos fijamos en las dos fuentes más recientes, la *Adult Edu-*

cation Survey de 2011 (AES) de Eurostat (los círculos del Gráfico 1), o el *Programme for the International Assessment of Adult Competencies* de 2013 (PIAAC), de la OCDE (los triángulos en el Gráfico 1), constatamos que las tasas de aprendizaje de los adultos ascienden en España hasta el 37,7% o al 45,9%, respectivamente. Con ambas fuentes llegaríamos a conclusiones similares: estamos en la media europea y por encima de nosotros se sitúan unos países que bien podríamos tomar como referencia para mejorar nuestro sistema de aprendizaje. Estos países serían principalmente los escandinavos, Suiza y Holanda, pero también Francia y el Reino Unido.

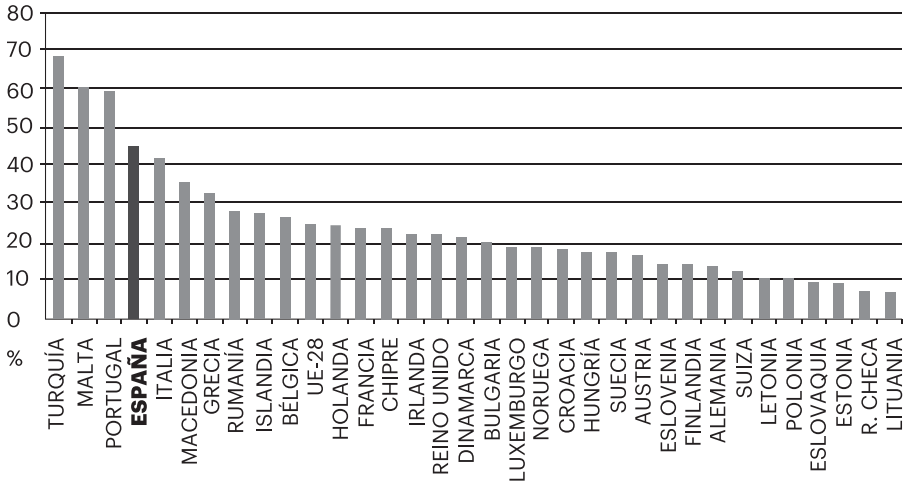
En segundo lugar, sería más conveniente tener una medida del *stock* de formación incorporado a la fuerza de trabajo. Sin embargo, para ello, necesitaríamos datos que nos permitieran contabilizar la inversión en formación acumulada en cada adulto, así como su depreciación. No disponemos de esta información, pero sí tenemos otra que nos indica que la formación además es muy heterogénea y que varía en duración, intensidad, nivel y tipo de aprendizaje. Por este motivo, en este trabajo, analizaremos las diferencias que nos separan de ocho países que tomaremos como referencia para investigar de dónde proceden y a qué podrían deberse dichas diferencias: Dinamarca, Finlandia, Noruega, Suecia, y los mencionados Suiza, Holanda, Francia y Reino Unido. La mayoría de los datos que utilizaremos procederán de la EU-LFS. A pesar de las limitaciones que acabamos de señalar, tiene la ventaja de disponer de muestras de tamaño suficientemente amplio como para estudiar estas diferencias en detalle, así como la evolución en el tiempo.

Quizás el hecho más criticable de tomar la tasa de participación como indicador de referencia es que las necesidades de formación de algunos países, y muy en particular de España, son mucho mayores que las de los países que sí cumplen ampliamente con los objetivos trazados por la Comisión Europea.

Repasemos rápidamente cuáles son estas necesidades. En primer lugar, hemos de tener en cuenta nuestro retraso educativo por las sucesivas olas de abandono escolar temprano. Así, España presenta una población adulta con uno de los porcentajes más alto de personas (43,4%) con nivel educativo bajo (un nivel igual o inferior a ISCED2)[2]. Según la EU-LFS, sólo Portugal, Turquía y Malta nos superarían con porcentajes superiores al 50%. La media de la UE-28 se sitúa en el 24% y países como los escandinavos o incluso los de Europa central y del Este se encuentran por debajo del 20%. Estas diferencias con los demás países europeos se reproducen para todas las edades debido a nuestras altas tasas de abandono escolar temprano.

2. Para una descripción detallada de esta nomenclatura, consultar el apéndice que acompaña esta publicación, *Indicadores básicos sobre el estado del sistema educativo español*. 2015.

GRÁFICO 2. POBLACIÓN ADULTA CON BAJO NIVEL EDUCATIVO (ENSEÑANZA SECUNDARIA OBLIGATORIA O MENOS, ISCED 0-2, 25-64 AÑOS, 2014)



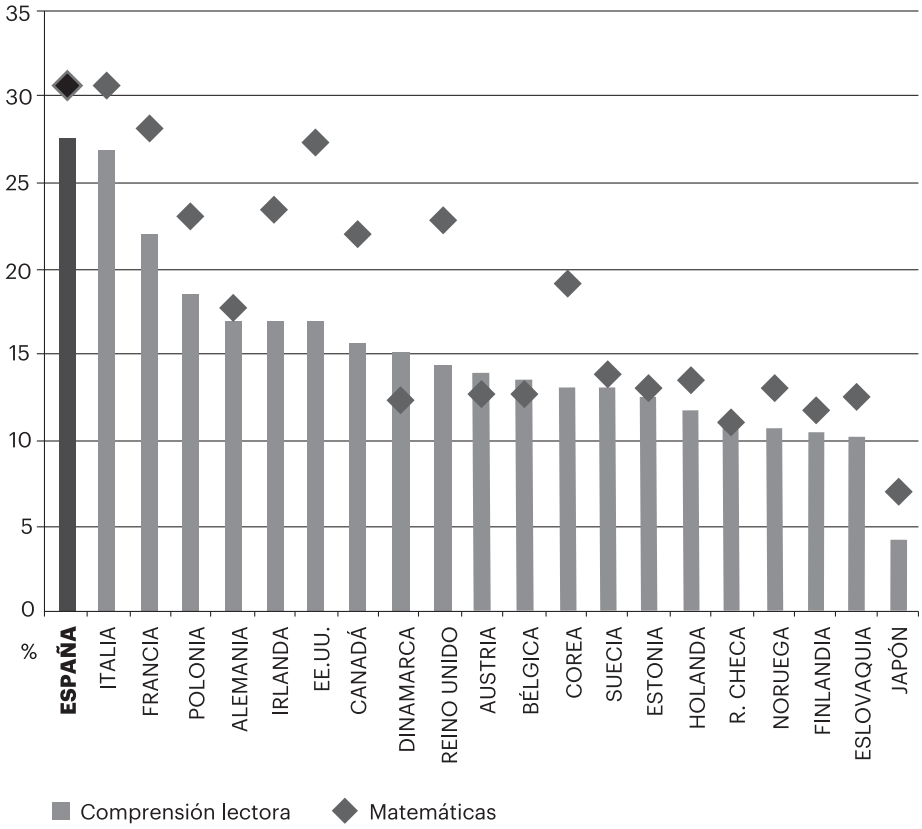
Fuente: *European Labour Force Survey* (EUROSTAT).

Dentro de este grupo de personas con bajo nivel educativo, se encuentra un colectivo de tamaño considerable (un 14,4%), que ni tan siquiera ha finalizado el último curso de la enseñanza obligatoria, habiendo adquirido un título equivalente a la Educación Primaria.

No obstante, estos indicadores de nivel educativo sólo nos dan una visión parcial de nuestras deficiencias formativas. Las cualificaciones adquiridas en el sistema educativo se pueden volver obsoletas con el paso del tiempo, y las personas con un bajo nivel educativo pueden adquirir cualificaciones a lo largo de la vida, como por ejemplo gracias a la experiencia laboral, por lo que, finalmente, el nivel de cualificación real puede ser distinto entre países para el mismo nivel educativo.

El *Programa Internacional para la Evaluación de las Competencias de la Población Adulta* (PIAAC 2013) reveló un panorama muy sombrío acerca de las competencias de la población adulta residente en España en dos de las habilidades básicas, las de comprensión lectora y las de competencia matemática, presentando niveles realmente muy bajos, sólo comparables con los italianos entre los 17 países de la UE: un 27,5% en comprensión lectora y un 30,6% en matemáticas de los adultos españoles entre 16 y 65 años se sitúa en el nivel 1 o menor que 1 en la escala PIAAC.

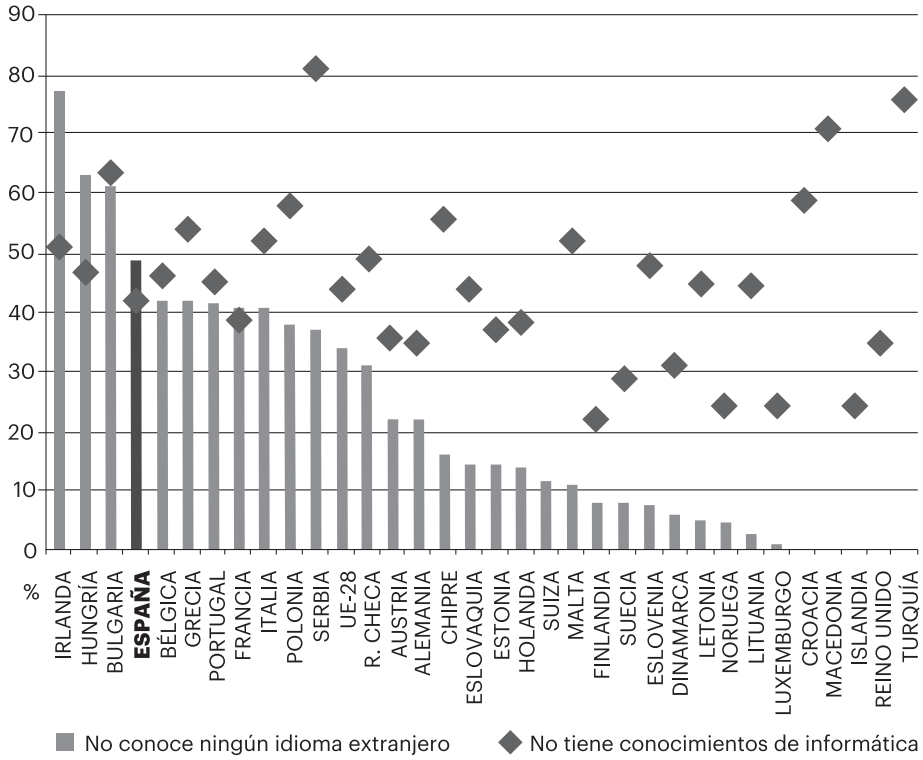
GRÁFICO 3. POBLACIÓN ADULTA CON UN NIVEL DE COMPETENCIAS EN MATEMÁTICAS Y COMPRENSIÓN LECTORA BAJO O MUY BAJO (<=1)



Fuente: *PIAAC* (OCDE).

Además, nuestras competencias en informática y en idiomas son también deficitarias y se revelan como un claro objetivo a atender desde la formación permanente. La AES de 2011 nos mostró que éramos uno de los países con uno de los menores niveles de conocimiento de idiomas de Europa. En concreto, un 49% de la población adulta (25-64 años) no conocía ningún idioma extranjero. Las *Estadísticas de la Sociedad de la Información* de Eurostat indican que, en 2014, un 41% de la población de 25 a 64 años no había utilizado nunca un ordenador o no sabía realizar ninguna de las operaciones TIC básicas, es decir, se encontraba aún cerca del analfabetismo digital.

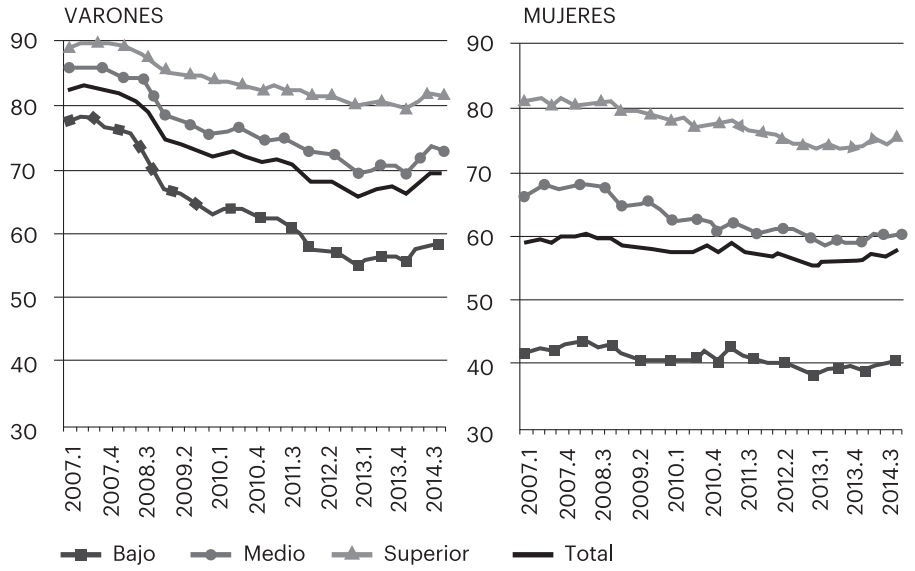
GRÁFICO 4. POBLACIÓN ADULTA QUE NO CONOCE NINGÚN IDIOMA EXTRANJERO (2011) Y POBLACIÓN SIN CONOCIMIENTOS DE INFORMÁTICA (2014)



Fuente: *Adult Education Survey* (2011) *Information Society Statistics* (EUROSTAT).

Durante este período de recesión, se han ampliado las necesidades de formación de la población española, esencialmente las relacionadas con su empleabilidad. Esta etapa ha tenido un impacto muy grande en el empleo de los varones menos educados. Como se puede ver en el Gráfico 5, sus tasas de empleo han caído 20 puntos porcentuales. Las mujeres de bajo nivel educativo, por su lado, probablemente no hayan sufrido un impacto tan grande por la recesión, pero presentan unas tasas de empleo muy bajas (un 40%).

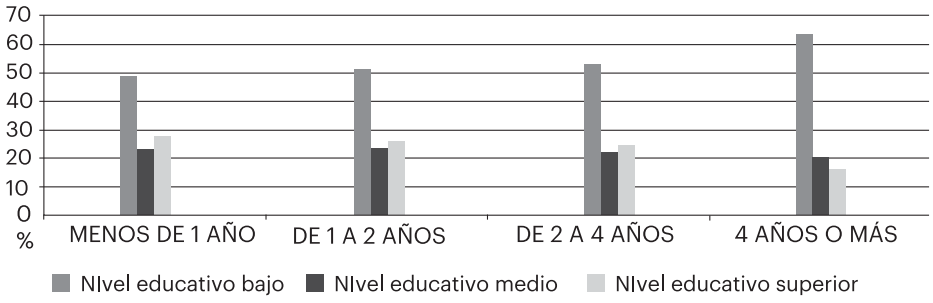
GRÁFICO 5. TASAS DE EMPLEO POR SEXO Y NIVEL EDUCATIVO (25-64 AÑOS), 2007-2014



Fuente: *Encuesta de Población Activa* (EPA, INE).

En realidad, el deterioro laboral de las personas menos educadas no se debe en exclusiva a esta recesión, sino que es fruto de una tendencia que ha marcado estas tres últimas décadas. Sin embargo, la situación actual es verdaderamente crítica. Las personas que no han ido más lejos de la educación obligatoria o tan siquiera la han alcanzado representan más de la mitad de los desempleados en España, y su peso es mayor entre los parados de muy larga duración (superior a dos 2 años). Estas personas muy probablemente requieran de programas formativos específicos para su reinserción laboral.

GRÁFICO 6. DISTRIBUCIÓN DEL PARO POR NIVELES EDUCATIVOS PARA DIFERENTE DURACIÓN DE LA BÚSQUEDA DE EMPLEO. 2014. 4º TRIMESTRE, (25-64 AÑOS)



Fuente: *Encuesta de Población Activa* (EPA, INE).

EL APRENDIZAJE PERMANENTE DURANTE LA RECESIÓN Y LAS DIFERENCIAS CON LOS PAÍSES DE REFERENCIA

En esta sección analizamos las características del aprendizaje de adultos en España desde dos perspectivas: su evolución durante la crisis y las diferencias que nos separan de los países en los que las tasas de participación en el sistema educativo y formativo son superiores a la nuestra. El objetivo es detectar cuáles son las deficiencias y en qué dirección deberían ir las políticas dirigidas a fomentar la participación en el aprendizaje permanente.

PARTICIPACIÓN EN EL APRENDIZAJE DE ADULTOS EN ESTA RECESIÓN: ENSEÑANZA REGLADA *VS* NO REGLADA

La literatura no es concluyente acerca de si la educación y formación de adultos, y la inversión en capital humano en general es contra-cíclica, pro-cíclica o incluso acíclica. Para los trabajadores que han perdido su empleo, la adquisiciónn de capital humano puede ser una vía (o la única vía) para la reinserción laboral. Cuando la recesión viene además acompañada de una reasignación de los recursos entre sectores y una adopción de nuevas tecnologías que demandan más educación genérica, la reinserción laboral puede requerir de una movilidad hacia otros sectores u ocupaciones, movilidad que sólo se produce con una mayor dotación de capital humano genérico. Por parte de las empresas, la formación de sus empleados también puede ser una vía para superar los *shocks* a los que se han visto sometidas durante la recesión, mediante la innovación en productos y procesos de producción.

Los costes de oportunidad de formarse también deberían ser menores tanto para los desempleados como para los empleados y las empresas. Para los desempleados, el tiempo dedicado a la formación supone una menor pérdida de ingresos a medida que aumenta la tasa de paro. Además, con el aumento del paro de larga duración también se incrementa el tiempo disponible para formarse. Para los empleados, la menor disponibilidad de oportunidades fuera de la empresa aumenta los incentivos para adquirir más capital humano específico, y para las empresas, los incentivos para la formación, en términos de horas destinadas a la misma en lugar de a la producción, también son menores. Todos estos factores deberían influir positivamente sobre las tasas de participación en los distintos sistemas educativos y formativos. Por estos motivos, se esperaría un mayor incremento en las tasas de participación de las personas con menores niveles educativos y mayores necesidades de mejorar sus competencias, y también de las empresas supervivientes que requieran realizar cambios en sus sistemas productivos.

Sin embargo, el impacto de la recesión en los ingresos del hogar obliga a las personas a renunciar a sus planes de formación para dedicarse con más intensidad a la actividad laboral (a la búsqueda de empleo o a trabajar en empleos que no hubieran deseado en una fase expansiva). Por otra parte, también han de hacer frente a mayores restricciones de crédito para formarse, restricciones que pueden ser mayores cuanto menores sean las probabilidades de encontrar empleo, esto es, cuanto mayores sean sus necesidades formativas. Por parte de las empresas, las restricciones financieras también pueden decidir aplazar las inversiones en equipos y en la adopción de nuevas tecnologías, y, en consecuencia, recortar sus inversiones en formación. Finalmente, estas restricciones financieras puede ser mayores o menores en función de las políticas públicas destinadas a fomentar la educación y formación de adultos, y también, de forma indirecta en las políticas de I+D.

En el caso español, la recesión ha sido larga y profunda. Se destruyeron empleos de forma masiva en sectores sobredimensionados (en especial en construcción e industrias y servicios auxiliares), por lo que cabría esperar que se haya producido un aumento sustancial en tasas de participación en programas de formación de los trabajadores desplazados de estos sectores, más aún en un contexto de claro retraso en el *stock* de capital humano de la población adulta. Por otra parte, también se produjeron importantes recortes en los gastos públicos en políticas activas, en especial de formación, y en I+D.

Aunque existen varios estudios para el caso de España (tanto específicos o en un contexto de comparación internacional) sobre los perfiles de los adultos que adquieren educación o formación, existe escasa evidencia empírica sobre el efecto neto de esta recesión[3]. Recientemente, Calero y Escardíbul (2014) compararon las tasas de formación no reglada de los empleados en el año 2007 y 2012 utilizando los datos de la EPA. Encuentran que se habría producido un aumento de las tasas de los hombres y una disminución en el caso de las mujeres, lo que interpretan como que existen diferencias de género en el acceso a la formación, y que estas barreras se habrían incrementado en el caso de las mujeres.

