# ECONOMÍA DE LA EDUCACIÓN

Antonio Cabrales
Ismael Sanz
(Coordinadores)

HAND books

FUNDACIÓN RAMÓN ARECES

# ECONOMÍA DE LA EDUCACIÓN

**Antonio Cabrales**
**Ismael Sanz**
(Coordinadores)

# ECONOMÍA DE LA EDUCACIÓN

FUNDACIÓN
RAMÓN ARECES

Reservados todos los derechos.

Ni la totalidad ni parte de los libros pueden 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 Editorial Centro de Estudios Ramón Areces, S.A.

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

© EDITORIAL CENTRO DE ESTUDIOS RAMÓN ARECES, S.A.
Tomás Bretón, 21 – 28045 Madrid
Teléfono: 915 398 659
cerasa@cerasa.es
www.cerasa.es

© FUNDACIÓN RAMÓN ARECES
Vitruvio, 5 – 28006 MADRID
www.fundacionareces.es

ISBN-13: 978-84-9961-456-4
Depósito legal: M-8534-2024

Impreso por:
Campillo Nevado, S.A.
Desierto de Tabernas, 8
28320 Pinto (Madrid)

Impreso en España / Printed in Spain

# ÍNDICE

PRÓLOGO ..... XVII

Capítulo 1. INDICADORES DEL SISTEMA EDUCATIVO ESPAÑOL 1

Resumen ..... 1

1. Indicadores de escolarización y entorno educativo en España ..... 2
   1.1. Alumnado, tasas de matriculación ..... 2
      1.1.1. Educación Infantil de primer y segundo ciclo ..... 2
      1.1.2. El descenso del alumnado matriculado en segundo ciclo de Infantil y Primaria ..... 7
      1.1.3. La expansión de la FP Media en España y la estabilidad de Bachillerato ..... 9
      1.1.4. Más población con estudios superiores en España que en la UE 13
   1.2. Profesorado y entorno de aprendizaje ..... 17

2. Indicadores de financiación educativa ..... 20
   2.1. Gasto total en educación, público y privado ..... 20
   2.2. Becas y ayudas públicas ..... 25

3. Indicadores de resultados educativos ..... 30
   3.1. Tasa de abandono temprano educativo ..... 30
   3.2. Tasas de graduación en ESO, FP y Bachillerato ..... 35
   3.3. Tasas de repetición en Primaria y Secundaria en España en comparación con OCDE y UE ..... 38
   3.4. Resultados de los estudiantes españoles en pruebas internacionales de Primaria. PIRLS y TIMSS ..... 40
   3.5. Resultados de los estudiantes españoles en Secundaria: PISA ..... 45

4. Educación y empleo ..... 53

Bibliografía ..... 59

Capítulo 2. MÉTODOS EMPÍRICOS: BREVE INTRODUCCIÓN A LA INFERENCIA CAUSAL ..... 63

1. Introducción ..... 63

2. ¿Qué efecto queremos conocer? ..... 64
2.1. Definiciones previas ..... 64
2.1.1. Aquello cuyo efecto queremos conocer ..... 64
2.1.2. La variable de resultado ..... 65
2.2. Efecto causal ..... 66
2.2.1. Efecto total o de forma reducida, efecto directo o estructural y mecanismos que producen esos efectos ..... 66
3. Principales dificultades para conocer el efecto causal ..... 68
3.1. No se observa la situación contrafáctica ..... 68
3.2. Sesgo de selección ..... 69
4. Ensayos o experimentos aleatorios controlados (*randomized control trials*) 71
4.1. Limitaciones en la implementación práctica ..... 72
4.2. Problemas de validez externa del experimento ..... 73
5. Controlar por variables observables: regresión con controles, método de emparejamiento (*propensity score matching*) ..... 75
5.1. Regresión con controles ..... 76
5.2. Método de emparejamiento basado en la probabilidad de tratamiento (*propensity score matching*) ..... 78
5.3. Comparación entre la regresión y el método de emparejamiento. 79
5.4. Cuando no se cumple el supuesto de selección sólo en observables. 80
6. Método de diferencias en diferencias (DiD) ..... 80
6.1. DiD básico y supuesto de identificación ..... 81
6.2. Extensiones del método de DiD básico ..... 82
6.2.1. Varios periodos: efectos distintos en el tiempo y distintas tendencias ..... 82
6.2.2. DiD controlando además por diferencias observables ..... 82
6.3. Importancia de las pruebas de falsificación y placebos ..... 83
7. Variables instrumentales ..... 84
8. Método de regresión discontinua (RDD) ..... 88
9. Modelos estructurales: investigando los mecanismos que explican los efectos totales de una política ..... 95
10. Efectos heterogéneos ..... 97
11. Expresiones formales de algunos de los parámetros y estimadores ..... 99
11.1. Resultados potenciales, efecto causal, sesgo de selección y ensayos. 99
11.2. Selección sólo en observables ..... 99