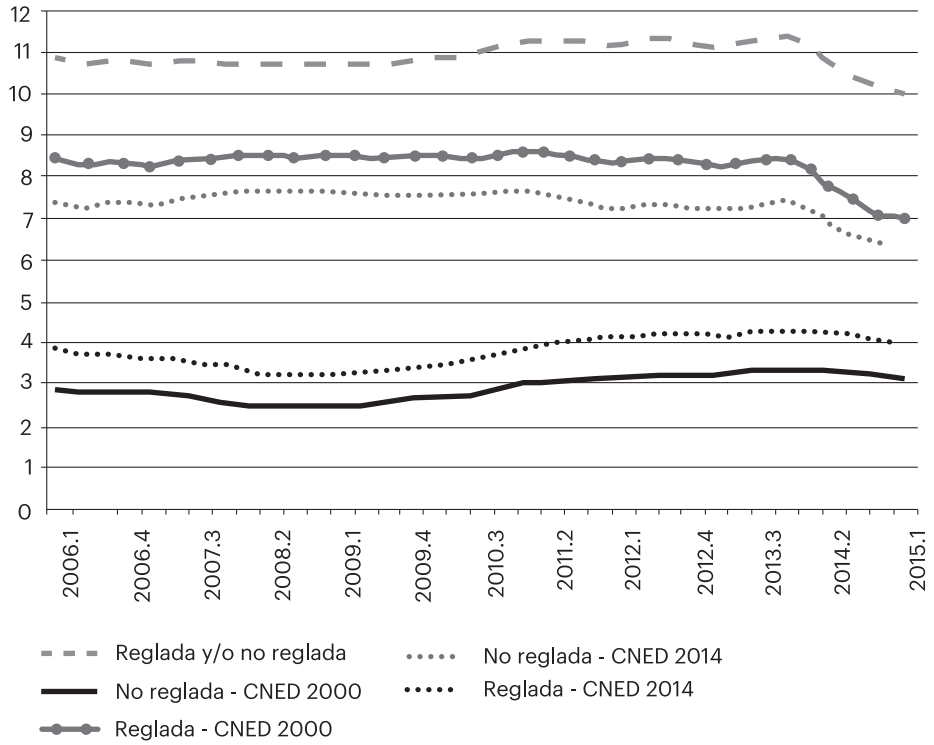
No obstante, los cambios en el aprendizaje de los adultos durante esta recesión es un proceso bastante más complejo, que debería haber afectado tanto a la educación o formación reglada (incluida en los planes oficiales de estudios) como a la formación no reglada (cursos impartidos por academias, en centros de trabajo, dirigidos a desempleados, seminarios, conferencias, clases particulares, etc.), y tan-

3. Por ejemplo, Abellán y Felgueroso (2005), Albert *et al.* (2005 y 2010), Cabrales (2014), Carvajal (2013), Calero y Escardíbul (2014), Caparros et. (2009,), Felgueroso y Jiménez (2009 y 2011), Peraita (2005), por citar sólo algunos realizados por investigadores españoles en los últimos 10 años.

to a los empleados como a los inactivos y los desempleados. De hecho, uno de los resultados más sorprendentes que se obtiene cuando se miran las tasas de participación de los adultos en los dos sistemas es que haya permanecido bastante estable desde el inicio de la recesión (en torno al 11% en media anual). De ahí que, a primera vista, podría parecer que el aprendizaje de los adultos es acíclico. Sin embargo, la evolución ha sido bien distinta según si distinguimos entre estudios reglados o formación no reglada, y por situación laboral.

En el Gráfico 7, se muestran las tasas de participación distinguiendo entre los sistemas reglado y no reglado. Cómo puede observarse, como mucho, la tasa de participación total habría aumentado en unas pocas décimas entre el año 2007 y el año 2013, y éste aumento procedería esencialmente de un ligero crecimiento de la participación en el sistema reglado.

GRÁFICO 7. TASAS DE PARTICIPACIÓN EN LA EDUCACIÓN Y FORMACIÓN REGLADA Y NO REGLADA (25-64 AÑOS, 2006.1-2015.1, MEDIAS ANUALES)



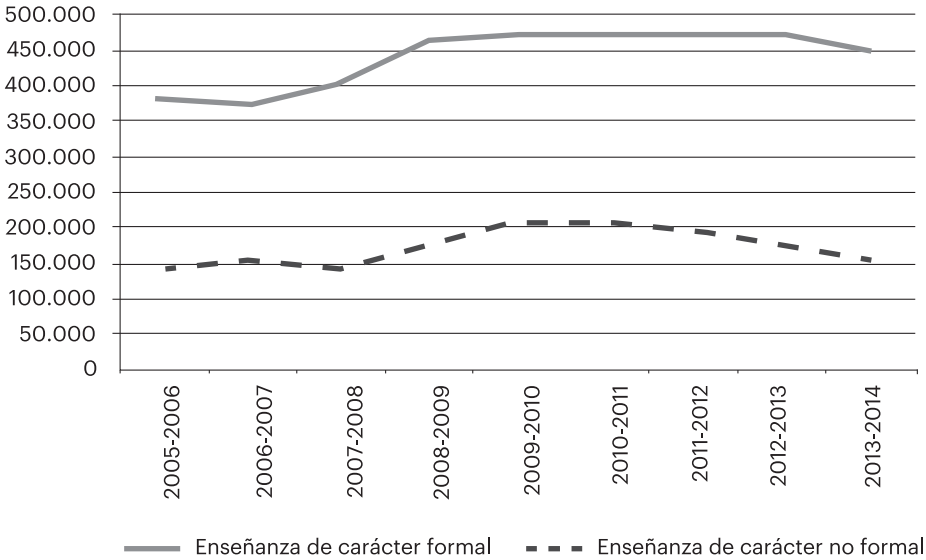
Fuente: EPA, INE.

La caída que se registra desde el año 2014 no tiene una explicación clara. Se ha producido un cambio en la Clasificación Nacional Educativa (CNED-2014), pero no debe-

ría haber influido en las preguntas sobre participación en la educación y formación reglada y no reglada, que permiten calcular la tasa de participación en el sistema educativo o formativo. Además, en la EPA se ha producido un traspaso de algunos ítems de formación no reglada a la reglada que explicaría también la discontinuidad hacia este último sistema. Este traspaso afecta sobre todo a la formación que conduce a la obtención de certificados de profesionalidad (CP) y a la asistencia a las escuelas oficiales de idiomas (EOI), incluidos en la EPA como parte del sistema no reglado antes del 2014. En el Gráfico 7, hemos elaborado una homogeneización, incluyendo ambos, los CP y la participación en las EOI, en el sistema reglado (antes del 2014) y en el sistema no reglado (desde el año 2014), pero seguimos observando una disminución notable (de más de 1 punto porcentual) en la formación no reglada.

La caída de la participación en el aprendizaje permanente no reglado o informal puede también mostrarse con los datos de educación de adultos del Ministerio de Educación, Cultura y Deporte. Según estos datos, los alumnos matriculados en educación de adultos informal se habrían reducido en un 25% desde el curso 2010-2011, después del aumento significativo que se produjo al inicio de la crisis. Este mismo aumento se produjo en la enseñanza formal, que también cesó de aumentar en ese mismo curso, cayendo incluso ligeramente. Estas caídas, además, no pueden explicarse en exclusiva por la disminución de alumnos extranjeros, en especial en la de la formación no reglada, dado que la diminución de alumnos extranjeros ha sido sustancialmente menor que la caída total de matriculados.

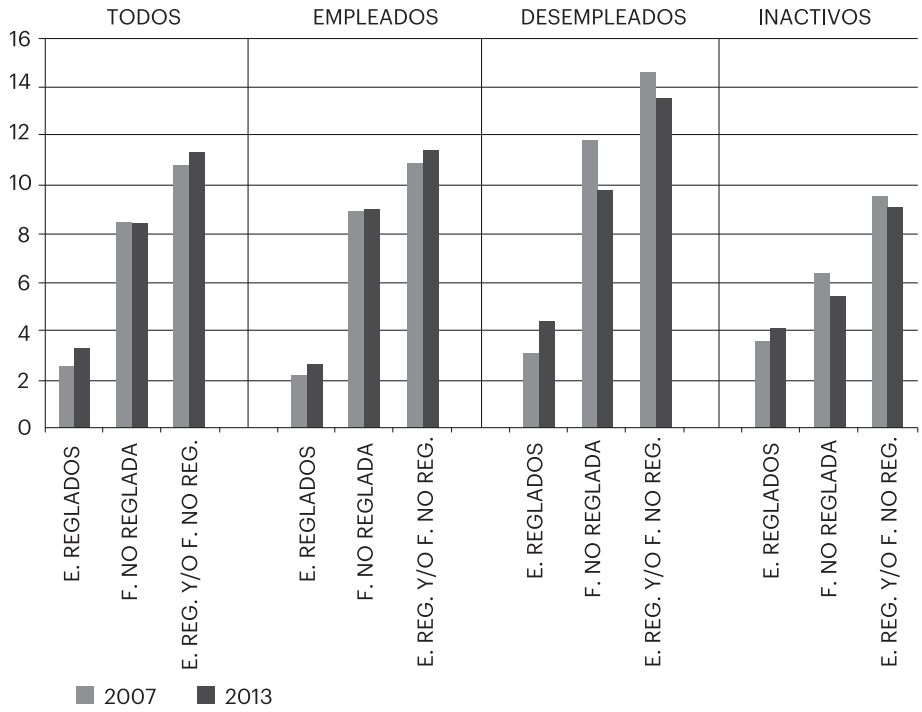
GRÁFICO 8. ALUMNOS MATRICULADOS EN LA EDUCACIÓN DE ADULTOS



Fuente: Estadísticas del MECD.

Además, cuando distinguimos por situación laboral, observamos una caída sustancial de la formación no reglada entre los desempleados, y algo menor entre los inactivos. Aun así, al tener los desempleados mayor tasa de participación en la formación que los demás y, debido al aumento tan dramático de la tasa de paro, se ha mantenido la tasa de participación total sin distinción de la situación laboral. En cuanto a la formación reglada, observamos que el pequeño aumento entre el inicio y el final de la recesión se ha producido en todas las situaciones laborales y, en consecuencia, refleja un ligero aumento en la tasa de participación total.

GRÁFICO 9. TASAS DE PARTICIPACIÓN EN ESTUDIOS O FORMACIÓN REGLADOS Y NO REGLADOS SEGÚN SITUACIÓN LABORAL (MEDIAS DE LOS AÑOS 2007 Y 2014, 25-64 AÑOS)



Fuente: EPA, INE.

Estos fenómenos pueden ser fruto de varios hechos. En primer lugar, la caída de la tasa de participación de los desempleados en la formación no reglada puede deberse a los recortes tan sustanciales en la financiación de cursos de formación para el empleo, en especial desde el año 2011. Por otra parte, otra posible explicación es que el aumento del desempleo haya reducido el coste de oportunidad de los estudios reglados, y se haya producido una sustitución de la formación no reglada por los estudios reglados. También puede ocurrir que las personas que hayan engro-

sado las listas del paro durante la recesión fueran las menos propensas a adquirir formación no reglada.

ESTIMACIÓN DE LOS DETERMINANTES DEL APRENDIZAJE DE ADULTOS EN LOS SISTEMAS REGLADO Y NO REGLADO EN ESPAÑA, Y EN LOS PAÍSES DE REFERENCIA

A continuación, nos centraremos en la comparación de las tasas de formación de adultos en los sistemas reglado y no reglado en ocho países de referencia. Excepto en el Reino Unido, en el resto del grupo de países de referencia se ha producido un aumento de las tasas de participación en educación y formación de los adultos en el período 2007-2013 y, en todos estos países, el aumento se ha producido esencialmente en la formación no reglada. Es, precisamente, en ese tipo de formación donde se producen las mayores diferencias con España, aunque con los países escandinavos y Holanda las diferencias también son notables en la formación reglada.

CUADRO 1. TASAS DE PARTICIPACIÓN DE LOS ADULTOS EN EL APRENDIZAJE PERMANENTE. EDUCACIÓN Y FORMACIÓN REGLADAS *VS* NO REGLADAS (25-64 AÑOS 2007 Y 2013)

	REGLADA Y/O NO REGLADA		REGLADA		NO REGLADA	
	2007	**2013**	**2007**	**2013**	**2007**	**2013**
España	10.6	11.1	2.4	3.0	8.5	8.4
Dinamarca	29.0	31.4	5.7	6.5	24.8	26.6
Suiza	26.8	30.4	5.1	4.7	23.6	27.3
Finlandia	23.4	24.9	8.2	8.9	16.8	17.7
Reino Unido	20.0	16.1	5.5	5.1	18.0	13.9
Suecia	8.6	28.1	5.9	7.1	13.6	23.2
Noruega	18.0	20.4	5.1	6.2	13.6	15.2
Holanda	16.6	17.4	6.8	6.5	9.8	10.9
Francia	6.1	17.7	0.7	1.5	5.5	16.5

Fuente: EU-LFS muestras submuestras anuales (EUROSTAT).

Más allá de las diferencias en tasa de participación en los sistemas de aprendizaje reglados o no reglados con estos países de referencia, podemos preguntarnos si

existen además diferencias en el perfil de los no participantes y en el de los participantes en cada uno de estos sistemas. Para comprobarlo, realizamos un ejercicio econométrico consistente en la estimación de los determinantes en cada sistema formativo estimando un modelo *logit multinomial* con tres situaciones alternativas: no adquirir formación permanente, adquirir formación reglada y adquirir formación no reglada (en exclusiva, es decir, no adquirir a la vez formación reglada, en cuyo caso estaría encuadrado en la situación anterior).

Las variables independientes como determinantes de la participación son las tradicionales en la literatura: las características individuales (la edad, el nivel educativo, el estado civil, la nacionalidad y la situación laboral, distinguiendo además entre distintas antigüedades en las empresas en el caso de los asalariados y la duración del desempleo en el caso de los desempleado). Además también se han incorporado variables que intentan recoger el ciclo económico, como las tasas de paro agregadas de la región de residencia y la tasa de paro nacional del nivel educativo correspondiente; se consideran también efectos fijos para cada año de la muestra. Las estimaciones se han realizado para España y para los 8 países de referencia con una tasa de participación superior a la nuestra. Para hacer las estimaciones se han utilizado los microdatos de la EU-LFS en sus submuestras anuales para el período 2007-2013. Los resultados que se muestran en estos cuadros son *los ratios de riesgo relativos* (RRR), siendo la situación de referencia la de no adquirir formación permanente. Los resultados obtenidos en este ejercicio se muestran en los Cuadro A-1 y A-2.

Podemos observar que la relación entre la edad y la participación en el sistema reglado no es muy distinta de la de los demás países. Existe una relación decreciente (a mayor edad, menor participación) y los ratios de riesgo relativos no son muy distintos a los demás, con la excepción quizás de Francia (con menores RRR) y el Reino Unido (con mayores RRR). Se constatan mayores diferencias en la participación en el sistema no reglado. España sigue manteniendo la relación negativa con la edad aunque no sea tan pronunciada como en el caso del sistema reglado. Por ejemplo, entre las personas de 30 a 49 años, la disminución en la participación no es tan clara, ni lo son tampoco las diferencias entre los más jóvenes y los de más edad, con unos RRR entre 0,47 y 0,51 por 0,03-0,05 en el caso de la reglada. Aun así, el comportamiento decreciente con la edad no es tan claro en los demás países. Por ejemplo, en Suiza los RRR son muy cercanos a 1 (mostrando que no hay casi diferencia por edades). Algo parecido ocurre en el Reino Unido, y en Dinamarca, Finlandia y Noruega los RRR se mantienen bastante estables. Están por debajo de 1, mostrando una menor participación que los de 25 a 29 años, pero luego no existen casi diferencias por encima de esta edad.

CUADRO A-1. RESULTADOS DE LAS ESTIMACIONES DEL LOGIT MULTINOMIAL. EDUCACIÓN O FORMACIÓN REGLADA. (RATIOS DE RIESGO RELATIVOS)

	ESPAÑA		DINAMARCA		FINLANDIA		NORUEGA		SUECIA	
25-29 años										
30-34 años	0.36	***	0.30	***	0.48	***	0.41	***	0.46	***
35-39 años	0.22	***	0.18	***	0.35	***	0.32	***	0.34	***
40-44 años	0.20	***	0.13	***	0.23	***	0.25	***	0.27	***
45-49 años	0.15	***	0.11	***	0.17	***	0.19	***	0.19	***
50-54 años	0.09	***	0.07	***	0.11	***	0.12	***	0.12	***
55-59 años	0.05	***	0.03	***	0.05	***	0.04	***	0.06	***
60-64 años	0.03	***	0.00	***	0.01	***	0.00		0.02	***
Nivel educativo bajo										
Nivel educativo medio	6.36	***	1.18	***	2.67	***	1.80	***	1.52	***
Nivel educativo alto	10.12	***	1.50	***	3.04	***	3.43	***	3.69	***
Mujer	1.00		1.44	***	1.35	***	1.40	***	1.74	***
Soltero/a										
Casado/a	0.43	***	0.77	***	0.67	***	0.91	***	1.07	***
Separado/a	0.76	***	0.87	***	1.02		1.22	***	1.11	***
Nacionales										
UE-15	0.56	***	1.34	***	1.18		0.81	**	1.02	
Europeo no UE-15	0.36	***	0.99		1.14		0.90		1.52	***
No Europeo	0.56	**	0.95		1.00	***	0.76		1.00	***
Asalariado/a antig. 0-1 año										
Asalariado/a antig. 2-5 años	0.82	***	0.87	***	3.56	***	0.97		0.65	***
Asalariado/a antig. 6 o +años	0.92	**	0.63	***	3.05	***	0.93	*	0.73	***
Autónoma/a	0.59	***	0.33	***	1.28	***	0.69	***	0.54	***
Otro tipo de empleo	1.50	***	0.83		1.55		2.31	***	1.41	***
Parado/a 0-6 meses	1.50	***	0.93		3.23	***	2.09	***	2.76	***
Parado/a 6 m-1 año	1.54	***	0.57	***	1.16		0.98		2.70	***
Parado/a 1 año o más	1.57	***	0.72	***	1.37	**	2.08	***	4.00	***
Inactivo/a	3.67	***	3.96	***	3.81	***	4.11	***	4.95	***
Tasa de paro regional	1.01	***	1.30	***	0.97	*	1.05		0.97	***
T. Paro reg. Nivel educativo	1.01	**	0.91	***	1.00		1.01		1.01	***
Año 2007										
Año 2008	0.91	***	1.10	***	0.85	***	1.08		1.03	*
Año 2009	0.87	***	0.76	***	0.81	***	1.06		1.06	***
Año 2010	0.91	***	0.78	***	0.85	***	0.94		1.19	***
Año 2011	1.00		1.12	***	0.87	***	0.96		1.18	***
Año 2012	0.87	***	1.06	*	0.87	***	1.15	***	1.27	***
Año 2013	0.94		1.25	***	0.88	***	1.16	***	1.34	***
Constante	0.02	***	0.15	***	0.05	***	0.07	***	0.09	***
N	481896		477272		237163		114648		1397226	
Pseudo-R2	0.1248		0.1116		0.1484		0.0976		0.1027	

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO A-1. RESULTADOS DE LAS ESTIMACIONES DEL LOGIT MULTINOMIAL. EDUCACIÓN O FORMACIÓN REGLADA. (RATIOS DE RIESGO RELATIVOS)

	SUIZA		FRANCIA		HOLANDA		UK	
25-29 años								
30-34 años	0.49	***	0.20	***	0.48	***	0.75	***
35-39 años	0.33	***	0.06	***	0.37	***	0.67	***
40-44 años	0.27	***	0.06	***	0.35	***	0.61	***
45-49 años	0.25	***	0.04	***	0.31	***	0.50	***
50-54 años	0.15	***	0.02	***	0.24	***	0.38	***
55-59 años	0.07	***	0.01	***	0.15	***	0.22	***
60-64 años	0.03	***	0.00	***	0.10	***	0.12	***
Nivel educativo bajo								
Nivel educativo medio	2.89	***	6.73	***	1.79	***	2.27	***
Nivel educativo alto	6.75	***	23.88	***	2.76	***	4.12	***
Mujer	1.10	***	0.79	***	1.13	***	1.56	***
Soltero/a								
Casado/a	0.50	***	0.39	***	0.76	***	0.83	***
Separado/a	0.80	***	0.83	***	1.04		1.09	***
Nacionales								
UE-15	0.67	***	1.42	***	1.09	*	1.35	***
Europeo no UE-15	0.55	***	1.55	***	1.86	***	1.14	***
No Europeo	1.17		1.77	***	1.08		1.27	*
Asalariado/a antig. 0-1 año								
Asalariado/ a antig. 2-5 años	0.94	**	0.61	***	0.91	***	0.89	***
Asalariado/a antig. 6 o +años	0.65	***	0.47	***	0.76	***	0.70	***
Autónoma/a	0.55	***	0.42	***	0.59	***	0.61	***
Otro tipo de empleo	0.60	***	0.08	***	0.57	***	0.90	
Parado/a 0-6 meses	0.96		0.65	***	0.93		1.02	
Parado/a 6 m-1 año	1.10		0.43	***	0.81	**	1.18	**
Parado/a 1 año o más	0.88		0.56	***	0.94		1.40	***
Inactivo/a	1.15	***	6.91	***	1.15	***	1.21	***
Tasa de paro regional	0.88	***	0.92	***	0.86	***	1.04	***
T. Paro reg. Nivel educativo	1.04	***	1.06	***	1.01		1.01	
Año 2007								
Año 2008	0.98		0.97		0.92	***	1.00	
Año 2009	0.93	**	0.93	**	1.09	***	0.82	***
Año 2010	0.94	*	0.85	***	1.21	***	0.73	***
Año 2011	0.84	***	0.94	*	1.33	***	0.62	***
Año 2012	0.83	***	0.96		1.43	***	0.63	***
Año 2013	0.88	***	3.27	***	1.65	***	0.65	***
Constante	0.14	***	0.01	***	0.23	***	0.05	***
N	314045		1775873		446116		355085	
Pseudo-R2	0.0921		0.127		0.0569		0.0723	

CUADRO A-2. RESULTADOS DE LAS ESTIMACIONES DEL LOGIT MULTINOMIAL. EDUCACIÓN O FORMACIÓN NO REGLADA. (RATIOS DE RIESGO RELATIVOS)

	ESPAÑA		DINAMARCA		FINLANDIA		NORUEGA		SUECIA	
25-29 años										
30-34 años	0.77	***	0.79	***	0.86	***	1.01		0.83	***
35-39 años	0.74	***	0.74	***	0.86	***	1.08	**	0.81	***
40-44 años	0.72	***	0.76	***	0.80	***	1.06		0.85	***
45-49 años	0.71	***	0.75	***	0.76	***	1.06		0.86	***
50-54 años	0.63	***	0.77	***	0.81	***	1.04		0.84	***
55-59 años	0.51	***	0.69	***	0.71	***	0.91	**	0.77	***
60-64 años	0.47	***	0.70	***	0.77	***	0.68	***	0.68	***
Nivel educativo bajo										
Nivel educativo medio	2.04	***	2.99	***	1.57	***	1.64	***	1.44	***
Nivel educativo alto	3.98	***	5.78	***	2.90	***	2.94	***	2.32	***
Mujer	1.23	***	1.84	***	1.41	***	1.06	***	2.00	***
Soltero/a										
Casado/a	0.80	***	1.07	***	0.96	**	1.03		1.02	***
Separado/a	0.92	***	1.05	***	1.15	***	1.06	*	0.97	***
Nacionales										
UE-15	0.66	***	0.97		1.02		1.06		0.81	***
Europeo no UE-15	0.49	***	0.91	***	1.09		0.70	***	0.64	***
No Europeo	0.67	***	1.64	***	1.00	***	0.79		1.00	***
Asalariado/a antig. 0-1 año										
Asalariado/a antig. 2-5 años	0.95	**	1.12	***	7.80	***	0.95		1.01	*
Asalariado/a antig. 6 o + años	1.10	***	1.21	***	9.12		1.00		1.09	***
Autónoma/a	0.66	***	0.99		2.87	***	0.66		0.80	***
Otro tipo de empleo	0.79	***	0.73	***	1.97	***	0.22		0.70	***
Parado/a 0-6 meses	1.33	***	1.04	*	2.36	***	0.63		1.17	***
Parado/a 6 m-1 año	1.67	***	1.43	***	2.23	***	0.52		2.05	***
Parado/a 1 año o más	1.51	***	1.35	***	2.35		0.79		4.20	***
Inactivo/a	0.94	***	0.80	***	1.61	***	0.18		0.61	***
Tasa de paro regional	1.00	***	0.76	***	0.95	***	0.99		0.98	***
T. Paro reg. Nivel educativo	0.99	***	1.30	***	1.01		1.01		1.00	
Año 2007										
Año 2008	1.00		1.11	***	0.80	***	0.97		1.33	***
Año 2009	0.97	*	1.26	***	0.80	***	0.89	***	1.34	***
Año 2010	0.97		1.13	***	0.82	***	0.93	**	1.53	***
Año 2011	0.92	***	1.08	***	0.87	***	0.91	***	1.57	***
Año 2012	0.87	***	1.01		0.90	***	1.03		1.73	***
Año 2013	0.92	***	0.97		0.94	**	1.05	*	1.85	***
Constante	0.08	***	0.09	***	0.02	***	0.10	***	0.09	***
N	481896		477272		237163		114648		1397226	
Pseudo-R2	0.1247		0.1116		0.1484		0.0976		0.1027	

[CONTINÚA EN PÁGINA SIGUIENTE]

CUADRO A-2. RESULTADOS DE LAS ESTIMACIONES DEL LOGIT MULTINOMIAL. EDUCACIÓN O FORMACIÓN NO REGLADA. (RATIOS DE RIESGO RELATIVOS)

	SUIZA		FRANCIA		HOLANDA		UK	
25-29 años								
30-34 años	0.99		0.87	***	0.87	***	0.93	***
35-39 años	1.00		0.86	***	0.86	***	0.99	
40-44 años	1.03		0.83	***	0.81	***	1.03	
45-49 años	1.06	***	0.77	***	0.78	***	1.01	
50-54 años	1.05	*	0.70	***	0.72	***	0.98	
55-59 años	0.96	**	0.51	***	0.61	***	0.90	***
60-64 años	0.83	***	0.42	***	0.44	***	0.83	***
Nivel educativo bajo								
Nivel educativo medio	2.64	***	1.62	***	1.73	***	1.80	***
Nivel educativo alto	5.83	***	3.02	***	2.33	***	3.49	***
Mujer	1.28	***	1.22	***	1.04	***	1.51	***
Soltero/a								
Casado/a	0.89	***	1.01	*	0.95	***	0.95	***
Separado/a	0.93	***	1.10	***	1.16	***	0.99	
Nacionales								
UE-15	0.78	***	0.76	***	0.89	**	1.00	
Europeo no UE-15	0.51	***	0.90	**	1.22	***	0.69	***
No Europeo	0.91		0.81	*	1.09		1.25	
Asalariado/a antig. 0-1 año								
Asalariado/a antig. 2-5 años	1.02		1.03	**	0.99		0.87	***
Asalariado/a antig. 6 o + años	1.05	***	1.28	***	0.99		0.92	***
Autónoma/a	0.88	***	0.85	***	0.68	***	0.58	***
Otro tipo de empleo	0.59	***	0.52	***	0.48	***	0.65	***
Parado/a 0-6 meses	0.85	***	1.03		0.75	***	0.45	***
Parado/a 6 m-1 año	1.02		1.34	***	0.82	***	0.50	***
Parado/a 1 año o más	0.65	***	1.09	***	0.77	***	0.68	***
Inactivo/a	0.45	***	0.96	***	0.43	***	0.30	***
Tasa de paro regional	0.91	***	0.97	***	0.83	***	0.96	***
T. Paro reg. Nivel educativo	1.00		0.99	***	1.02		1.02	***
Año 2007								
Año 2008	1.10	***	0.93	***	0.90	***	1.00	***
Año 2009	0.86	***	0.89	***	1.00		0.94	***
Año 2010	1.39	***	0.78	***	1.24	***	0.90	***
Año 2011	1.30	***	0.86	***	1.41	***	0.64	***
Año 2012	1.24	***	0.88	***	1.55	***	0.65	***
Año 2013	1.25	***	3.45	***	2.39	***	0.63	***
Constante	0.14	***	0.05	***	0.15	***	0.11	***
N	314045		1775873		446116		355085	
Pseudo-R2	0.0921		0.127		0.569		0.0723	

En todos los países también se producen aumentos significativos en las tasas de participación en relación con el nivel educativo pero, en el caso de la reglada, España muestra (con Francia) aumentos mucho más significativos que los demás países.

En cuanto al género, se observa que en España no hay diferencias significativas entre hombres y mujeres en la participación en el sistema reglado, aunque si se producen en el no reglado a favor de las mujeres. Aquí también diferimos de los demás países dado que, a excepción de Francia, las mujeres están más presentes tanto en el reglado como en el no reglado.

Finalmente, con respecto a la nacionalidad, España se comporta de manera muy diferente al resto de países de referencia en cuanto a la participación de los inmigrantes en nuestros sistemas de formación. Tanto en el reglado como en el no reglado se observa en España una menor participación de los extranjeros, especialmente de los europeos no procedentes de la UE-15. En países como Dinamarca, Finlandia, Francia, Suecia o Reino Unido las tasas de participación en la educación y formación regladas son mayores para los extranjeros que para los nativos. En las no regladas, también ocurre algo similar. Ningún país tiene tantas diferencias (negativas) en la participación en el aprendizaje permanente no reglado entre extranjeros y nativos como España, y en varios países la diferencias son positivas, incluso para las personas con nacionalidades no-europeas.

En España, las situaciones laborales como determinantes de la participación en el aprendizaje permanente también resultan diferentes al resto de países. En nuestras estimaciones, la persona de referencia está asalariada con una antigüedad en la empresa inferior a dos años. Observamos que la antigüedad está relacionada negativamente con la participación en la reglada y positivamente con la participación en la no reglada. La relación negativa ocurre en todos los países excepto en Finlandia (donde la relación es positiva) y en Noruega, donde la relación no es estadísticamente significativa.

Coincidimos con el resto de países en la menor participación en el aprendizaje permanente de los autónomos, tanto en el sistema reglado como en el no reglado. En el caso de los inactivos ocurre todo lo contrario, dado que en todos los países estos tienen una mayor participación en la reglada que los asalariados y parados, mientras que en la no reglada ocurre todo lo contrario (excepto en Finlandia).

Finalmente, se constata una mayor participación en el sistema reglado en el caso de los desempleados, con pocas diferencias según la duración del desempleo. En esta circunstancia, la situación es muy variada entre países, siendo Suiza el caso extremo ya que no muestra diferencias entre parados y asalariados con escasa antigüedad. En cuanto a la participación en el sistema no reglado, la tasa de formación

es también significativamente mayor que para las demás situaciones laborales, y parece mayor para los que llevan más de seis meses de duración en el desempleo. De nuevo podemos encontrar una gran variedad de resultados por países, no teniendo ninguno RRR tan altos como España para los desempleados excepto Finlandia y Suecia, donde además la relación entre participación y duración del paro es más intensa aún.

La estimación de las diferencias entre situaciones laborales para cada país no puede ocultar las diferencias tan sustanciales que existen en las tasas medias entre estos países y España: cerca de un 30% de los empleados daneses o suizos participan en actividades de formación no reglada, más de 3 veces la tasa estimada para España y, en Suecia y Finlandia, las tasas superan el 20%, resultando especialmente altas las tasas de participación en la formación reglada de los empleados finlandeses.

En cuanto a los no empleados, en Suecia un 45% de los desempleados sigue algún tipo de estudio o formación, y en Dinamarca un 33,5%, frente al 13,2% de España. El 17% de los desempleados suecos sigue una educación reglada, frente al 4% en España. Pero en ese país, incluso una parte muy notable de los inactivos sigue una formación reglada (un 18,1%), frente al 3,8% de España.

PECULIARIDADES DEL MODELO DE APRENDIZAJE DE ADULTOS ESPAÑOL

En este apartado, centraremos nuestra atención en ilustrar en detalle cuáles son las peculiaridades del modelo español de aprendizaje de adultos desde una perspectiva comparada.

Los programas destinados a la adquisición de competencias básicas

Siguiendo un reciente informe de Eurydice, los programas dirigidos a la adquisición de competencias básicas puede dividirse, al menos, en tres grupos principales:

(1) Programas directamente orientados a la educación inicial o reglada que permitan finalizar el nivel correspondiente a la enseñanza secundaria de nivel inferior (ISCED 2).
(2) Programas no reglados destinados a adquirir competencias básicas incluyendo la comprensión lectora o alfabetismo, las matemáticas y las TIC.
(3) Programas de acceso a niveles educativos superiores y adquisición de competencias básicas en un contexto laboral (formación profesional para el empleo, políticas activas), y otros como la educación popular.

Según este mismo informe, España no dispone de programas orientados a proveer a nuestros adultos de competencias educativas básicas más allá de la vuelta a la escuela para cursar un nivel de enseñanza del sistema educativo reglado. Prácticamente todos los países de Europa occidental si tendrían algún tipo de programa al margen de la educación inicial dirigido a que los adultos puedan adquirir un conjunto de competencias básicas (de comprensión lectora, matemáticas y otros contenidos como informática). Estos programas no están directamente vinculados con la formación inicial o reglada, y tienen una línea de financiación propia. No son programas destinados a completar los estudios más allá de la ESO, ni orientados únicamente a combatir el analfabetismo digital, ni tampoco van encaminados a alfabetizar a los inmigrantes en el idioma nativo. Bien es cierto que, como reconoce el propio informe, algunas Comunidades Autónomas tienen programas de este tipo, pero no existe un plan integral nacional que no sea la vuelta a la enseñanza inicial oficial o reglada, con accesos a sus distintos niveles y pruebas para la obtención de títulos.

Los países con programas del tipo (2) se imparten como formación no reglada, aunque varios países los reconocen en sus sistemas de cualificaciones. En la mayoría de los países estos programas están a cargo de los Ministerios de Educación pero en otros son considerados como parte de las políticas activas de empleo, como es el caso de Francia en el que el 90% de los participantes en el programa de competencias clave son desempleados. En Noruega, el programa de competencias básicas para la vida laboral se orienta a los empleados en cooperación con los empleadores.

En cualquier caso, los programas de las Comunidades Autónomas tampoco tendrían gran alcance. En la última EPA del año 2014, aparecen 4,2 millones de personas adultas de edades comprendidas entre los 25 y 64 años que no consiguieron finalizar los estudios correspondientes a la primera etapa de secundaria (lo que denominamos ESO). De estos, sólo una proporción muy modesta, el 3,5%, realizaba algún tipo de curso, ya sea en el sistema reglado (1%) o no en el no reglado (2,5%). Podríamos pensar que el grueso de este colectivo son personas de edad avanzada, generaciones que tuvieron muy bajas tasas de escolaridad, o incluso procedentes de otros países con sistemas educativos menos avanzados. Sin embargo, esta no sería razón suficiente, y aun así, 850.000 personas nativas menores de 45 años estarían en esta situación. De estas, sólo un 6% (3% en la reglada y otro 3% en la no reglada) estaría adquiriendo algún tipo de formación.

Los datos del MECD indican que sólo unas 68.000 personas se matricularon en los cursos de enseñanzas iniciales para adultos en el curso 2013-2014, un 51,5% menos en el curso 2006-2007. La matriculación en la enseñanza secundaria, sin embargo, ha aumentado significativamente: un 25,6% en la forma presencial y un

86,3% a distancia. Sin embargo, el número de alumnos sigue siendo también relativamente bajo, en torno a 178.000. También han aumentado muy significativamente todas las matriculaciones en las pruebas de acceso a los ciclos de formación profesional y a la universidad.

En Felgueroso (2014), se muestra que se ha producido un aumento muy considerable de la demanda de este tipo de estudios de Formación Profesional entre población de más de 20 años (creció un 131% desde 2007, siendo ya un 42% los estudiantes de este ciclo educativo). Como puede observarse en el Cuadro 2, las tasas de retorno a la educación reglada (interanuales) han crecido en unos 3 puntos desde el inicio de la recesión, pero aún siguen siendo relativamente bajas (algo menos del 11%), especialmente para aquellos que sólo tienen una titulación de ESO o incluso menos. Por ello, la Formación Profesional para el empleo sigue siendo clave para estos colectivos que parecen totalmente desconectados de la educación reglada.

CUADRO 2. RETORNO A LOS ESTUDIOS REGLADOS (% DE PERSONAS QUE NO ESTUDIAN EN EL AÑO T-1 Y QUE REALIZAN ESTUDIOS REGLADOS EN EL AÑO T SOBRE TOTAL PERSONAS QUE NO ESTUDIAN EN AÑO T-1 (16-29 AÑOS)

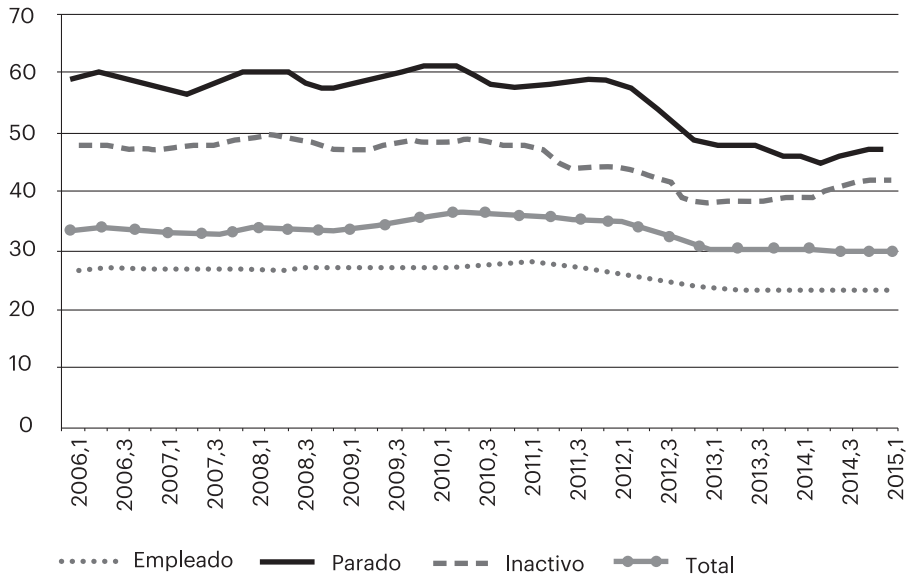
NIVEL DE ESTUDIOS	2006-2007	2008-2009	2010-2011	2012-2013
Menos de ESO	4,8	5,0	6,9	7,6
Título de ESO	3,8	4,6	5,4	5,8
Bachillerato	17,8	18,7	19,0	23,2
FP de grado medio	3,2	4,0	6,5	7,1
FP de grado superior	6,1	7,4	10,9	11,5
Estudios universitarios	12,3	14,6	15,7	15,2
Total	7,9	8,8	10,1	10,9

Fuente: EPA submuestra anual.

Finalmente, la EU-LFS también pone de manifiesto el gran déficit de España en la educación reglada de adultos con menor nivel educativo siendo, como se recordó en la primera sección de este capítulo, uno de los países con más población adulta. En el año 2013, sólo un 0,20% de la población adulta (de 25 a 64 años) realizó estudios reglados iniciales, un 0,4% hizo estudios de nivel medio por casi un 3% de población que realizó estudios superiores (ISCED5 e ISCED6). Las cifras correspondientes a los estudios de la primera etapa de secundaria o de un nivel inferior son menores que en Holanda, Suecia o Reino Unido, por ejemplo, mientras que en estudios medios tenemos unas tasas menores que cualquiera de los 8 países, la mitad que en Finlandia y Holanda, por ejemplo. En estudios de nivel superior

también estamos por debajo de Suecia, Finlandia y Dinamarca. No obstante, lo más sorprendente es que tengamos una tasa tan baja de estudiantes adultos en los menores niveles, sabiendo que cerca de un 15% no ha alcanzado la ESO, y que también sea tan baja en los niveles medios dado que el déficit aquí es aún más considerable (29%).