11.3. DiD básico y supuesto de identificación ..... 100
11.4. Variables instrumentales: estimador de Wald, Supuestos y LATE.. 101
11.5. Regresión Discontinua..... 101
Bibliografía..... 102

Capítulo 3a. EL IMPACTO DEL TAMAÑO DE LA CLASE EN EL RENDIMIENTO ACADÉMICO..... 107
Resumen ..... 107
1. Introducción..... 107
2. La asociación entre el tamaño de clase y el rendimiento educativo: la primera generación de artículos no causales ..... 112
3. El impacto causal del tamaño de las clases en el rendimiento educativo 116
3.1. Experimentos aleatorios controlados..... 116
3.2. Regresiones de discontinuidad ..... 119
3.3. Variables instrumentales..... 122
3.4. Leyes de reducción del tamaño de la clase..... 127
3.5. Comparación de ratios de los alumnos por clase y rendimiento por país a partir de PISA, TIMSS o PIRLS..... 130
4. Estimaciones empíricas del impacto de la reducción del tamaño de la clase. Efectos en el largo plazo y heterogeneidad ..... 132
5. Tutorías en pequeños grupos ..... 140
6. Conclusiones ..... 143
Bibliografía..... 145

Capítulo 3b. EL IMPACTO DE LA FAMILIA EN LA EDUCACIÓN . 153
Resumen ..... 153
1. Introducción..... 153
2. Evidencia sobre los padres y su influencia en la educación..... 154
3. Intervenciones de política para aumentar la implicación parental..... 157
4. Estilos parentales y educación: Doepke y Zilibotti, 2017..... 160
4.1. La decisión del estudiante..... 161
4.2. La elección entre persuasión y permisividad..... 162
4.3. La elección entre autoritarismo y permisividad ..... 164

4.4. Evidencia sobre estilos culturales y relación con los entornos económicos ..... 165
5. Transmisión cultural: Bisin y Verdier, 2000 y 2001 ..... 167
5.1. Culturas de capital humano y evolución cultural ..... 168
5.2. La decisión del estudiante ..... 168
5.3. Decisiones de la familia ..... 168
5.4. Evidencia sobre transmisión cultural ..... 170
6. Familias y recursos: Albornoz et al., 2018 ..... 170
6.1. Utilidad de los estudiantes, familias y profesores ..... 171
6.2. La estructura y equilibrio del juego de políticas educativas ..... 173
6.3. Estática comparada e implicaciones de política del modelo ..... 174
6.4. Evidencia sobre familias y recursos educativos ..... 178
7. Evidencia adicional sobre familias y educación ..... 180
Bibliografía ..... 181

Capítulo 3c. EL EFECTO DE LOS COMPAÑEROS DE CLASE (*PEER EFFECTS*) ..... 187
1. Introducción ..... 187
2. ¿Cómo estimar el efecto de los compañeros de clase? ..... 190
3. Métodos para estimar el efecto de los compañeros de clase ..... 194
3.1. Variación no experimental ..... 194
3.1.1. Movimientos exógenos de personas ..... 194
3.1.2. Variación aleatoria entre cohortes ..... 195
3.1.3. Asignación aleatoria de compañeros de cuarto ..... 196
3.2. Variación experimental ..... 197
3.2.1. Experimentos aleatorios ..... 197
3.2.2. Método de tratamiento parcial (*Partial Population Approach*) ... 198
3.3. Reconciliando la variación natural y experimental ..... 199
4. ¿Qué conocemos sobre el efecto de los compañeros de clase en educación? 201
4.1. Rendimiento académico ..... 201
4.1.1. Compañeros de clase ..... 201
4.1.2. Familia ..... 203
4.1.3. Vecinos ..... 204
4.2. Trayectoria educativa ..... 204
4.2.1. Compañeros de clase ..... 205