GRÁFICO 10. HORAS MENSUALES MEDIAS DEDICADAS A LA FORMACIÓN NO REGLADA POR SITUACIÓN LABORAL (ESPAÑA, 25-64 AÑOS, 2006.1-2015.1, MEDIAS ANUALES)



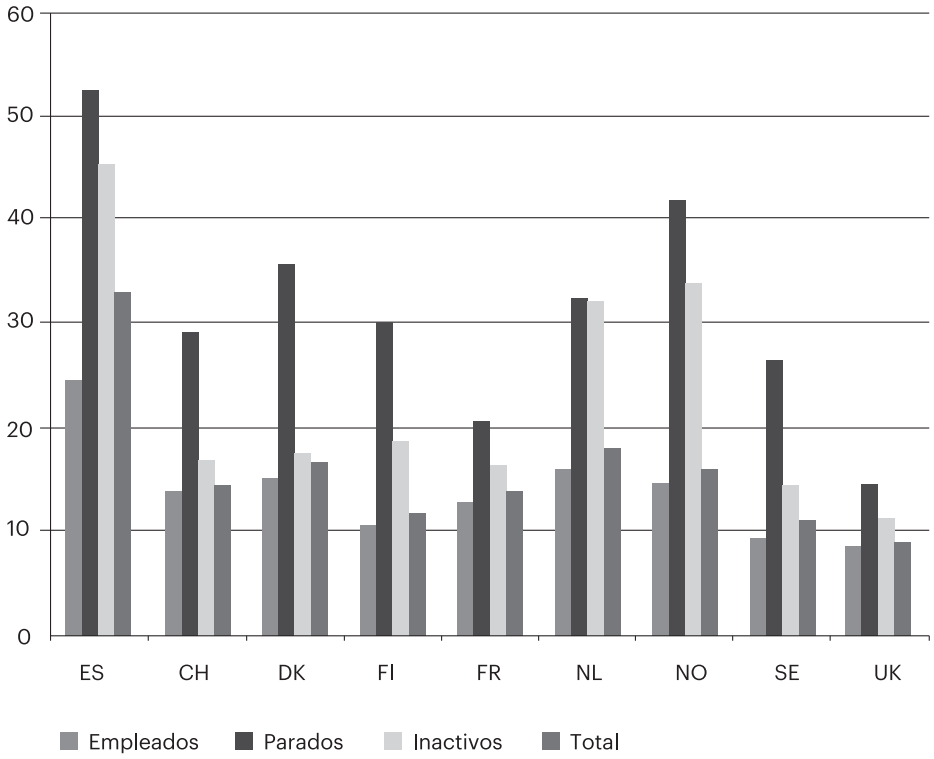
Fuente: EPA (INE).

Intensidad de la participación: decrece pero sigue siendo muy superior al resto de países de referencia

Además, de la tasa de participación hay que considerar la intensidad de la misma. Ésta se puede medir por el número de horas de formación realizadas al mes. En la EPA y en la EU-LFS sólo se da información sobre la enseñanza y formación no reglada. Se observan dos hechos notables. En primer lugar, se evidencia una caída en las horas medias semanales dedicadas a la formación no reglada para todas las ocupaciones desde el año 2011, en especial para los no empleados. El número de horas medias para todos, sin distinguir la situación laboral, había aumentado an-

tes del año 2011 debido al efecto composición de tener más parados, dada la mayor intensidad de formación de los mismos. Las horas semanales medias de los desempleados cayeron en más de 10 horas hasta el año 2014, y también cayeron las de los empleados en más de 5 horas.

GRÁFICO 11. HORAS MEDIAS MENSUALES DEDICADAS A LA FORMACIÓN NO REGLADA POR SITUACIÓN LABORAL (25-64 AÑOS, 2013)



Fuente: EU-LFS submuestra anual (EUROSTAT).

Otro hecho muy destacable en relación con las horas de estudio y formación no reglada es que España se encuentra muy por encima de los otros 8 países, para las tres situaciones laborales de referencia. Los datos del Gráfico 11 proceden de la submuestra anual de la EU-FLS para año 2013, teniendo en cuenta parte de la disminución de horas que se ha producido en la segunda fase de la recesión. Las diferencias son especialmente considerables en los parados y los inactivos.

Políticas activas y formación para el empleo

Tanto la reducción de la tasa de participación no reglada de los desempleados, como la reducción de las horas de trabajo en este tipo de formación seguramente vienen explicadas en parte por los recortes que se han realizado en el gasto en políticas activas y, en especial, en la formación para el empleo. Según los presupuestos del Servicio Público de Empleo Estatal, entre los años 2011 y 2014, la partida destinada a la formación se redujo en unos mil millones de euros, esto es en un 34,5%. Se disminuyó también en un 32,4% la formación de ocupados, en un 31,3% la de parados y un 57% la destinada a las escuelas taller. Las aportaciones del Estado cayeron un 87% y se verificó también una caída de las cuotas de formación (un 10,2%), lo cual está relacionado con la destrucción de empleo neto y la disminución de los salarios medios.

En España, en el año 2013, el 80% de la formación no reglada era formación para el empleo actual o para un empleo futuro. Estos porcentajes aumentan significativamente entre los empleados (84%) y, sobre todo, entre los parados (87%), mientras que cae hasta el 43% entre los inactivos. Las cifras totales son parecidas a las de los suizos, finlandeses y algo inferior a los daneses, y casi el doble de las registradas en Francia y en el Reino Unido (no disponemos de información para Suecia, Noruega y Dinamarca). Sin embargo, también es aquí resultado de un efecto composición: entre los inactivos es mucho más alta en España que en Suiza o Finlandia, y también entre los desempleados. Un factor determinante es la edad de los inactivos españoles, dado que, por encima de los 50 años, la formación para el empleo cae por debajo del 20%.

Resulta difícil realizar una comparación actualizada del gasto público en políticas activas de formación entre España y estos países, dado que los datos disponibles llegan sólo hasta el 2011, es decir, antes de que se iniciara el gran recorte en este tipo gasto. No obstante, ya en el año 2011, podía observarse una peculiaridad del modelo español relacionado con la distribución de las transferencias públicas a los beneficiarios de las políticas activas, a las empresas y a los centros proveedores de los servicios de formación.

Como se puede observar en el Cuadro 3, en el año 2011, España aún tenía un gasto público en políticas activas destinadas a la formación similar a la media europea en su relación con el PIB, incluso por encima de países como Holanda, Suecia o parecido al de Noruega. Sin embargo, las diferencias sustanciales se encontraban, primero, en el gasto por persona activa, segundo, en el porcentaje de participación de estas personas en los cursos de formación y, en tercer lugar, en los destinatarios del gasto transferido. En España, la mayor parte del gasto iba destinada a los proveedores de los servicios de formación y, sin embargo, el porcentaje del gasto

público directamente percibido por los individuos participantes en estas políticas activas era prácticamente nulo.

En Dinamarca, el país con mayor tasa de participación en el aprendizaje de adultos, casi el 100% del gasto es transferido a los participantes directamente y, en la mayoría de los países restantes, una parte muy sustancial del gasto se realiza también de esta manera. Incluso en Suecia, donde la formación reglada tiene más peso que la no reglada, los individuos reciben directamente la mitad del gasto que se destina a esta. Dada la reducción que han experimentado los fondos para la formación, es probable que la asignación a proveedores se encuentre ya en niveles más cercanos a los países de referencia, no obstante una de nuestras principales diferencias con los demás países sigue siendo la no transferencia directa a los individuos para que puedan costearse su formación.

CUADRO 3. GASTO PÚBLICO EN POLÍTICAS ACTIVAS DE FORMACIÓN, PARTICIPACIÓN Y TRANSFERENCIAS (AÑO 2011)

	% GASTO PÚBLICO EN FORMACIÓN SOBRE EL PIB	% GASTO/PIB SEGÚN TIPO DE TRANSFERENCIA			GASTO POR PERSONA QUE DESEA TRABAJAR EN €-PPS	PARTICIPANTES (%)
		INDIVIDUOS	EMPRESAS	PROVEEDORES		
España	0,188	0,005	0,048	0,136	317,55	5,0
Francia	0,353	0,173	0.093	0,086	1.717	14,9
Holanda	0,132	0,041	0,064	0,027	827,17	19,8
Finlandia	0,517	0,369	0,017	0,131	2.031,17	16,1
Suecia	0,085	0,042	0,000	0,043	394,47	2,9
Noruega	0,187	0,148	0,000	0,039	1.972,83	10,9
Dinamarca	0,496	0,485	0,011	0,000	2.117,81	16,7

Fuente: EUROSTAT.

El Cuadro 3 es relevante porque define en gran parte la política de formación para el empleo que se ha seguido en España en las últimas décadas y que probablemente sea transformada profundamente por los cambios recientes introducidos por el actual gobierno, a través del RDL 4/2015 de Reforma Urgente del Sistema de Formación Profesional para el Empleo en el ámbito laboral, y que se iniciaron con la Reforma Laboral del año 2012.

El nuevo modelo de formación para el empleo prevé varios cambios, algunos de los cuales ya están en marcha. En primer lugar, la oferta formativa se realizará en régimen de concurrencia competitiva entre los centros de formación. Este prin-

cipio debería cambiar la asignación de los fondos de formación a los proveedores, abandonando un modelo monopolizado por los agentes sociales. En segundo lugar, se prevé implantar una cuenta de formación que debería acompañar al trabajador a lo largo de su carrera profesional, así como poner en marcha un cheque-formación en el ámbito de la formación para desempleados. Esta última medida debería acercarnos un poco más a los países de referencia por sus efectos en el tipo de transferencias del gasto público en formación. Y, en tercer lugar, se evaluarán permanentemente la calidad y el impacto real de la formación.

INCENTIVOS A LA EDUCACIÓN Y FORMACIÓN DE ADULTOS

En este apartado intentaremos analizar estos dos últimos cambios a la luz de la evidencia disponible.

La intervención pública en los procesos de educación y formación de adultos se justifica por las mismas razones que su intervención en la educación inicial. El gobierno puede considerar el aprendizaje de adultos como un bien socialmente deseable que no se está adquiriendo de forma óptima. Por otra parte, también puede perseguir objetivos distributivos y fijar sus objetivos en que los colectivos más desfavorecidos para que puedan competir mejor en el mercado de trabajo. También se puede estar produciendo una infra-inversión porque el mercado de formación no funcione de forma eficiente, por ejemplo, por deficiencias en la información. Finalmente, en la formación se producen también las limitaciones bien conocidas en la financiación realizada por las empresas, debido a los riesgos de pirateo en la formación genérica o, simplemente, porque las de menor tamaño no pueden asumir los gastos en formación. Por ello, es necesaria la intervención de los gobiernos financiando esencialmente la formación genérica de las personas más necesitadas, por sus déficits en competencias y para que puedan hacer frente a restricciones de liquidez.

Tal cómo hemos ilustrado en la sección anterior, la evidencia empírica muestra que los adultos con menor nivel educativo y menores competencias básicas en general y las personas de edad avanzada son las que tienen menos probabilidad de participar en los procesos de aprendizaje a lo largo de la vida, y sin embargo, también son los que más lo necesitan. Estos hechos ocurren en todos los países, incluso en los más avanzados, sin embargo, en España estos problemas se producen con mayor intensidad.

Los gobiernos pueden intervenir en el mercado de formación de varias formas y pueden jugar varios roles. En un extremo, pueden ofrecer la formación directamente, al igual que las escuelas públicas, o intervenir mediante la financiación de la formación. Oosterbeek (2013) distingue tres tipos de instrumentos: los instru-

mentos fiscales, los fondos de formación sectoriales y los subsidios, incluyendo los bonos y las cuentas de formación o educación. Diferencia, a su vez, dos tipos de instrumentos fiscales: deducciones de los gastos de formación de los pagos tributarios de las empresas, o permitiendo que los trabajadores los deduzcan de sus impuestos sobre la renta.

Como recuerda Oosterbeek, las escasas evaluaciones existentes de este tipo de instrumentos fiscales proceden de Holanda. En concreto, Leuven y Oosterbeek (2004) evaluaron un cambio en el sistema impositivo holandés que permitió que las empresas tuvieran un mayor descuento fiscal para los trabajadores mayores de 40 años, generando una discontinuidad en los costes de formación en esta edad. Los resultados sugirieron que no se produjo el aumento en la participación de la formación de estos colectivos. En cuanto a los incentivos directos mediante deducción en el impuesto sobre la renta, Leuven y Oosterbeek (2007) mostraron que un aumento del 10% en las deducciones por formación incrementaba la tasa de participación en 0,33 puntos porcentuales.

Los fondos de formación sectoriales son bastante habituales en los países europeos, incluyendo España. Se suelen realizar mediante una aportación en forma de cuotas a la Seguridad Social y permiten una mayor integración de las PYMES en los procesos de formación. Sin embargo, no existen evaluaciones relevantes de este tipo de formación.

Del instrumento del que sí disponemos abundante información sobre su impacto y en el que nos detendremos más debido a la previsión de cambios que se pueda llegar a instaurar en España, son los subsidios y, más en concreto, la modalidad de las cuentas o bonos de formación.

Una de las decisiones importantes que deben realizarse en los programas de formación destinados a las personas más desfavorecidas es si los programas han de dictar programas de formación directos y específicos o si es mejor instrumentarlos con bonos que permiten que los beneficiarios elijan por sí mismos el programa de formación.

Los principales argumentos a favor de los bonos son que aumentarían la libertad en la elección de los demandantes y la competencia entre proveedores de formación. La posibilidad de elegir el proveedor de acuerdo con sus preferencias debería llevar a unos mejores emparejamientos entre los demandantes y los proveedores de formación, lo que incrementaría la eficacia de la participación. Además, la libertad de elección puede incentivar la competencia entre los proveedores, dado que deberán hacer frente regularmente a la demanda de los participantes en los procesos de aprendizaje en lugar de un contrato a largo plazo con el Estado.

En el caso de los trabajadores empleados, los bonos permiten que puedan adquirir formación, incluso aquellos que no han sido elegidos por los empleadores para participar en los cursos financiados o gestionados por la empresa, aun siendo subvencionada.

Por otra parte, el argumento de que incrementar las oportunidades de elección siempre aumenta la utilidad es bastante discutible. Puede que el beneficiario del bono no haga la elección más correcta desde una perspectiva social. Una de las razones para que se produzcan malas decisiones son los fallos en la información. De hecho, según Barnow (2009) se pueden producir tres fallos de información potencial. En primer lugar, los participantes pueden tener insuficiente información acerca de la demanda de ocupaciones y salarios con las que son retribuidas. El segundo fallo es la insuficiente información acerca de la eficacia de éxito de los proveedores en la colocación y, por último, puede que no tengan una percepción correcta de sus capacidades para varias ocupaciones y programas de formación. Los dos primeros fallos se pueden resolver proporcionando información a los participantes en el programa, pero el tercero requiere además de orientación específica de los desempleados. En efecto, los orientadores deberían tener un mejor conocimiento de los cursos de formación y la calidad de los proveedores de formación debido a su mayor experiencia previa con los desempleados.

Estados Unidos lleva más de medio siglo de experiencia en el desarrollo de programas de formación basados en bonos. La denominada "G.I. *Bill*" (1944) fue un primer ejemplo de bono de formación por el que el gobierno ofrecía apoyo económico para que los veteranos de guerra accedieran al sistema educativo por un periodo de 45 meses durante los diez años posteriores a la finalización de sus servicios militares al país. Este tipo de prestación podía ser usada ya sea para cubrir los costes directos de la educación o los costes de vida.

Bound y Turner (2002) han estudiado los efectos del G.I. *Bill* sobre el nivel educativo de los veteranos de la Segunda Guerra Mundial. En concreto, trataron de evaluar si el aumento de las matrículas universitarias en la post-guerra se puede atribuir a una democratización de los estudios superiores, o fue debido a los beneficios educativos que acompañaban el servicio militar. Para ello, utilizaron la estructura de sorteo por lotería para disponer de grupos de tratamiento y control (veteranos y no veteranos), así como los sucesivos remplazos que se produjeron a lo largo de la guerra. Encontraron efectos significativamente positivos sobre los años de universidad completados y sobre la probabilidad de obtener una titulación superior. Es decir, el G.I. *Bill* posibilitó que se produjera una transformación educativa y social importante en Estados Unidos, no sólo porque elevó el nivel educativo de la población, sino también porque cambió la distribución de la edu-

cación entre los varones, al permitir el acceso a los estudios superiores a una más amplia capa social.

Desde entonces, se han probado varios tipos de bonos en Estados Unidos. Barnow (2009) hace una revisión de lo que se ha aprendido de esta experiencia. Una de las conclusiones es que el impacto de los programas con bonos de formación, para personas de renta baja, es menor que el de los programas sin bonos. Para los trabajadores desplazados, la evidencia es mixta. Por otra parte, cuando se utilizan los bonos, tanto el asesoramiento como la información sobre los resultados obtenidos por los proveedores pueden mejorar el rendimiento de los bonos.

Las numerosas evaluaciones de los programas con bonos anteriores a esta ley permitieron establecer las condiciones mínimas necesarias para que tengan cierta eficacia. En primer lugar, los estados tuvieron que crear una lista de proveedores de formación elegibles que se ajustara a los requerimientos del estado sobre tasas de colocación y salarios que se consiguen una vez adquirida la formación. En segundo lugar, se requirió que los programas financiaran ocupaciones de alta demanda. En tercer lugar, los programas locales pudieron establecer límites en el coste y en el tiempo para los beneficiarios. Y finalmente, los programas locales pudieron usar a orientadores para guiar a los parados sobre cómo hacer uso de los bonos.

La mayoría de los países que llegaron a implementar un bono de formación realizaron una evaluación *ex-ante* mediante un piloto o aplicación en la "vida real". Este tipo de experimento permite disponer de información sobre la viabilidad de este tipo de bono, así como establecer las reglas que permitan que su aplicación fuera lo más exitosa posible. Una de las grandes limitaciones de las evaluaciones *ex-post* es que resulta muy difícil reconstruir un entorno con los parámetros evaluados en retrospectiva, y porque una aplicación universal del bono de formación no permite luego disponer de un grupo de control que pueda ser utilizado como referencia.

Suiza desarrolló un experimento a gran escala a modo de evaluación *ex-ante* para conocer si merecía la pena instaurar un bono para la formación de adultos. Se hizo entre los años 2006 y 2007. En este país la educación para adultos está muy orientada hacia el libre mercado, en comparación con otros países desarrollados.

Dolores Messer y Stefan Wolfer (2009) explican muy bien cómo se hizo este experimento y sus principales resultados. En el año 2006, unas 2.400 personas, elegidas aleatoriamente, se beneficiaron de un sistema de bonos de formación para que pudieran seguir el curso de su elección en un período de 6 meses. Otras 14.000 personas formaron parte de la muestra de control. Ninguno de los dos grupos supo en ningún momento que formaba parte de un experimento. El grupo de tratamiento se dividió en seis grupos según recibieran o no asesoramiento para la elección del curso y según el valor de bono (200, 750 o 1500 francos suizos). El seguimiento

de los participantes en el experimento se hizo con la EPA Suiza que entrevistó a las personas seleccionadas en su muestra durante 5 años sucesivos.

El problema era intentar que el escenario fuera lo más real posible, aunque la política activa no se aplicase aún en la práctica. Cada uno de los beneficiarios recibió una carta de la Oficina Federal Suiza de Estadísticas con el bono para la formación, diciendo que sería una recompensa por participar en la EPA. La mitad de los beneficiaros recibieron un número de teléfono donde recibirían un asesoramiento gratuito si así lo deseasen.

Messel y Woffers analizaron si los bonos habían tenido un impacto positivo en la adquisición de formación de los beneficiarios. También evaluaron la importancia del peso muerto de los bonos, es decir, el porcentaje del valor del bono que los beneficiarios hubieran pagado en cualquier caso para adquirir la formación (aunque no hubiesen recibido el bono). Para contestar a esta pregunta simplemente compararon los valores esperados de participación en la formación entre los miembros del grupo de control y el de tratamiento.

Para toda la población obtuvieron un aumento significativo de 6 puntos porcentuales en la participación en cursos de formación. El peso muerto estimado resultó francamente elevado: un 60% de los que utilizaron el bono, se habrían formado de cualquier manera. En media, el peso muerto equivaldría a un 50% del valor del bono. Pero, observaron una relación significativamente positiva entre el peso muerto y el nivel educativo. Los menos educados utilizaron menos los bonos, pero más de 2/3 lo hicieron en "bona fide", mientras que los universitarios utilizaron mucho más el bono, pero con un elevado peso muerto en el 91% de los casos.

Por otra parte, la "sensibilidad" de la participación en la formación, atribuida al valor de los bonos resultó significativamente positiva. Finalmente, con respecto a la influencia del asesoramiento para los cursos de formación, los efectos no resultaron significativos, quizás porque no fue activo, sino que los beneficiarios debían solicitarlo por iniciativa propia.