4.2.2. Familia ..... 206
4.2.3. Vecinos ..... 206
4.3. Resultados no académicos ..... 207
4.3.1. Compañeros de clase ..... 207
4.3.2. Familia ..... 208
4.3.3. Vecinos ..... 209
5. Conclusiones y discusión ..... 210
Bibliografía ..... 213

Capítulo 3d. PROFESORES: MUCHO MÁS QUE UNA PROFESIÓN 219
Resumen ..... 219
1. Introducción: la profesión más importante del mundo ..... 219
2. Profesorado y capital humano ..... 223
2.1. El profesorado desde una perspectiva del capital humano ..... 223
2.2. ¿Por qué el profesorado es tan importante para el desempeño de los alumnos? ..... 226
2.3. Profesores, escuelas y su impacto más allá del aula ..... 229
3. Profesorado: quiénes son, dónde están y el papel de las políticas públicas
3.1. ¿Cómo identificar un buen profesor? ..... 231
3.2. ¿Qué impacto tienen las condiciones laborales de los docentes sobre su efectividad? ..... 233
3.3. El continuo de la carrera profesional docente: desde la selección a la jubilación ..... 237
4. Los equipos directivos y su impacto en la escuela ..... 240
5. Conclusiones ..... 242
Bibliografía ..... 244

Capítulo 3e. LOS CONOCIMIENTOS DE IDIOMAS EXTRANJEROS 251
Resumen ..... 251
1. Introducción ..... 251
2. Los beneficios económicos de conocer otros idiomas ..... 255
2.1. Beneficios económicos a nivel individual ..... 255
2.2. Beneficios económicos a nivel de empresa y de país ..... 257

3. Programas que implementan el uso de un idioma extranjero como idioma de la enseñanza (programas bilingües) en España y en otros países ..... 258
3.1. La educación bilingüe en España ..... 259
3.2. La educación bilingüe en otros países ..... 260
4. La evaluación de los programas de educación bilingüe ..... 263
4.1. Metodología y resultados de la evaluación de un programa de educación bilingüe en España ..... 263
4.1.1. Descripción de los datos y la metodología de estimación ..... 263
4.1.2. Resultados empíricos ..... 267
4.2. Metodología y resultados de evaluaciones de programas de educación bilingüe en otros países ..... 268
4.2.1. Los Países Bajos ..... 268
4.2.2. Estados Unidos ..... 270
5. Otros ejemplos de programas o reformas que afectan el idioma de la enseñanza en las escuelas ..... 272
6. Conclusiones ..... 274
Ejercicios ..... 276
Bibliografía ..... 277

Capítulo 4a. LA EDUCACIÓN INFANTIL ..... 281
Resumen ..... 281
1. Introducción ..... 281
2. La Educación Infantil en España ..... 285
3. Marco conceptual ..... 289
3.1. La función de producción educativa ..... 289
3.2. Determinantes de la inversión de las familias en Educación Infantil ... 294
3.3. Retos de la estimación empírica del impacto de la Educación Infantil . 298
4. Medición del impacto de la Educación Infantil ..... 300
4.1. Evaluaciones experimentales ..... 300
4.2. Evaluaciones cuasiexperimentales ..... 303
5. Conclusiones ..... 306
Bibliografía ..... 308

Capítulo 4b. LA EDUCACIÓN POST-OBLIGATORIA ..... 315
Resumen ..... 315

1. Introducción ..... 316
2. La elección de estudios post-obligatorios en España en los últimos 60 años 317
3. ¿Merece la pena seguir estudiando? ..... 327
3.1. Marco teórico ..... 327
3.2. Marco empírico de la decisión de educarse y los rendimientos de la educación ..... 330
3.2.1. Seguir estudiando durante el boom inmobiliario 1995-2008 335
3.2.2. Evidencia descriptiva de los rendimientos tras la crisis financiera 337
4. ¿FP o Universidad? ..... 342
4.1. Marco teórico y evidencia para otros países ..... 342
4.2. Evidencia descriptiva para España ..... 346
5. ¿FP tradicional o dual? ..... 349
5.1. Marco teórico ..... 351
5.2. Evidencia del impacto en la carrera laboral de la FP tradicional o dual 352
6. Conclusiones ..... 357
Bibliografía ..... 359

Capítulo 4c. EDUCACIÓN TERCIARIA: ACCESO, FINANCIACIÓN Y RENTABILIDAD ..... 365
1. Introducción ..... 365
2. Los mercados de enseñanza superior y la elección educativa ..... 369
2.1. El mercado de Educación Superior: FP *vs* Universidad ..... 370
2.2. Elección de área de estudio ..... 374
2.3. Abandono en la Universidad ..... 377
2.4. Epítome: mercados que no existen pero que deberían existir ..... 379
3. Elitismo, discriminación y acceso ..... 380
4. La rentabilidad privada de la educación universitaria ..... 383
5. La financiación de los estudios universitarios ..... 387
5.1. Elementos redistributivos de la financiación mediante tributación general ..... 394
5.2. Restricciones de crédito ..... 395
5.3. Efectos comportamentales en las decisiones ..... 397
5.4. La economía política de la financiación de las Universidades ..... 398
5.5. La financiación universitaria en España ..... 400

6. El gobierno de la Universidad ..... 401
6.1. La gestión del talento ..... 402
6.2. El talento del gestor ..... 403
Ejercicios ..... 404
Bibliografía ..... 405

Capítulo 5a. LA ELECCIÓN ESCOLAR ..... 415
Resumen ..... 415
1. Introducción ..... 415
2. Determinantes de la demanda de escuelas ..... 417
2.1. Rendimiento absoluto *vs* valor añadido ..... 419
2.2. Multidimensionalidad ..... 422
3. Mecanismos de asignación ..... 424
3.1. Propiedades ..... 427
3.2. Estudios empíricos ..... 429
3.3. Elección escolar y segregación ..... 433
4. La elección universitaria ..... 436
5. Conclusiones ..... 438
Bibliografía ..... 439

Capítulo 5b. DESIGUALDAD EDUCATIVA ..... 441
Resumen ..... 441
1. Introducción ..... 441
2. Medidas de desigualdad educativa ..... 444
3. Evidencia en España ..... 447
3.1. Abandono escolar y características familiares ..... 447
3.2. Abandono escolar y desigualdades regionales ..... 453
3.3. España en perspectiva con otros países europeos ..... 455
3.4. Desigualdad en el acceso a la Educación Universitaria ..... 457
3.5. Desigualdad en resultados escolares ..... 459
3.6. Desigualdad en cuanto a repetición de curso ..... 462
3.7. Desigualdad educativa en puntuaciones dentro y fuera de España. 464
4. Consecuencias futuras de la desigualdad educativa actual ..... 466

4.1. Desigualdad educativa y movilidad social en un contexto de cambio de estructura productiva ..... 467
5. Causas de la desigualdad y posibles soluciones ..... 468
6. Otras desigualdades y comentarios finales ..... 471
Bibliografía ..... 473

Capítulo 5c. BRECHAS DE GÉNERO EN EDUCACIÓN ..... 479
Resumen ..... 479
1. Introducción ..... 480
2. Qué factores explican un mayor porcentaje de fracaso escolar entre los chicos ..... 486
2.1. Habilidades no cognitivas ..... 487
2.2. Familias y factores socioeconómicos ..... 490
2.3. Escuela y profesorado ..... 493
2.4. Condiciones en el mercado laboral y coste de oportunidad de seguir estudiando ..... 496
3. Qué factores explican una menor participación de las mujeres en STEM 498
3.1. Ventaja comparativa ..... 498
3.2. El rol de la cultura y las normas sociales ..... 501
3.3. Estereotipos de género ..... 504
3.4. Ausencia de referentes en STEM ..... 507
3.5. El rol de los contextos competitivos ..... 510
4. Conclusiones ..... 512
Bibliografía ..... 513