Experimentos similares se han realizado en otros países como Holanda con resultados no muy diferentes. El experimento holandés se realizó en el período 2006-2008, y su tamaño fue menor que en el caso suizo. Unas 600 personas de bajo nivel educativo se beneficiaron de un bono de formación de 1.000 euros que se podían gastar en un período de dos años (en lugar de 6 meses como en Suiza). Hidalgo *et al.* (2014) analizaron el impacto de estos bonos en la participación en la formación y en los resultados en el mercado de trabajo de sus beneficiarios. Por una parte, encontraron que las tasas de participación aumentaron 5 puntos porcentuales (pasando del 37% al 42%) en el primer año y hasta unos 17 puntos porcentuales en el segundo año. El peso muerto estimado, al igual que en el experimento suizo, fue

muy elevado: un 60%, esto es, más de la mitad de la participación podría haberse realizado sin necesidad de incentivos públicos a la formación.

CONCLUSIONES

En este trabajo hemos mostrado que España tiene considerables deficiencias en las competencias básicas de su población adulta, en gran parte derivadas de un retraso educativo profundo al que se han ido sumando sucesivas generaciones de jóvenes que abandonaron la escuela en edad temprana.

A pesar de los avances que se han producido en el aprendizaje de adultos a lo largo de las dos últimas décadas, aún se tienen que producir importantes progresos si se quiere algún día alcanzar a los países que hoy son referencia en la materia. No se trata únicamente de converger en las tasas medias de participación, sino en otros aspectos más preocupantes. En concreto, existe una desigualdad muy profunda entre las personas que adquieren educación y formación en edades adultas y las que no lo consiguen. Existen desigualdades por edades, y sobre todo por niveles educativos, siendo la participación en el aprendizaje permanente de los menos educados realmente testimonial. En este sentido, se deberían tomar las medidas oportunas, fijando objetivos claros en las políticas públicas orientados hacia estos colectivos. Entre ellos, cabría el desarrollo de un programa de adquisición de competencias básicas al margen del sistema educativo, pero con un reconocimiento oficial. La mayoría de los países de Europa occidental disponen de este tipo de programas de ámbito nacional, con la excepción de España. Estos programas cuentan con sus propias líneas de financiación y suelen estar integrados en la formación no reglada, incluso en forma de políticas activas.

Otra deficiencia importante de nuestro modelo de aprendizaje de adultos es el excesivo número de horas mensuales destinadas a la formación no reglada, lo que muy probablemente dificulta la conciliación del aprendizaje de adultos con la vida laboral y familiar, y en consecuencia, permitiría explicar la baja tasa de participación en este tipo de aprendizaje.

En cualquier caso, uno de los hechos más sorprendentes es que no se haya producido un aumento de la tasa de participación a lo largo de esta recesión, a pesar de la intensidad y duración de la misma, y de las necesidades de mejora de las competencias de tantos trabajadores que han sido desplazados. Muy probablemente, los recortes públicos en educación y políticas activas de formación hayan tenido parte de culpa. De hecho, desde el año 2011 se observan mayores caídas en la formación no reglada, que constituye el pilar del aprendizaje de adul-

tos. Esta caída de la financiación pública del aprendizaje permanente ha venido a reforzar las restricciones financieras de parte de la población con más necesidades de formación.

Finalmente, otro aspecto relevante para el caso español es la distribución de la financiación de las políticas activas de formación. Están esencialmente orientadas hacia la transferencia a las empresas y a los proveedores de formación, dejando de lado una opción muy utilizada en los países que nos deberían servir de referencia, de realizar transferencias a los participantes en los cursos de formación. Este es, sin embargo, uno de los cambios previstos en la reforma actual de la formación para el empleo a través de la creación de bonos de formación.

La evidencia disponible en los países de nuestro entorno indica que estos bonos tienen un importante peso muerto, es decir, que acaban subvencionando también a personas que hubieran adquirido la formación en cualquier caso. Por lo tanto, sería importante implementar un experimento previo que indicase quiénes son los colectivos a los que deberían ir dirigidos estos bonos y qué cantidades serían necesarias para elevar la probabilidad de que participen en la educación y formación. Además, la experiencia de otros países también recomienda que se evalúe la eficacia de los proveedores de formación en el impacto que tiene y que esta información esté disponible para los usuarios de los bonos. Finalmente, también recomienda que los bonos vengan acompañados de una inversión en formación y contratación de orientadores que puedan guiar activamente a los usuarios de los bonos en sus decisiones de formación.

Por último, una cosa es que los bonos de formación u otras políticas activas destinadas a incentivar la participación en el aprendizaje permanente acierten o consigan mejorarlo, y otra que la formación adquirida sea eficaz. Solo los países que están en la parte alta del *ranking* en las tasas de participación han realizado evaluaciones consecuentes, aplicando los métodos de evaluación más avanzados, en general basados en experimentos. En este sentido, España está aún muy lejos de fijar las condiciones para que se realicen este tipo evaluaciones. Quizás este sea el momento para empezar a plantearlas. La intención de implementar bonos de formación y los principios de evaluación y transparencia establecidos en la reforma de la formación para el empleo del año 2015 avanzan en esta dirección.

REFERENCIAS BIBLIOGRÁFICAS

Abellán, C. y F. Felgueroso (2005): "Formación no reglada y productividad". *Presupuesto y Gasto Público*, 39, 2/2005, 73-102.

Albert, C., C. García Serrano y V. Hernanz (2005): "Firm-Provided Training and Temporary Contracts". *Spanish Economic Review*, 7 (1), 67-88.

Albert, C., C. García-Serrano y V. Hernanz (2010) "On-the-job training in Europe: Determinants and wage returns". *International Labour Review*, 149 (3), 315–341.

Barnow, B.S: (2009): "Vouchers in U.S. vocational training programs: an overview of what we have learned". ZAF, 32, 71-84.

Bound, J. y Turner, S. (2002). "Going to war and going to college: Did World War II and the G.I. Bill increase educational attainment for returning veterans?" *Journal of Labor Economics*, 20:784–815.

Cabrales, A., J.J. Dolado y R. Mora (2014): "Dual labour market and (lack-of) on-the-job training: PIAAC evidence from Spain and other EU countries". *Estudios sobre la Economía Española*, 2014/14, Fedea.

Calero, J. y J.O. Escardíbul (2014): "Barriers to non-formal, professional training in Spain in periods of econòmic growth and crisis. An analysisi with special attention to the effect of the previous human capital of workers". *Document de Treball de l´IEB*, 2014/12.

Caparrós, A., L. Navarro y M.F. Rueda, Mario (2009) "Análisis de la incidencia y duración de la formación laboral financiada por empresas y trabajadores". *Cuadernos de Economía*, 32 (89), 83-112.

Carvajal, J. (2013): "Choosing to invest in human capital through adult education". Master Thesis CEMFI N0 1304, CEMFI .

Cedefop (2009). *Using tax incentives to promote education and training. Technical report*, Cedefop, Panorama series.

Eurydice (2015): *Adult Education and Training in Europe: Widening Access to Learning Opportunities*. Eurydice Report. Luxembourg: Publications Office of the European Union.

Florentino, F. y S. Jiménez (2009): "The New Growth Model: How and with Whom?" FEDEA, DT 39-2009.

Felgueroso, F. y S. Jiménez (2011): "Balance y retos de la formación professional en España". *Reformas Necesarias para potenciar el crecimiento de la economía Española*, Thomson Reuters/Civitas, Madrid.

Hidalgo, D., H. Oosterbeek y D. Webbink (2014): "The impact of training vouchers on low-skilled workers". *Labour Economics*, 31, 117-128.

Leuven, E. y H. Oosterbee (2004): "Evaluating the effect of tax deductions on training. *Journal of Labor Economics*". 22,461–488.

Leuven, E. y H. Oosterbeek (2007): "The responsiveness of training participation to tax deductibility". Mimeo.

Messer, D. y S. C. Wolter (2009): *Money matters - Evidence from a large-scale randomized field experiment with vouchers for adult training*. CESifo Working Paper 2548.

Oosterbeek, H. (2013): "The finance of adult learning". EENEE Analytical Report Nº 15, European Commission.

Peraita, C. (2005): "Firm-sponsored training in regulated labour markets: evidence from Spain". *Applied Economics*, 37 (16), 1885-1898.

EVALUAR PARA MEJORAR: HACIA EL SEGUIMIENTO Y LA EVALUACIÓN SISTEMÁTICA DE LAS POLÍTICAS EDUCATIVAS

DANIEL SANTÍN
GABRIELA SICILIA
Universidad Complutense de Madrid

RESUMEN

El futuro de cualquier sistema educativo debería estar guiado tanto por una continua aspiración de mejorar sus resultados como por una mayor transparencia, donde se ofrezca más y mejor información sobre las variables educativas relevantes para que los agentes implicados en las políticas educativas puedan tomar mejores decisiones. El seguimiento y la evaluación de las políticas educativas realizados de forma sistemática debe ser la herramienta fundamental que permita conseguir ambos objetivos. Los autores describen cuáles son las claves para poner en marcha un sistema integral de evaluación sistemática de la calidad, la eficacia, la eficiencia y la equidad del sistema educativo no universitario español. Para ello se discute la información que sería necesario recoger y las principales técnicas de seguimiento y evaluación que deben acompañar el proceso educativo. Además de exponer la necesidad de evaluar, los autores también dedican un espacio a presentar las ventajas de las dos metodologías que en los últimos años se han impuesto para realizar la evaluación y el seguimiento de los centros y los programas educativos: la evaluación de impacto y la medición de la eficiencia y la productividad total de los factores. Cualquier programa o intervención educativa que quiera ponerse en marcha y sea susceptible de ser evaluada debería ser diseñada de forma que su evaluación pueda ser implementada mediante estas técnicas. En este sentido, a modo de ejemplo se hace una reflexión sobre cómo debería haber sido evaluado el Programa Escuela 2.0. Finalmente, en el trabajo se explorará además cómo debería articularse institucionalmente la evaluación educativa dentro de un contexto donde las competencias educativas están descentralizadas.

INTRODUCCIÓN

Si acudimos al diccionario de la Real Academia Española evaluar es "estimar, apreciar, calcular el valor de algo". Definido de esta manera, resulta del máximo interés conocer cómo debemos calcular el valor de las políticas públicas que se llevan a cabo en España. Por otro lado, la evaluación o el cálculo del valor de una política o programa público, debe realizarse teniendo en cuenta tanto los recursos invertidos como los resultados alcanzados. Por tanto, la evaluación de un programa educativo está íntimamente ligada con la medición de su eficiencia, que desde un punto de vista Paretiano implica la ausencia de desperdicio de los recursos aplicados para conseguir un objetivo determinado.

La importancia de la eficiencia en la articulación de las políticas públicas resulta un objetivo clave y más aún en el actual contexto de elevado déficit público y persecución de la estabilidad presupuestaria. En este escenario la sociedad simplemente no puede permitirse disponer de recursos ociosos que puedan aumentar su rendimiento en usos alternativos. Tal es la importancia de la eficiencia en la actuación pública que prácticamente todos los países recogen en sus Constituciones el mandato de ejecutar sus presupuestos de forma eficiente. En España, el artículo 31.2 de la Carta Magna afirma que: "El gasto público realizará una asignación equitativa de los recursos públicos y su programación y ejecución responderán a los criterios de eficiencia y economía". Para cumplir con este mandato legal, la evaluación de las distintas políticas públicas es la herramienta de la que disponemos para medir el grado de cumplimiento de este objetivo de eficiencia dentro de un proceso que debería ser sistemático para racionalizar el gasto público y mostrar a los ciudadanos los resultados de las políticas aplicadas.

La evaluación de las políticas públicas en general, pero de la política educativa en particular, es un proceso no exento de problemas debido a distintos factores. En primer lugar, la educación formal reglada es un proceso productivo complejo que consiste en transmitir unos conocimientos en sentido amplio con el objetivo de que el alumno los aprenda y que los mismos le sean relevantes para adquirir nueva formación hasta completar su desarrollo personal y profesional. En segundo lugar, existen distintos agentes implicados en el proceso educativo como alumnos, padres, profesores, directores de centro o la administración pública, cada uno con sus propios objetivos, intereses e incentivos que a veces no están bien definidos y que no siempre coinciden. En tercer lugar, el producto de la educación carece de precios de mercado por lo que el valor de los resultados generados resultan difícilmente cuantificables. En cuarto lugar, los logros educativos generan valor de forma directa para el individuo que la obtiene en forma de capital humano (Becker, 2009),

que se traduce en una mayor productividad que en media generará mejor salario y empleabilidad esperados a lo largo de su ciclo vital. Finalmente, la educación genera además distintas externalidades positivas tanto económicas como sociales. Por un lado, el sistema productivo resulta fortalecido en su conjunto a través de beneficios tales como mayor productividad, ocupación y crecimiento económico. De estos resultados se benefician no sólo los individuos sino también la sociedad, dado que los individuos con más ingresos serán también los futuros pagadores de más impuestos que contribuirán a desarrollar políticas de gasto más generosas que mejoren la cohesión social. Por otro lado, la educación favorece aspectos más sociales como son la participación y consolidación de los valores democráticos o la reducción de la marginalidad, la exclusión y la criminalidad (Lochner y Moretti, 2001) con los consiguientes ahorros en protección social.

Tradicionalmente la evaluación del sistema educativo ha sido parcial y ha estado centrada en garantizar el cumplimiento de los compromisos de gasto y el acceso al sistema. Si bien el control presupuestario sigue resultando útil como mecanismo de control y fiscalización para garantizar que el presupuesto es ejecutado respetando la norma, resulta insuficiente como herramienta de evaluación. En particular debe volver a subrayarse que de una correcta ejecución presupuestaria no puede ser inmediato concluir que los recursos aplicados tengan algún impacto en los objetivos a corto, medio y largo plazo perseguidos, más allá del número de alumnos que han recibido el servicio, o que hayan sido aplicados de forma eficiente. Es por ello que junto al seguimiento de los costes del sistema educativo existe, desde hace un par de décadas, un creciente interés por medir los resultados educativos obtenidos en términos de las competencias que adquieren los alumnos. Para ello se han puesto en marcha proyectos de evaluación internacional del rendimiento a través de pruebas de conocimiento a los que cada vez se adhieren más países. España se ha ido incorporando poco a poco a las evaluaciones de estos proyectos. En 1995 participó en la primera oleada del TIMSS (*Third International Mathematics and Science Study*) promovido por la IEA (*The International Association for the Evaluation of Educational Achievement*) y recientemente participó nuevamente en la oleada de TIMSS-PIRLS 2011 (*Progress in International Reading Literacy Study*). Además, España ha participado en las cinco oleadas que hasta la fecha han tenido lugar de la evaluación internacional que dirige la OCDE conocida como Proyecto PISA (*Programme for International Student Assessment*).

Este proceso de conversión de la evaluación orientada a los recursos en una mayor atención hacia el análisis de los resultados se puede observar en la evolución de la legislación educativa en España. Tanto la Ley General de la Educación (LGE) de 1970 como la Ley Orgánica de Ordenación General del Sistema Educativo (LOG-

SE) de 1990 eran leyes orientadas a garantizar la igualdad de los recursos educativos en todas las escuelas. Así, en aquellos años era más importante regular la escolarización obligatoria, el acceso a la función docente, la dotación mínima de recursos humanos y materiales o la mejora de la preparación del profesorado. En definitiva se trataba de que no existieran escuelas *de primera y de segunda*, sino que todas ellas contaran con unos estándares de recursos adecuados. Posteriormente, tanto la Ley Orgánica de Calidad de la Educación de 2002 como la Ley Orgánica Educativa de 2006 van más allá, al proponer una evaluación orientada a los resultados (o evaluación de diagnóstico). La LOCE, que no llegó a desarrollarse, preveía incluso la realización de pruebas objetivas tanto en primaria como en la secundaria obligatoria y la LOE fijaba las edades de 10 y 14 años como los momentos en los que esta evaluación de diagnóstico era llevada a cabo. Recientemente, la Ley Orgánica para la Mejora de la Ley Educativa (LOMCE) de 2013 prevé cuatro evaluaciones individualizadas (no de diagnóstico) en tercero y sexto de Primaria, en cuarto de la Educación Secundaria Obligatoria y al finalizar el Bachillerato.

Cabe destacar aquí que la realización de pruebas objetivas recoge solo una parte del verdadero producto de la educación recibida por un individuo, que es muy difícil de medir debido a su intangibilidad inherente. La educación no se compone únicamente de la capacidad de repetir información y resolver problemas con más o menos aciertos sino que comprende además otras habilidades no cognitivas como la capacidad de razonar, los buenos modales en la sociedad, la adaptabilidad, el trabajo en equipo y la capacidad de juicio crítico, todos ellos de muy difícil cuantificación aunque sin duda relevantes (Levin, 2012). A pesar de esta naturaleza multi-producto de la educación, la mayoría de las evaluaciones en economía de la educación utilizan los resultados de rendimiento en pruebas objetivas acerca de los conocimientos que el alumno debería tener en una determinada edad por distintas razones. Entre ellas, cabe destacar que son difíciles de falsificar y responden a las actuales exigencias administrativas de calidad para pasar curso e incluso acceder a la universidad. Pero quizás, siguiendo a Hoxby (1999), la más importante sea que tanto políticos como padres de alumnos usan este criterio para evaluar el producto educativo y utilizan la información que se deriva para elegir escuela e incluso lugar de residencia. En este contexto de evaluación del producto educativo, existe un acuerdo prácticamente unánime en todos los países desarrollados en considerar que disciplinas como las matemáticas o la capacidad lectora, son una parte importante del currículo escolar y se consideran competencias vehiculares para la educación de los alumnos.

Teniendo en cuenta estas consideraciones previas, podemos definir la evaluación educativa como el proceso generador de información relevante que permite

relacionar los recursos, las actividades y los resultados a corto, medio y largo plazo de los centros y programas educativos de una administración observando su evolución en el tiempo para utilizar sus resultados para la toma de decisiones que permitan mejorar la eficacia, la eficiencia y la equidad del sistema educativo.

Las ventajas de la evaluación de cualquier política son numerosas. Por ejemplo la australiana *Steering Committee for the Review of Commonwealth/State Service Provision* (2013) señala las siguientes:

- Ayuda a definir y clarificar el cumplimiento de objetivos.
- Hace que el desempeño sea transparente y permite que tanto ciudadanos como políticos valoren si se cumplen los objetivos de las políticas y programas a lo largo del tiempo.
- Anima a la continua mejora en el desempeño.
- Anima a promover los valores de la eficiencia en el uso de los recursos.
- Promueve la coordinación de las políticas entre las distintas administraciones.
- Facilita la comparación de resultados (*yardstick competition*) entre unidades lo que incentiva la mejora del desempeño.
- Identifica a las mejores unidades para aprender de sus prácticas y tener así una referencia (*benchmarking*) de actuación.

El objetivo de este capítulo es introducir cuáles son los aspectos clave y las principales herramientas disponibles de evaluación de la educación no universitaria. No se trata de desgranar el detalle de todas las metodologías de evaluación, pues para ello ya existen múltiples manuales y referencias en economía de la educación, sino de presentar el razonamiento lógico de cómo debería institucionalizarse la evaluación educativa y las técnicas más empleadas para llevar a cabo su correcta ejecución.

Para ello el siguiente apartado está dedicado a presentar el diseño de una evaluación educativa, la cadena de resultados y los distintos indicadores implicados y necesarios para evaluar. A continuación, se detalla por un lado cómo realizar una evaluación de centros educativos utilizando para ello métodos de fronteras de producción y por otro lado los elementos esenciales de una evaluación de programas educativos mediante técnicas de evaluación de impacto. La última parte del capítulo está destinada a presentar cómo se implementó el programa Escuela 2.0. y cómo se podría haber evaluado. Finalmente, el capítulo concluye con unas reflexiones finales exponiendo algunas recomendaciones para articular la evaluación educativa en España.

EL DISEÑO DE UNA EVALUACIÓN EDUCATIVA: CADENA DE RESULTADOS, INDICADORES Y LÍNEA DE BASE

El sistema educativo español debe incorporar las principales técnicas de seguimiento y evaluación orientadas a resultados como pilar fundamental donde asentar el salto cualitativo que demanda la sociedad. Toda la información generada debe además ser transparente a la sociedad civil y a los agentes que forman el sistema educativo para que éstos puedan llevar a cabo una toma de decisiones adecuada.

La evaluación educativa debe estructurarse en una cadena de resultados para analizar tres grandes áreas: eficacia, eficiencia e impacto. La eficacia es una medida del grado de cumplimiento de los indicadores de calidad, resultados y equidad definidos por la política educativa. La eficiencia permite saber si los recursos destinados para cumplir con estos objetivos han sido, o no, los óptimos y por qué. Finalmente la evaluación de impacto permite determinar si existe causalidad para concluir si determinadas políticas y programas educativos han permitido o no mejorar los indicadores de resultados educativos de la población destinataria. En definitiva, el seguimiento y la evaluación permiten informar a los gestores del sistema educativo si éste está funcionando adecuadamente o no. Con esta valiosa información los responsables técnicos y políticos podrán proponer cambios, plantear nuevas acciones, suprimir otras y en definitiva contribuir a maximizar la eficacia, la eficiencia y el impacto de los recursos destinados a educación. Sin una evaluación propia no será posible tomar decisiones racionales basadas en evidencias y en el mejor de los casos se aplicarán recetas que resultaron exitosas en otros países. A continuación se desarrollan los principales pasos y herramientas que debemos utilizar para realizar una correcta evaluación de un sistema educativo[1].

CADENA DE RESULTADOS

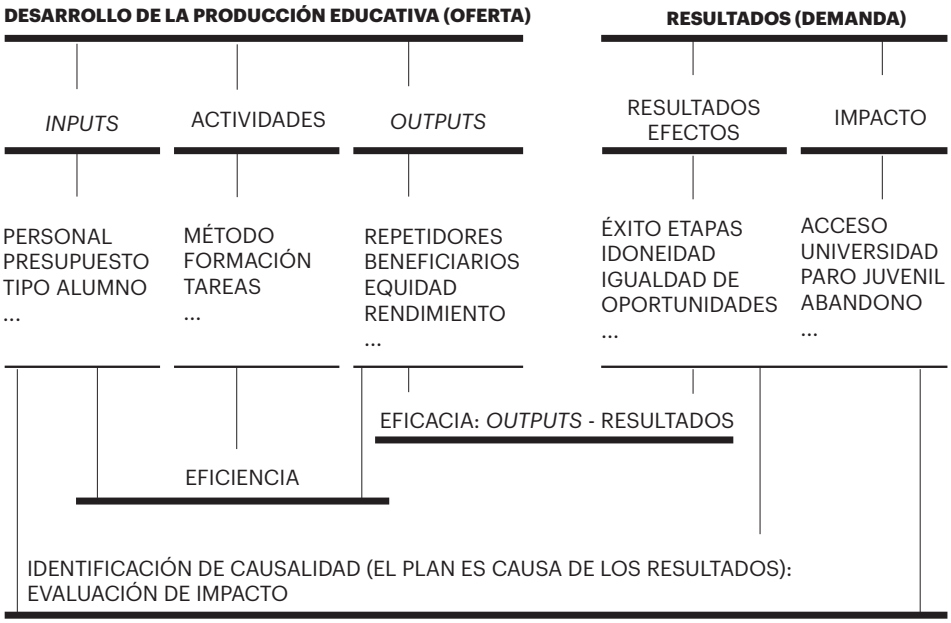
El primer paso que es necesario dar para poner en marcha una evaluación educativa es una cadena de resultados que describa claramente el marco lógico de cómo las políticas y los programas educativos van a conseguir alcanzar los objetivos planteados que deben estar bien definidos y ser cuantificables. Una cadena de resultados está compuesta por los siguientes eslabones:

1. Esta descripción no comprende todas las técnicas de evaluación disponibles. Es por ello que debe quedar claro que en este trabajo se exponen las técnicas cuantitativas que consideramos que más información aportarían a un proceso de evaluación. Sin embargo, otras técnicas de evaluación, fundamentalmente cualitativas, también pueden ser aplicadas para complementar los resultados del análisis cuantitativo.

- *Inputs*: Son los recursos necesarios para obtener una cantidad de producción. Necesariamente deben incluir recursos humanos y de capital en sentido amplio.
- Actividades: Son aquellas tareas necesarias para transformar los recursos en resultados. No deben confundirse con los *outputs* ya que estas actividades no son un fin en sí mismas sino los medios para conseguir este fin.
- *Outputs*: Producción final perseguida por el programa o la unidad productiva. Se obtienen una vez que se han empleado los *inputs* y se han ejecutado todas las actividades.
- Resultados: Resultados a corto y medio plazo que alcanza la población beneficiaria una vez que hace uso de los *outputs* entregados.
- Impacto: Grandes y últimos objetivos del programa a los cuales contribuyen las políticas educativas.

La Figura 1 ilustra de forma esquemática cómo se organizaría la cadena de resultados para mostrar la lógica causal de la política educativa.

FIGURA 1. LA CADENA DE RESULTADOS EN UNA POLÍTICA O PROGRAMA EDUCATIVO



Fuente: Elaboración propia a partir de Gertler *et al*. (2011).

La cadena de resultados describe una causalidad del tipo "Si..., entonces...". Por ejemplo, "Si en la escuela se refuerza el número de horas dedicadas a lectura en edades tempranas, entonces mejorará el rendimiento académico no solo en lectura sino también en otras disciplinas" (Machin y McNally, 2008). En este caso se necesitarían recursos para elaborar materiales específicos y una vez distribuido el material éste se entregaría a los alumnos realizando distintas actividades en clase. A partir de trabajar más horas de lectura en clase y en el hogar se esperaría que empiecen a mejorar los niveles de rendimiento académico tanto en cada curso como al final de cada etapa educativa. Cualquier política o programa de innovación educativa debería seguir una lógica similar a esta cadena dotada con sus indicadores correspondientes.

Por otro lado, en la cadena de resultados educativa se pueden distinguir dos fases que se relacionan directamente con las tareas posteriores de seguimiento y evaluación. En primer lugar, los centros de educación no universitaria, colegios e institutos, llevan a cabo cada curso y en cada nivel educativo un proceso productivo en sus alumnos del tipo, *inputs* –actividades– *outputs*; de la siguiente manera. Los colegios a partir de los *inputs* con los que cuentan (recursos escolares, nivel socioeconómico y estructura familiar de los alumnos, conocimientos previos), implementan una serie de actividades (metodología docente, implicación de los padres, tareas en el hogar, pruebas de evaluación, etc.) para transformar e incrementar los conocimientos de los alumnos a lo largo del curso. Debe tenerse en cuenta que el *output* final que se entrega a los alumnos es tan solo el resultado final del curso, el cual es controlable por la unidad productora, pero no el resultado final de la educación que se obtendrá una vez que cada cohorte finalice el proceso educativo en cada etapa. Cabe también destacar que esta tarea es llevada a cabo de forma repetida en el tiempo por parte de las unidades productivas, escuelas e incluso distintos profesores dentro de la misma escuela. Este proceso debe ser evaluado a través de técnicas de seguimiento y en particular de medición de la eficacia y la evolución de la productividad del proceso productivo.

En segundo lugar, una vez que los alumnos terminan el ciclo educativo podemos empezar a medir los resultados a corto, medio y largo plazo de la educación entregada por los centros. Estos resultados son los que produce el sistema educativo en forma de tasas de éxitos y rendimiento medio al final de cada etapa (corto y medio plazo) o acceso a la educación superior y al mercado de trabajo (largo plazo). Además, la medición y evolución de los resultados a lo largo del tiempo nos sirve para evaluar la causalidad de los distintos programas educativos que hayan sido puestos en marcha por las autoridades educativas o por centros innovadores. Las técnicas de evaluación de impacto permiten analizar esta relación causal entre los recursos del sistema educativo invertidos por el programa y los resultados finales que se producen en los alumnos a corto, medio y largo plazo.

INDICADORES

Una vez definida la cadena de resultados en una evaluación orientada a resultados a continuación es necesario definir y clarificar los objetivos del sistema educativo, y crear indicadores que permitan su medición y evolución en el tiempo. A la hora de definir y construir indicadores podemos seguir el criterio de UNDP (2009, p. 58) que señalan que un buen indicador debería ser SMART (en inglés un acrónimo para *Specific, Measurable, Achievable, Relevant* y *Time-bound*), es decir, debería cumplir las siguientes características:

- Específico: lo más cercano al concepto que se quiere medir.
- Medible: tanto si el indicador es cuantitativo o cualitativo, debe ser cuantificado de forma clara y no de forma ambigua en relación con lo que se mide.
- Alcanzable: que pueda ser alcanzable y atribuible al programa o política.
- Relevante: que el indicador aunque imperfecto realmente se aproxime a los objetivos perseguidos y sea relevante para la toma de decisiones.
- Con un plazo establecido: los indicadores deben poder obtenerse con la frecuencia deseada a un coste razonable.

El sistema de indicadores de resultados debería además, en la medida en que sea posible, lograr el máximo consenso y aceptación por parte de todos los agentes implicados que participan en la evaluación. En particular, quien evalúa, quien financia la evaluación y quien es evaluado, deberían alcanzar un compromiso en el que quede claro que se evalúa para mejorar y no para fiscalizar. Por otro lado los indicadores deben ser multidimensionales, fáciles de entender, recoger y revisar, y además contemplar aspectos ligados tanto a la cantidad como a la calidad pero también a la "no-calidad", en forma de repetición de curso, quejas y reclamaciones por parte de los padres, etc. Con ello se trata de evitar que los gestores se centren únicamente en unos pocos indicadores que primen su desempeño y desprecien el resto (este efecto es lo que Smith (1995) denomina "visión de túnel").

LÍNEA DE BASE O ESTUDIO DE DIAGNÓSTICO

Una vez definida la cadena de resultados y antes de empezar el proceso de evaluación resulta necesario establecer un diagnóstico o *línea de base* de los indicadores cuyos valores queremos modificar. Por tanto, la línea de base es la primera medición de todos los indicadores contemplados en el diagnóstico que permite establecer el 'punto de partida' y de comparación del proyecto o intervención. Suele tener

un carácter cuantitativo y puede recurrir tanto a fuentes primarias, producidas *ad-hoc*, como a secundarias, generalmente utilizando otras encuestas o la información contenida en datos administrativos. Sin información de partida será difícil comprobar los progresos reales de las medidas implementadas. La línea de base permite, mediante los datos obtenidos, caracterizar de forma más precisa a la población objetivo del proyecto, y con ello incluso se podrían reformular los objetivos para ganar mayor pertinencia, eficacia, eficiencia e impacto.

ÁREAS DE LA EVALUACIÓN EDUCATIVA: EFICACIA, EFICIENCIA E IMPACTO

EFICACIA

El nivel más básico, aunque no por ello menos importante en una evaluación educativa, es la recogida de todos los indicadores definidos en la cadena de resultados, especialmente de los referidos a los *outputs* y los resultados a corto plazo[2]. La eficacia mide el grado en el cuál los *outputs* y los resultados deseados son alcanzados. Esta definición lleva implícita que cada indicador debe estar asociado a un objetivo esperado. En este sentido, en las últimas décadas la efectividad de un programa educativo se medía a través del cumplimiento presupuestario. Si bien es necesario seguir recogiendo información acerca de si los recursos asignados y las actividades educativas programadas se corresponden con las observadas en la ejecución, resulta aún más necesario poner la atención en los resultados obtenidos (cantidad, calidad, equidad entendida como igualdad de oportunidades educativas y posibles resultados no deseados) ya que la correcta ejecución del gasto es condición necesaria pero no suficiente para que el Plan tenga éxito.

El control de la eficacia debería realizarse mediante un seguimiento de los indicadores clave que permita comparar la realidad del sistema educativo con los objetivos establecidos por la política educativa en el marco de una Ley o programa. A modo de ejemplo la Figura 2 muestra un esquema para llevar el seguimiento de un conjunto de colegios de Educación Primaria.

2. Debe advertirse que a veces resulta complicado delimitar perfectamente si un indicador es un *output* o un resultado. Es por ello que si bien la eficacia se corresponde con el cumplimiento de los *outputs*, en la Figura 1 y en esta discusión se extiende también el concepto a resultados a corto plazo.

FIGURA 2. SEGUIMIENTO DE LA EFICACIA DE COLEGIOS DE EDUCACIÓN PRIMARIA

INDICADORES (DEFINICIÓN)	LÍNEA DE BASE (LB) Y OBJETIVOS PRÓXIMOS DOS AÑOS			RECOGIDA DEL INDICADOR		
	LB	1	2	FRECUENCIA	INSTRUMENTO	RESPONSABLE
Presupuesto en personal (miles €)	49	50	52	Anual 31 Dic	Registros	Colegio
Presupuesto para gastos corrientes (€)	23	24	25	Anual 31 Dic	Registros	Colegio
Alumnos escolarizados (Nº)	280	285	295	Anual 30 Sep	Registros	Colegio
Alumnos por aula (Nº)	23	24	25	Anual 30 Sep	Registros	Colegio
Número de repetidores (%)	25	-	20	Bianual 31 Dic	Registros	Colegio
Rendimiento en matemáticas (prueba)	480	-	490	Bianual 1 Mar	Prueba Objetiva	Ministerio
Nivel estudios padres (años)	10	11	12	Anual 31 Dic	Encuesta	Consejería CC.AA.
...						

EFICIENCIA

La presencia de ineficiencias en la gestión educativa implica que sería posible una mejora paretiana de la misma, ya sea mediante la reducción de los recursos presupuestarios sin disminuir los *outputs* y resultados, o alternativamente, incrementando los resultados con los recusos disponibles. Las fuentes de la ineficiencia son diversas. En primer lugar, el marco legal puede condicionar los resultados si los objetivos no están claros y los incentivos del profesorado son limitados. En segundo lugar, los factores educativos pueden explicar parte de las diferencias si asumimos que no todos los métodos pedagógicos, por ejemplo la organización de la clase, la cantidad de tareas para el hogar, el desdoble de grupos, el tratamiento de la diversidad o la política de decidir la repetición de curso, son igualmente válidos. En tercer lugar, puede existir una incorrecta organización y/o gestión de los centros en ausencia de un marco competitivo (Levin, 1974; Nechyba, 2000; Woessman, 2001). Finalmente, la calidad, experiencia y carisma del profesor a la hora de transmitir conocimientos también puede explicar estos resultados[3].

3. Las fuentes de la ineficiencia son de difícil concreción ya que finalmente estos factores son las razones desco-

Para combatir la ineficiencia es necesario cuantificarla. Así, la medición de la eficiencia permite clarificar y relacionar cuales son los recursos, actividades y *outputs* de los colegios e institutos. Las ventajas de la evaluación del desempeño o la eficiencia de los centros educativos a través de su seguimiento son numerosas pero en particular permiten:

- Fomentar y animar la continua mejora en el desempeño.
- Promover los valores de la eficiencia en el uso de los recursos.
- Facilitar la competencia por comparación de resultados (*yardstick competition*) entre colegios lo que incentiva la mejora del desempeño.
- Identificar a las mejores unidades (*benchmarking*) para aprender de sus prácticas y tener así una referencia de actuación.

El seguimiento de la eficiencia y la evolución en el tiempo de la productividad total de los factores para los centros educativos, si se realiza de forma sistemática, permitiría:

- Elaborar un documento de buenas prácticas de gestión que puedan ser trasladadas al resto de centros.
- Analizar de forma detallada las causas que afectan al mal desempeño de los peores centros y en su caso tomar las medidas para que éstas no se reproduzcan.
- Estudiar la evolución del desempeño en el tiempo de tal forma que el seguimiento funcione como un sistema de diagnóstico y detección temprana de los problemas que puedan surgir en cada centro.
- Introducir innovaciones y mejoras en la gestión cuyos efectos puedan ser después cuantificados y en caso de éxito ser extendidos a otras unidades gestoras.

Para medir la eficiencia educativa es necesario estimar la función o frontera de producción a partir de los *inputs* o recursos escolares utilizados y las cantidades de *outputs* producidos[4]. Para ello, el enfoque teórico normalmente utilizado está basado en la función de producción educativa propuesta por Levin (1974), Hanushek

---

nocidas que explicarían por qué la producción es menor que la potencial dado un nivel de recursos. Leibenstein (1966) lo explicaba así: "For a variety of reasons people and organisations normally work neither as hard nor as effectively as they could. In situations where competitive pressure is light, many people will trade the disutility of greater effort, or search the utility of feeling less pressure and of better interpersonal relations".

4. En ocasiones la estimación de la función de producción tiende a identificarse con una función de producción promedio que sería resultado de una aproximación econométrica sencilla mediante mínimos cuadrados ordinarios. Sin embargo, cualquier libro de microeconomía define una función de producción de la siguiente manera: "A single *output* technology is commonly described by means of a production function f(z) that gives the maximum amount q of *output* that can be produced using *input* amounts $(z_1,...,z_{L-1}) > 0$" (Mas-Colell *et al.* 1995, p.129).

(1979) y más recientemente por Hanushek *et al.* (2013) que se expresa en la Ecuación 1:

$$A_i = f(B_i, S_i) \qquad (1)$$

donde el subíndice *i* se refiere al centro educativo. A_i representa el vector de *outputs* educativos, en general asociados al promedio de los resultados académicos obtenidos por los alumnos en pruebas objetivas. Por otra parte, los *inputs* educativos se dividen en B_i, que representa las características socioeconómicas medias de los hogares de los que provienen los alumnos; aproximadas normalmente por el nivel educativo y profesional de sus padres y el vector S_i que hace referencia a los recursos educativos del centro. Por las razones comentadas con anterioridad, la estimación de dicha función de producción educativa es llevada a cabo frecuentemente considerando la posible existencia de comportamientos ineficientes. En este caso, se estimaría una frontera productiva (Ecuación 2) basada en la observación de las mejores prácticas productivas, es decir, en aquellos centros educativos que maximizan los *outputs* de sus alumnos a partir de los *inputs* que utilizan. La medida de la ineficiencia sería por tanto, la distancia de cada centro a la frontera construida:

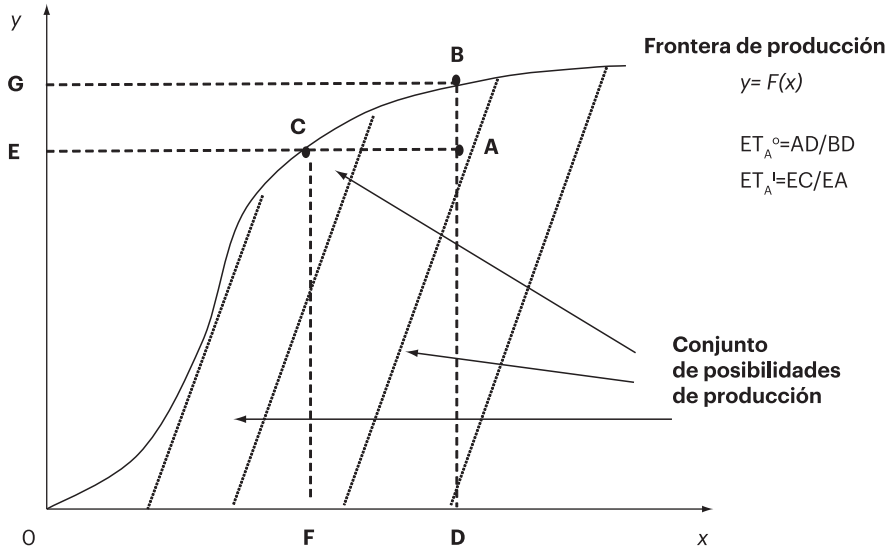
$$A_i = f(B_i, S_i) \times u_i \qquad (2)$$

donde u_i representa el nivel de eficiencia de cada escuela. Si u_i=1, ello supondría que el centro educativo *i* sería plenamente eficiente. Es decir, dado su nivel de *inputs* educativos, el nivel socioeconómico de sus alumnos y los recursos escolares con los que cuenta, y dada la tecnología existente este centro estaría maximizando el nivel de *outputs* que obtiene y por tanto, llevando a cabo una buena gestión de los recursos con los que cuenta. Valores de u_i<1, señalarían que el centro es ineficiente y por tanto la cuantía del índice de eficiencia indicaría la cantidad de *output* que se incrementaría si se eliminara la ineficiencia[5]. Veamos un ejemplo donde se ilustra (Figura 3) una función de producción sencilla que únicamente utiliza un único *input* para producir un único *output*.

En la Figura 3 observamos una función de producción muy simple $y = F(x)$ que permite transformar el *input* x en el *output* y mediante una tecnología $F(x)$. Las unidades de producción B y C pertenecen a la frontera productiva, esto es, dados los recursos que tienen disponibles no es factible que puedan producir más. Sin embargo, la unidad A pertenece al conjunto de posibilidades de producción. Es por tanto fac-

5. Una escuela con un índice de eficiencia; u_i=0,9 indicaría que dados los *inputs* con los que cuenta la escuela y dado el desempeño que están obteniendo otros centros similares con niveles de recursos parecidos, la escuela evaluada podría aumentar en un 10% todos sus *outputs* mediante una mejor gestión.

FIGURA 3. CONCEPTO Y MEDICIÓN DE LA EFICIENCIA TÉCNICA EN TÉRMINOS DE *INPUT* Y DE *OUTPUT*



tible que existan unidades, como A, con un desempeño por debajo de la combinación de puntos definida por la frontera. El índice de eficiencia de la unidad A puede ser medido mediante una orientación *output* o mediante una orientación *input*. En el primer caso, $ET_A^O = AD/BD$ expresa la eficiencia en términos de *output* indicando en cuánto podría la unidad A aumentar su *output* dados los recursos de los que dispone. En el segundo caso, $ET_A^I = EC/EA$ indica la eficiencia en términos de *input*, es decir, dada la producción de E unidades en cuánto podría reducir sus recursos para llegar a la frontera de producción[6].

La frontera de producción no es un concepto estático y puede cambiar a lo largo del tiempo. Así, el hecho de que la gestión de un servicio haya sido correcta durante un ejercicio no garantiza que lo será el siguiente. El sistema de seguimiento debe controlar el desempeño de las unidades de forma continuada en el tiempo para con ello poder anticiparse a los problemas que puedan surgir durante la ejecución del Plan. Para realizar este seguimiento continuado es necesario medir también la evolución de la productividad total de los factores, esto es, la evolución de la ratio entre la suma ponderada de *outputs* y de recursos utilizados por las unidades evaluadas.

Las técnicas analíticas empleadas para estimar la frontera de producción y medir la eficiencia y los cambios en la productividad total de los factores deben adaptarse a las peculiaridades de la producción educativa. Fundamentalmente esto pasa

6. Prácticamente la totalidad de los trabajos que estiman fronteras de producción educativa utilizan una orientación al *output*.

por reconocer la ausencia de precios para buena parte de la producción y el carácter multidimensional del *output*. Además, la tecnología de producción F(.) que permite entregar un servicio tan complejo es completamente desconocida, lo que aconseja la utilización de aproximaciones flexibles en su formulación y que no impongan supuestos demasiado restrictivos sobre la frontera de producción.

Los principales métodos de medición de la eficiencia técnica pueden ser divididos en dos grandes grupos: los métodos no paramétricos, basados en modelos de optimización matemática y los métodos paramétricos o econométricos[7]. El principal método no paramétrico utilizado en evaluación de políticas públicas es el análisis envolvente de datos (DEA son las siglas en inglés correspondientes con *Data Envelopment Analysis*). Su principal ventaja es que al no asumir ninguna forma funcional acerca de la tecnología que relaciona factores productivos y resultados y con supuestos poco restrictivos, como convexidad y monotonocidad, es capaz de trazar la frontera productiva. Los métodos paramétricos son igualmente válidos pero necesitan hacer supuestos clave acerca de varios aspectos del análisis como la forma funcional de la tecnología, la función de distribución del error o la distribución de la ineficiencia. A cambio permiten calcular elasticidades para conocer la importancia de cada *input* (Crespo-Cebada *et al.*, 2014).

En cuanto a la medición de los cambios en la productividad total de los factores, la opción más empleada en el análisis de centros educativos es nuevamente la aproximación de fronteras de producción, que tienen en cuenta explícitamente el posible comportamiento ineficiente de las unidades analizadas. El cálculo del índice de Malmquist es uno de los métodos más utilizados para analizar la evolución de la productividad total de los factores y sus componentes a lo largo del tiempo (Maragos y Despotis (2004), Brennan *et al.* (2014) y Essid *et al.* (2013)). El éxito de esta medida radica en que solamente necesita datos relativos a cantidades, no es necesario realizar supuestos sobre la forma funcional de la función de producción y permite descomponer la productividad total de los factores de los centros en cambios de productividad debidos a progreso tecnológico y a la mejora de la eficiencia (Coelli, *et al.*, 2005).

Por las consideraciones anteriormente remarcadas relacionadas con la actividad del sector público, los métodos no paramétricos[8] han sido mucho más utilizados para evaluar centros educativos (Worthington, 2001).

---

7. No es el objetivo de este capítulo realizar una descripción detallada de todas las metodologías disponibles. Para ello puede consultarse Coelli *et al.* (2005).
8. Para una descripción detallada sobre los pasos necesarios para aplicar en la práctica un análisis no paramétrico se recomienda la lectura de Emrouznejad y De Witte (2010).

EVALUACIÓN DE IMPACTO

Las políticas educativas deberían estar basadas en evidencias científicas empíricas que muestren una relación causal entre la política o el programa que se desee aplicar y el efecto perseguido. Para cumplir con este objetivo las técnicas de inferencia causal permiten cuantificar el impacto que tiene una política pública sobre una población. Una evaluación de impacto es por tanto la medición del cambio en un indicador relevante (éxito escolar, rendimiento académico, etc.) para los beneficiarios que sea atribuible a un programa educativo. Una evaluación de impacto consiste en realizar un experimento o cuasi-experimento social y para ello requiere la definición de una situación contrafactual o también llamado grupo de control. Lo que se desea evaluar es el comportamiento del indicador de interés con y sin programa. Idealmente, requiere comparar la situación *ex-ante* (Línea de base), previa a la introducción de una política social concreta y *ex-post*, una vez finalizado el programa o en algún punto intermedio del mismo.

La gran ventaja de las técnicas de evaluación de impacto sobre otras técnicas de evaluación del tipo "antes - después" o "con - sin" (programa) es que permite distinguir correlación de causalidad. Obtener evidencia científica de causalidad del tipo $X \rightarrow Y$ es costoso (tiempo, recursos, análisis, etc.). De hecho, en muchas ocasiones sólo será posible hacer estudios de correlaciones basados en regresiones múltiples.

El experimento social

La idea básica del experimento social es la siguiente. Sea un grupo de individuos que por simplicidad sólo pueden recibir dos tratamientos (D):

- Si no recibe el tratamiento (grupo contrafactual) $D=0$
- Si recibe el tratamiento (grupo tratado) $D=1$

Los individuos son asignados a los dos grupos de forma aleatoria. Después del programa observamos:

- Resultado de los tratados Y_1 en el grupo tratado $D=1$
  $E[Y_1|D=1]$
- Resultado de los no tratados Y_0 en el grupo no tratado $D=0$
  $E[Y_0|D=0]$

Nuestro objetivo teórico es determinar el efecto que el tratamiento, o en nuestro caso el programa educativo, tuvo sobre el grupo tratado, es decir:

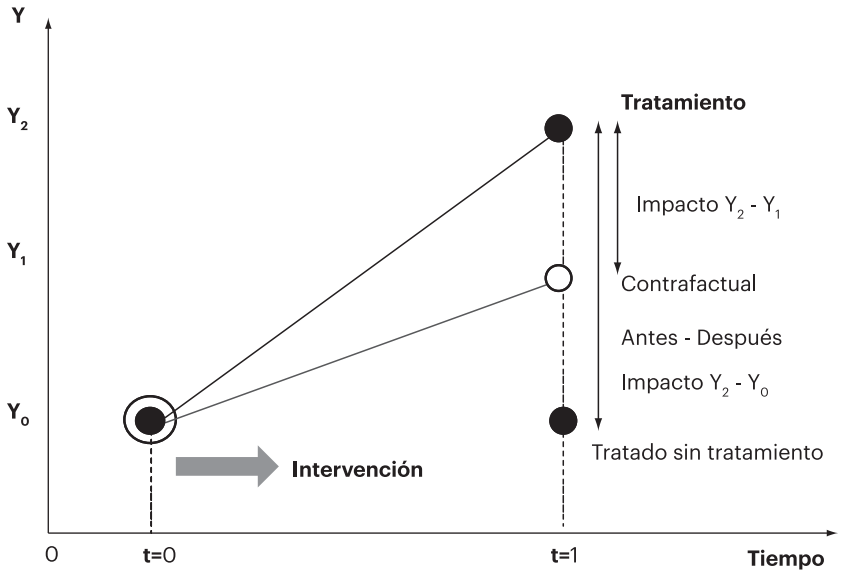
$$E[\Delta|D=1] = E[Y_1|D=1] - E[Y_0|D=1]$$

$E[Y_0|D=1]$ es inobservable, es decir, no es posible saber cuál hubiera sido el resultado del grupo tratado si no hubiera sido tratado. En la práctica, la solución pasa por asignar de forma aleatoria a los individuos a los dos grupos mediante un experimento controlado con grupos formados por individuos idénticamente distribuidos en todas las variables. De esta forma se logra que Y (el resultado) y D (si el individuo recibió o no el tratamiento) sean independientes. Por tanto, mediante la aleatorización se consigue que:

$$E[Y_0|D=1] = E[Y_0|D=0]$$

La Figura 4 muestra este proceso. Inicialmente tenemos un grupo de escuelas o alumnos en t=0 y mediante un estudio de diagnóstico recogemos información de línea de base de la variable que queremos estudiar Y_0. Al no haber ningún programa previo diferente los dos grupos presentarán el mismo valor medio. Posteriormente se produce una intervención educativa mediante la implantación de un determinado programa (reducción del tamaño de la clase, continuar con el mismo profesor varios cursos o cambiar cada curso, utilizar ordenador para enseñar matemáticas, etc.) en uno de los grupos. Tras realizarse esta intervención se vuelve a obtener el valor

FIGURA 4. EVALUACIÓN DE IMPACTO DE UN PROGRAMA



de la variable de interés que tiene el grupo de individuos tratado en el periodo t=1 que ahora presenta un valor Y_2.

En un estudio de evaluación clásico "antes-después", si sólo tenemos estos dos puntos (puntos negros de la Figura) no sabemos si la variación en la variable de interés Y a lo largo de este tiempo Y_2-Y_0 es debida a los efectos del programa o a otros motivos de la compleja realidad social (otras intervenciones públicas, factores exógenos, etc.).

Por esta razón, en una evaluación de impacto se designa aleatoriamente a un grupo de la población como grupo de control (o contrafactual) que esencialmente tiene las mismas características que la población beneficiaria salvo que no recibe el programa para que nos indique qué valor presentarían esos individuos en t=1 si el programa no se hubiera realizado. Esta situación viene dada por el punto Y_1, por lo tanto el impacto que tiene este programa vendría dado por la diferencia de los valores entre Y_2 e Y_1.

Determinar el grupo contrafactual que servirá para hacer la comparación no es tarea fácil y resulta por tanto el elemento clave en el diseño de la evaluación de impacto de un programa educativo. Es necesario recalcar que el grupo de control debe ser lo más parecido al grupo tratado salvo en que no resultó beneficiario del programa que se quiere estudiar.

El sesgo de selección

Uno de los problemas más relevantes que pueden afectar a la medición del impacto de un programa es el llamado sesgo de selección. Se produce sesgo de selección cuando los individuos no fueron asignados a los grupos de tratamiento y control de forma aleatorizada. El problema del sesgo de selección en una evaluación de impacto se produce cuando los individuos en los dos grupos (tratado y control) tienen determinadas características que no pueden ser observadas por el evaluador y afectan tanto a la decisión de participar en el programa como a sus resultados. Existen fundamentalmente dos tipos de sesgos:

1. Autoselección o heterogeneidad no observable: características individuales no observables entre los grupos como la motivación, habilidad, perspectivas, información, etc. que influyen tanto en pertenecer al grupo tratado como en los resultados del programa.
2. Heterogeneidad observable: por alguna razón hay diferencias significativas en características observables que afectan bien a recibir el tratamiento o bien a los resultados entre los grupos tratado y de control. Por ejemplo en

edad, renta, nivel educativo, tamaño de población, etc. Los resultados del análisis se pueden controlar si se dispone de esta información.

En el caso de una evaluación tradicional, además del caso de la evaluación "antes-después" es habitual analizar individuos tratados y no tratados de forma *ex-post*; o evaluación "con-sin" el programa, sin atender a las causas que determinan el tratamiento. En este caso, se controla por características observables y mediante una regresión múltiple se introduce una variable *dummy* que indica si el individuo fue beneficiario o no del programa. Por ejemplo, consideremos un curso de formación para jóvenes desempleados. Una forma de evaluar este programa sería un año después calcular la tasa de empleo y el salario medio de los que siguieron el curso y de los que no siguieron ningún curso. Sin embargo, es posible que exista un sesgo de selección y que ambos grupos no sean comparables si los individuos que decidieron participar son aquellas personas que tienen más que ganar con el programa y/o están más motivadas con encontrar empleo. Existe por tanto una diferencia no observable entre los dos grupos que sesga los resultados, ya que no sabremos si los mejores resultados del grupo tratado son consecuencia de su más alta motivación o del programa de formación seguido[9]. En la práctica, la determinación del grupo tratado y del grupo de control se puede realizar utilizando distintas metodologías que se dividen en dos grandes categorías: los diseños experimentales (aleatorios) y los diseños cuasi-experimentales (o diseños no aleatorizados)[10].

Los diseños experimentales son los métodos más robustos de todas las metodologías de evaluación de impacto (Angrist, 2004). Partiendo de la población, el estudio se realiza de forma bietápica en dos partes. En primer lugar se toma una muestra aleatoria representativa de la población (validez externa). A continuación, se realiza una asignación al azar de la pertenencia al grupo de beneficiarios y al grupo de control, de los individuos objeto de estudio, generando dos grupos cuyas características socioeconómicas son estadísticamente equivalentes entre sí (validez interna). La principal ventaja de esta técnica es la simplicidad en la interpretación de resultados. El impacto del programa sobre los resultados se puede medir directamente como la diferencia entre las medias de las muestras del grupo de tratamiento y el grupo de control.

---

9. Además de la asignación aleatoria de la formación existen otras soluciones prácticas a este problema. Por ejemplo se puede establecer un programa con sobredemanda estableciendo un número de plazas limitadas o un punto de corte para el acceso al programa donde se pueda evaluar el programa en el entorno de este punto de corte.
10. Para una revisión detallada de las técnicas de evaluación de impacto tanto experimentales como cuasi-experimentales puede acudirse a Gertler *et al.* (2011) y Khandker *et al.* (2010). Para el lector no familiarizado se recomienda la lectura de Schlotter *et al.* (2011) que realizan una excelente descripción de las principales técnicas de evaluación de impacto en referencia al ámbito educativo.

La aleatorización presenta algunos problemas en cuanto a su diseño. En primer lugar, pueden existir problemas éticos o restricciones políticas si se considera que toda la población debe tener derecho al programa. En este caso las principales soluciones pasan por rotar los programas o por establecer programas piloto (*phase-in*) que resultan muy deseables ya que son baratos, permiten la innovación y la experimentación educativa y en caso de ser exitosos se pueden extender a las poblaciones que inicialmente sirvieron de grupo de control mitigando así los problemas éticos. En segundo lugar, puede existir "contagio" si los grupos tratado y no tratado interaccionan o si varios programas son aplicados simultáneamente en una población, en cuyo caso será difícil dirimir posteriormente cuál fue el que causó los efectos medidos.

En el caso de que la aleatorización no sea posible se puede recurrir a las técnicas de cuasi-experimentación. Igualmente, si sabemos que los individuos no fueron asignados de forma aleatoria a los grupos tratado y control tenemos que acudir a estas técnicas basadas en la búsqueda de un contrafactual. En la práctica, se recurre a la cuasi-experimentación de forma *ex-post* justo cuando se quiere llevar a cabo la evaluación pero el programa no fue diseñado para ser evaluado.

Los métodos cuasi-experimentales tienen la ventaja de ser más rápidos y baratos ya que su implementación se basa en datos ya existentes donde se dispone de información de beneficiarios y no beneficiarios. Sus desventajas son su menor fiabilidad respecto a los métodos de aleatorización si no se controla por todas las diferencias entre los dos grupos y requieren además métodos econométricos más complejos. Dependiendo del tipo de información disponible sobre los grupos tratado y contrafactual; "antes - después" de la intervención o sólo después, y de cómo se realizó la asignación de los individuos a los grupos se seleccionará la mejor técnica para cada caso concreto.

UN EJEMPLO: PLAN DE SEGUIMIENTO Y EVALUACIÓN (S&E) DEL PROGRAMA ESCUELA 2.0

En el presente apartado, utilizaremos un programa que fue emblemático en España, el programa Escuela 2.0 implementado entre 2009 y 2012 para ejemplificar, en términos generales, cómo debería diseñarse el proceso de S&E a la hora de implantar un programa educativo.

BREVE DESCRIPCIÓN DEL PROGRAMA ESCUELA 2.0

El Programa Escuela 2.0 tuvo por objetivo integrar las tecnologías de la información y la comunicación (TICs) en los centros educativos sostenidos con fondos públicos.

Para ello, el programa consistía en dotar a los alumnos y profesores con un ordenador portátil, proveer la conectividad a internet y la interconectividad dentro del aula para todos los equipos, promover la formación del profesorado y generar y facilitar el acceso a materiales digitales educativos ajustados a los diseños curriculares tanto para profesores como para el alumnado y sus familias. A su vez, se buscaba implicar a los alumnos y a las familias en la adquisición, custodia y uso de estos recursos.

El programa comenzó en el curso 2009-10 con 5º de Primaria, y posteriormente, se extendería a 6º de Primaria y a los dos primeros cursos de la ESO. El presupuesto inicial del programa previsto era de 200 millones de euros por curso, cofinanciados al 50% entre el Gobierno y las Comunidades Autónomas (CC. AA.) que participaron. En 2012 el programa fue cerrado y sustituido por un nuevo programa de cooperación territorial de TICs. Por tanto, en los tres años que se implementó el programa se invirtieron alrededor de 600 millones de euros.

Existen numerosos estudios en la literatura de la Economía de la Educación que demuestran que más recursos tecnológicos (ordenadores, salas digitales, conexión a internet, etc.) no producen *per se* mejores resultados en el aprendizaje de los alumnos (Angrist and Lavy, 2002; Fuchs and Woessmann, 2004; Banerjee *et al.*, 2007; Warschaver 2008). Una posible explicación radica en que introducir la tecnología para "informatizar" procesos y prácticas educativas existentes no implica una innovación real en el proceso de aprendizaje. Esto es, el éxito del programa Escuela 2.0 requiere además de la provisión de los recursos tecnológicos de acciones y actividades que impliquen un cambio en las prácticas educativas tradicionales que aseguren una adecuada y eficaz utilización de los recursos para lograr los resultados y que se definan. En este sentido, un diseño *ex-ante* de la evaluación, un oportuno plan de seguimiento durante la implementación del programa y una adecuada evaluación *ex-post* hubieran sido fundamentales para conocer si efectivamente el programa fue implementado eficientemente y ha logrado impactar positivamente en los resultados educativos de los estudiantes. A continuación, se presenta una propuesta de algunos elementos básicos que conforman un adecuado plan de S&E de una intervención educativa que hubieran podido aplicarse al programa Escuela 2.0.

ELEMENTOS DE UN PLAN DE S&E PARA EL PROGRAMA ESCUELA 2.0

Implementación de un plan piloto

Los programas de alto contenido tecnológico como el Escuela 2.0 implican la inversión de una gran cantidad de recursos económicos. Hubiera sido por tanto adecuado instaurar el programa inicialmente solo en algunos centros seleccionados aleatoria-

mente y realizar un seguimiento y evaluación del plan piloto. Como hemos comentado, la idea de llevar a cabo una experiencia piloto es poder conocer efectivamente si el programa mejora el aprendizaje y las competencias de los estudiantes, y a su vez conocer cuáles son las formas de utilizar los recursos asignados que producen mejores resultados y cuáles no tienen efectos. Esta información permitiría en primer lugar, corroborar si el impacto es positivo y significativo, y en base a ello calcular cuáles serían los costes y beneficios de extender el programa a toda la población. Es decir, si tiene sentido o no aplicar el programa a nivel nacional o sobre qué colectivos tendría sentido aplicarlo. Por tanto, la evaluación piloto permite implementar el programa de forma eficiente.

Sin embargo, este plan piloto no se llevó a cabo con el programa Escuela 2.0. En ocasiones la idea de experiencia piloto se confunde con implementar un programa en etapas, y detectar y corregir sobre la marcha simplemente problemas de logística en la provisión de los recursos (ordenadores, instalación de redes y aulas digitales). Sin embargo, el plan piloto requiere un verdadero S&E de la experiencia para obtener información de si efectivamente se alcanzan los resultados esperados con el programa.

Definición de objetivos y línea de base del S&E, ¿qué queremos evaluar?

El S&E de un programa o política debe tener objetivos claros, esto es, definir inequívocamente qué cuestiones se quieren analizar y dar respuestas. Así por ejemplo, un plan de seguimiento del programa Escuela 2.0 debería poder responder a las siguientes interrogantes: ¿cuántos ordenadores se han entregado efectivamente?, ¿cuántos se estropean y en cuánto tiempo?, ¿cuáles son los centros que mejor resultado en competencia digital obtienen dados los recursos asignados? (*Benchmarking*), ¿qué actividades llevan a cabo estos centros?, ¿qué acciones están implementando los centros menos eficientes?

En el caso de la evaluación de impacto, las preguntas deberían enfocarse en los resultados de medio plazo, o en el impacto del programa: ¿qué impacto ha tenido Escuela 2.0 en el aprendizaje de competencia digital?, ¿qué impacto ha tenido Escuela 2.0 en el aprendizaje de otras disciplinas?, ¿ha tenido Escuela 2.0 algún impacto sobre la tasa de repetición o fracaso escolar?, ¿ha tenido Escuela 2.0 algún impacto sobre la tasa de abandono a temprana edad?, ¿cuál es la satisfacción de los padres y los profesores con el programa?

Junto a estas preguntas es necesario recabar información de la situación inicial antes de implementar el programa. Este hecho es crucial para poder llevar a cabo una evaluación de impacto del programa. En este caso, sería necesario disponer de información acerca del nivel de conocimientos en competencia digital (y otras competencias básicas) que los alumnos poseen antes de la intervención, así como de sus

características socioeconómicas. A su vez, para poder llevar a cabo una comparación justa de los logros alcanzado por cada centro educativo, sería necesario controlar el impacto de la intervención por otros factores que también afectan al proceso de aprendizaje y por tanto al rendimiento de los alumnos. En este sentido, la línea de base debe proporcionar también información, por ejemplo, acerca del nivel de conocimiento y experiencia previo que tienen los profesores en materia de TICs así como de la disponibilidad inicial de TICs en los centros educativos.

DEFINICIÓN DE LOS INDICADORES DE LA CADENA DE RESULTADOS DEL PROGRAMA

Indicadores de recursos (*inputs*)

Los *inputs* pueden definirse tanto en términos monetarios como en recursos físicos. En el caso del programa Escuela 2.0, no se asignó un presupuesto a cada unidad ejecutora (centro educativo), sino que se otorgaron los recursos físicos directamente y por tanto no se dispone de información presupuestal del programa a nivel de centro educativo. Los *inputs* deberían definirse entonces en términos físicos o no monetarios[11]. Sin ánimo de exhaustividad, algunos ejemplos de indicadores de *inputs* para el programa Escuela 2.0 podrían ser:

- Número de ordenadores entregados a los alumnos.
- Número de alumnos potenciales receptores de ordenador en el centro.
- Número de ordenadores entregados a los profesores.
- Número de profesores potenciales receptores de ordenador en el centro.
- Número de aulas digitales instaladas.
- Número de horas de formación a profesores.
- Características del alumnado:
    - Índice socioeconómico medio de los alumnos del centro (aproximado por el nivel educativo y recursos económicos de los padres).
    - Otras características familiares.
    - Resultados en pruebas de competencia digital y competencias básicas antes de aplicar el programa.
- Características (previas al programa) del centro:
    - Número de ordenadores en el centro.

11. Podría llevarse a cabo un ejercicio para recuperar esta cifra imputando el coste de los recursos otorgados a cada centro a partir del coste unitario de los ordenadores, aulas digitales, horas de formación, etc. que recibe cada centro educativo.

- Si contaban o no con aula digital.
- Número de horas de uso de los ordenadores por los alumnos.

- ...

Indicadores de actividades

Deberían reflejar las acciones emprendidas por cada centro con los recursos tecnológicos otorgados por el programa. Como hemos mencionado, los recursos en sí mismos no producen mejoras en el aprendizaje de los alumnos, sino a través de los métodos docentes que los emplean. Los indicadores de actividad son fundamentales sobre todo a la hora de llevar a cabo el seguimiento del programa, ya que estos indicadores permiten explicar o al menos disponer de hipótesis acerca de por qué se producen diferencias en los resultados del programa entre centros. Algunos ejemplos de actividades del programa Escuela 2.0 podrían ser:

- Número de horas de formación en cursos brindados a los docentes en el centro.
- Porcentaje de docentes que han asistido a los cursos de formación del centro.
- Resultados de los profesores tras una evaluación de sus competencias.
- El centro provee de contenidos educativos para el aprendizaje.
- El centro brinda apoyo institucional a proyectos educativos innovadores.
- El centro proporciona materiales digitales extras.
- Número de entrevistas, tutorías y contactos virtuales con las familias al mes.
- Número de horas de clase con portátil y/o en aulas digitales.
- El centro posee un coordinador TIC.
- En centro posee un técnico informático o servicio técnico online.
- El centro ha tenido alguna experiencia previa en programas TICs.
- Porcentaje de alumnos/docentes/familias que utilizan redes y portales.
- Uso de software libre/Uso de software convencional.
- ...

Indicadores de productos (*Outputs*)

Estos indicadores deben reflejar y cuantificar la producción final obtenida a partir de los recursos tecnológicos proporcionados por el programa y de las acciones llevadas a cabo en las aulas y centros. Como ya se ha señalado muchas veces, la distinción entre producto y resultado es difusa y, por tanto, su definición final dependerá de los objetivos establecidos en el programa y de qué se quiere evaluar. En este caso, si el objetivo del programa fuera mejorar las capacidades digitales de los estudiantes y que en el mediano plazo esto impacte positivamente en los resultados de otras

competencias, los indicadores de productos deberían al menos cuantificar el aprendizaje en competencia digital:

- Resultado en pruebas de competencia digital al finalizar el curso.
- Porcentaje de alumnos por encima del umbral mínimo de competencia digital.
- Diferencia en competencia digital entre alumnos del decil más alto y bajo.

Un enfoque complementario sería fijar los objetivos en función del uso efectivo de los recursos tecnológicos y del grado de aceptación y satisfacción con los mismos de los principales agentes involucrados en el proceso educativo:

- Número de ordenadores entregados que se conservan en correcto funcionamiento.
- Grado de satisfacción de las familias/profesores.
- Grado de satisfacción de los alumnos con el uso de TICs.
- Grado de satisfacción del profesorado con los cursos recibidos.

Indicadores de resultados

Como hemos mencionado, el objetivo a medio plazo de un programa educativo debiera ser mejorar los logros académicos de los estudiantes y por tanto la calidad del sistema educativo en su conjunto. En este sentido, los posibles indicadores de resultados del programa Escuela 2.0 serían:

- Resultado en pruebas de competencia digital al finalizar la etapa educativa.
- Resultados en pruebas objetivas de competencias básicas al finalizar la etapa educativa.
- Tasa de repetición en Educación Secundaria obligatoria.
- Diferencias en los resultados de las pruebas de competencias básicas y digitales entre distintos grupos de alumnos.

Indicadores de impacto

Finalmente, deberían establecerse cuáles son los grandes y últimos objetivos del programa. Algunos ejemplos serían:

- Tasa de abandono educativo.
- Porcentaje de alumnos que acceden a carreras con base tecnológica.
- Aumento del empleo en sectores con base tecnológica.

Una vez definidos los indicadores del plan de S&E es preciso determinar cómo se va a recabar la información necesaria antes, durante y después de la implementación del programa. Existen diversas fuentes y métodos de recopilación de la información y la selección final depende normalmente de los recursos disponibles. Para el caso del programa Escuela 2.0 la información podría obtenerse de las siguientes fuentes:

- Pruebas de conocimiento objetivas aplicadas a nivel nacional en competencia digital al inicio y al final del programa. Además de la prueba, sería adecuado aplicar una encuesta personal a los alumnos, directores y profesores para recabar las características iniciales comentadas anteriormente de centros y alumnos.
- Evaluaciones externas en competencias básicas, de la CC. AA. o a nivel nacional (Evaluaciones Generales de Diagnóstico) que se apliquen antes y después del programa.
- Evaluaciones internacionales en competencias básicas (PISA, TIMSS, PIRLS) que se hayan aplicado antes y después del programa.
- Registros administrativos.
- Otros métodos: observación de aulas, grupos de discusión, etc.

EVALUACIONES DEL PROGRAMA ESCUELA 2.0.

Hasta el momento existen dos grandes evaluaciones del programa Escuela 2.0 llevadas a cabo por grupos externos de investigación financiadas por el entonces llamado Ministerio de Ciencia e Innovación.

Proyecto tisce 2.0.

En el Proyecto *Las políticas de un "ordenador por niño" en España. Visiones y prácticas del profesorado ante el programa escuela 2.0. Un análisis comparado entre Comunidades Autónomas*[12], financiado por el Plan Nacional de investigación 2010 del entonces llamado Ministerio de Ciencia e Innovación, participan 13 universidades españolas y la U. de Coimbra de Portugal, y es impulsado por el grupo EDULLAB (U. de La Laguna). La evaluación está comprendida para ser ejecutada entre 2011 y 2014.

La evaluación tiene por objetivos en primer lugar conocer las opiniones, expectativas y valoraciones del profesorado. En segundo lugar, explorar el uso de prácticas de enseñanza o actividades didácticas con TIC en el aula, comparar entre

12. Para mayor detalle del proyecto ver http://tecedu.webs.ull.es/ticse20/index.php?lang=es.

CC.AA. y finalmente crear un observatorio sobre políticas TIC (países iberoamericanos). Para ello el proyecto encuesta a profesores y alumnos y realiza estudios de casos de aulas en distintas CC. AA. que aplican y no aplican el programa Escuela 2.0. con el fin de responder a las siguiente preguntas:

- ¿Qué efectos e impacto pedagógico y organizativo tendrá esta abundancia de tecnologías en las aulas ordinarias?
- ¿En qué medida cambiará la metodología de enseñanza desarrollada en la clase y el aprendizaje del alumnado?
- ¿Qué impacto tendrá sobre los libros de texto y materiales didácticos tradicionales?
- ¿Qué visiones y opiniones tiene el profesorado sobre este programa?
- ¿Existen diferencias en estas visiones y prácticas docentes por Comunidades Autónomas?

Proyecto grupo LOCALCOM

El proyecto *Las políticas públicas de impulso a la sociedad de la información en España: evaluación de las estrategias y actuaciones en nueve Comunidades Autónomas*[13] forma también parte del Plan Nacional de Investigación, financiado por el entonces llamado Ministerio de Ciencia e Innovación.

En este caso la metodología consiste en un modelo propio de análisis métrico de 24 indicadores (Índice Localcom, basado en tres categorías: oportunidades, infraestructuras y usos); y un modelo descriptivo para el estudio de las políticas autonómicas en TICs. Los indicadores se agrupan en siete dimensiones: pertinencia, efectividad, coherencia, permanencia, eficacia, impacto y eficiencia. Por ejemplo, en este caso los indicadores de eficacia, eficiencia e impacto son los siguientes:

Eficacia:

- ¿Se han obtenido distintos resultados según la tipología de los centros?
- ¿Cuál es la titularidad de los contenidos educativos digitales?
- ¿Se han implementado entornos de aprendizaje virtuales para el diseño, seguimiento y evaluación de actividades educativas?

Eficiencia:

- ¿Se utilizaron herramientas o recursos existentes?

13. Para mayor detalle del proyecto ver http://centresderecerca.uab.cat/oic/content/politicas-publicas-sociedad-informacion.

· ¿Se impulsó una entidad con participación de todos los agentes implicados?
· ¿El proyecto estuvo adscrito a una estructura institucional única?

Impacto:
· Número de ordenadores para estudiantes (acumulado en el tiempo).
· Número de aulas digitales (acumulado en el tiempo).
· Número de profesores formados (acumulado en el tiempo).

Como puede observarse, ambas evaluaciones son valiosas, ya que pondrán a disposición nueva información relevante acerca de la implementación del programa Escuela 2.0 y de los resultados obtenidos. Sin embargo, en ninguna de ellas el objetivo final es llevar a cabo una evaluación de la eficiencia de los centros o una evaluación de impacto sobre las competencias desarrolladas por los alumnos. No obstante, de la información que se derive será posible construir algunos de los indicadores de *inputs*, actividades, productos, resultados e impacto comentados anteriormente, que den lugar en el futuro a evaluaciones cuasi experimentales sobre el impacto que realmente ha tenido el programa en los alumnos beneficiarios del mismo.

REFLEXIONES FINALES

La institucionalización de la evaluación de los centros y programas educativos debería ser el eje que vertebre del sistema educativo español en los próximos años. Así, cualquier nuevo programa o intervención para la mejora de la educación debería ser acompañado desde su diseño de un documento que detalle cómo será evaluado.

El primer requisito para llevar a cabo una adecuada evaluación es disponer de información de calidad en diversos niveles del sistema educativo. En este sentido, la reciente introducción de pruebas individualizadas debería contribuir a mejorar la transparencia del sistema educativo y proveer información cualificada para que los diversos agentes implicados en el proceso educativo tomen mejores decisiones basados en la evidencia empírica. Los padres deberían tener información de los conocimientos que sus hijos efectivamente adquieren en la escuela que eligieron teniendo además en cuenta el nivel medio que alcanzan otros niños en otros colegios del barrio, municipio o municipios cercanos. A su vez, las autoridades educativas deberían usar esta información para llevar a cabo el seguimiento y la evaluación de impacto de las medidas e iniciativas educativas que se vayan poniendo en marcha.

En segundo lugar, España posee un grave problema de abandono y fracaso escolar. Es por ello que la evaluación continua debe ser un mecanismo esencial a

la hora de detectar a los alumnos con alto riesgo de fracaso escolar para intervenir tempranamente desde la escuela y evitar que la contingencia efectivamente suceda garantizando de este modo la igualdad de oportunidades educativas. La evaluación de impacto de los programas de intervención que se pongan en marcha en este sentido deben ser la herramienta que garantice tanto la consecución de los objetivos como la eficiencia de los recursos invertidos.

Finalmente, el seguimiento y la evaluación educativa debería ser realizada por una entidad externa a la administración educativa lo más independiente posible pero contando con la colaboración de las instituciones evaluadas responsables del diseño y ejecución de la intervención educativa. En el supuesto de que la evaluación se lleve a cabo a nivel regional o local, una opción podría ser crear Agencias Autonómicas de Evaluación Educativa que dependan exclusivamente del Presidente de la Comunidad Autónoma y que por lo tanto tengan plena capacidad e independencia para llevar a cabo las evaluaciones. La segunda opción sería la contratación de equipos de consultores externos independientes que realicen la evaluación en estrecha colaboración con las Consejerías de Educación e incluso incorporando en el equipo evaluador algún miembro de la propia Consejería que conozca el detalle del programa evaluado.

En definitiva, las políticas educativas deberían estar basadas en evidencias científicas empíricas que permitan identificar cuáles centros son los que mejor desempeñan y gestionan sus recursos y por qué y a su vez, establecer la relación causal entre un programa y el efecto perseguido. La importancia de evaluar se puede resumir utilizando la frase del célebre físico y matemático William Thomson quien afirmó: "Lo que no se define no se puede medir. Lo que no se mide, no se puede mejorar. Lo que no se mejora, se degrada siempre".

REFERENCIAS

Angrist, J. y Lavy, V. (2002). "New Evidence on Classroom Computers and Pupil Learning*". *The Economic Journal*, 112(482), 735-765.

Angrist, J. D. (2004). "American education research changes tack". *Oxford Review of Economic Policy*, 20(2), 198-212.

Banerjee, A. V., Cole, S., Duflo, E. y Linden, L. (2007). "Remedying education: Evidence from two randomized experiments in India". *The Quarterly Journal of Economics*, 122(3), 1235-1264.

Becker, G. S. (2009). *Human capital: A theoretical and empirical analysis, with special reference to education*: University of Chicago Press.

Brennan, S., Haelermans, C. y Ruggiero, J. (2014). "Nonparametric estimation of education productivity incorporating nondiscretionary inputs with an application to Dutch schools". *European Journal of Operational Research*, 234(3), 809-818.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J. y Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*: Springer.

Crespo-Cebada, E., Pedraja-Chaparro, F. y Santín, D. (2014). "Does school ownership matter? An unbiased efficiency comparison for regions of Spain". *Journal of Productivity Analysis*, 41(1), 153-172.

Emrouznejad, A. y De Witte, K. (2010). "COOPER-framework: A unified process for non-parametric projects". *European Journal of Operational Research*, 207 (3), 1573-1586.

Essid, H., Ouellette, P. y Vigeant, S. (2014). "Productivity, efficiency, and technical change of Tunisian schools: a bootstrapped Malmquist approach with quasi-fixed inputs". *Omega*, 42(1), 88-97.

Fuchs, T. y Woessmann, L. (2004). "What accounts for international differences in student performance? : CESifo working papers".

Gertler, P. J., Martínez, S., Premand, P., Rawlings, L. B. y Vermeersch, C. M. (2011). *Impact Evaluation in Practice; La Evaluación de Impacto en la Práctica*: World Bank Publications.

Hanushek, E. A. (1979). "Conceptual and empirical issues in the estimation of educational production functions". *Journal of human Resources*, 14(3).

Hanushek, E. A., Link, S. y Woessmann, L. (2013). "Does school autonomy make sense everywhere? Panel estimates from PISA". *Journal of Development Economics*, 104, 212-232.

Hoxby, C. M. (1999). "The productivity of schools and other local public goods producers". *Journal of Public Economics*, 74(1), 1-30.

Khandker, S. R., Koolwal, G. B. y Samad, H. A. (2010). *Handbook on impact evaluation: quantitative methods and practices*: World Bank Publications.

Leibenstein, H. (1966). Allocative efficiency vs. "X-efficiency". *The American Economic Review*, 392-415.

Levin, H. M. (1974). "Measuring efficiency in educational production". *Public Finance Review*, 2(1), 3-24.

Levin, H. M. (2012). "More than just test scores". *Prospects*, 42 (3), 269-284.

Lochner, L. y Moretti, E. (2001). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports: National Bureau of Economic Research.

Machin, S. y McNally, S. (2008). "The literacy hour". *Journal of Public Economics*, 92(5), 1441-1462.

Maragos, E. K. y Despotis, D. K. (2004)". Evaluating School Performance over Time in the frame of Regional Socio-Economic Specificities". *WSEAS Transactions on Mathematics*, 3(3), 664-670.

Mas-Colell, A., Whinston, M. D. y Green, J. R. (1995). *Microeconomic theory* (Vol. 1): Oxford university press New York.

Nechyba, T. J. (2000). "Mobility, targeting, and private-school vouchers". *American Economic Review*, 130-146.

Schlotter, M., Schwerdt, G. y Woessmann, L. (2011). "Econometric methods for causal evaluation of education policies and practices: a non-technical guide". *Education Economics*, 19(2), 109-137.

Smith, P. (1995). "On the unintended consequences of publishing performance data in the public sector". *International Journal of Public Administration*, 18(2-3), 277-310.

Steering Committee for the Review of Government Service Provision (2013). *Report on Government Services* 2013, Productivity Commission, Canberra.

Taylor-Powell, E. y Henert, E. (2008). "Developing a logic model: Teaching and training guide". *Benefits*, 3, 22.

UNDP (2009). *Handbook of Planning, Monitoring and Evaluating for Development Results*, New York.

Warschauer, M. (2008). "Laptops and literacy: A multi-site case study". *Pedagogies: An International Journal*, 3(1), 52-67.

Worthington, A. C. (2001). "An empirical survey of frontier efficiency measurement techniques in education". *Education Economics*, 9(3), 245-268.

Wößmann, L. (2001). "Why Students in Some Countries Do Better: International Evidence on the Importance of Education Policy". *Education Matters*, 1(2), 67-74.

Este libro recoge los resultados de un proyecto de investigación impulsado por las Fundaciones *Ramón Areces* y *Europea Sociedad y Educación*, en el que un grupo de economistas de la educación exploran aspectos fundamentales de los sistemas educativos en clave de competencia y competitividad.

FUNDACIÓN
RAMÓN ARECES

Vitruvio, 5 – 28006 Madrid
www.fundacionareces.es
www.fundacionareces.tv

Fundación Europea Sociedad y Educación
European Foundation Society and Education

José Abascal, 57 - 28003 Madrid
www.sociedadyeducacion.org