Capítulo 5d. HABILIDADES NO COGNITIVAS ..... 517
Resumen ..... 517
1. ¿Qué son las habilidades no cognitivas? ..... 518
2. ¿Cómo se miden las habilidades no cognitivas? ..... 520
2.1. Medición a través de encuestas ..... 521
2.1.1. Medidas de personalidad: los cinco grandes o "Big 5" ..... 521
2.1.2. Medidas de perseverancia a largo plazo o "Grit" ..... 522
2.1.3. Limitaciones de las medidas basadas en encuestas ..... 522
2.2. Medidas directas de comportamiento ..... 523

2.2.1. Usando datos administrativos ..... 523
2.2.2. Medidas de conducta a través de experimentos con tareas directas ..... 524
2.3. Usando metadatos para inferir medidas de habilidades no cognitivas . 525
2.3.1. El poder predictivo del esfuerzo de los estudiantes en encuestas 527
2.3.2. Diferencias internacionales, entre Comunidades Autónomas y por género ..... 528
3. ¡No te comas la golosina! ..... 531
4. Determinantes de las habilidades no cognitivas ..... 533
4.1. Genética ..... 533
4.2. Factores ambientales y renta familiar ..... 534
4.3. Estilos de crianza ..... 535
4.4. Tamaño de la familia ..... 537
4.5. Entorno escolar ..... 538
4.6. Cultura ..... 538
5. Intervenciones eficaces en la mejora de las habilidades no cognitivas .... 539
6. Conclusiones ..... 546
Ejercicios ..... 547
Bibliografía ..... 548

Anexo. ACTUALIZACIÓN DE LOS RESULTADOS DE PISA 2022 .. 551

Epílogo. LA ECONOMÍA DE LA EDUCACIÓN A TRAVÉS DEL LENTE DE ESPAÑA: LECCIONES Y POSIBILIDADES ..... 573

# Prólogo
# EDUCACIÓN, HISTORIA Y ECONOMÍA: PASADO, PRESENTE Y FUTURO DE LA EDUCACIÓN EN LA EUROPA MODERNA[1]

## ¿CUÁNDO SE FASTIDIÓ ESPAÑA, VARGUITAS?

La educativa es una de las cuatro revoluciones que ocurren en Europa alrededor de mediados del siglo XVIII[2]. Las otras son la científica, la política y la industrial. Son revoluciones que cambian el mundo para siempre y aseguran la supremacía europea en los próximos dos siglos[3]. Hasta entonces, las principales variables con las que los economistas medimos los resultados, la renta por habitante y la esperanza de vida son, a largo plazo, estacionarias[4]. Los imperios transoceánicos que se crean en Europa a partir del siglo XVI, en Portugal, España, Francia e Inglaterra, si bien enriquecen a las élites, son fundamentalmente incapaces de cambiar el nivel de vida de los ciudadanos, así como los usos y costumbres generales[5].

A principios del siglo XIX, ya con las antes citadas revoluciones en marcha, España todavía no era muy diferente de Francia y el Reino Unido. Angus Madison estimó que nuestro PIB per cápita en 1820 era un 90% del de Francia. Pero un

---

[1] Agradezco los comentarios de Carmen Beviá, Antonio Cabrales e Ismael Sanz a una versión preliminar de este texto.

[2] No es fácil resumir la historia de la educación en Europa en los últimos tres siglos. Yo lo he intentado en varios de mis artículos en *Nada es Gratis* que podéis encontrar en mi página web en la sección de "Divulgación" https://www.eco.uc3m.es/personal/cv/lcorchon.html

[3] Especialmente sobre Asia que hasta entonces había estado a la par o por encima de Europa en cualquier medida del funcionamiento de sus economías. De ahí el término "La Gran Divergencia" acuñado por Pomeranz (ver Broadberry y Zhai, 2023).

[4] Para el caso de España, la figura 1 del artículo de Prados de la Escosura y Sánchez Alonso (2023) es esclarecedora. Es probable que la esperanza de vida de las élites empezara a aumentar un siglo antes y que eso fuera una de las causas de las revoluciones posteriores (ver De la Croix y Licandro, 2015).

[5] Para una comparación entre estos imperios recomiendo el libro de María Elvira Roca Barea (2016).

siglo más tarde, en 1913, cuando estas revoluciones ya habían dado sus frutos, España era un país atrasado en donde su PIB per cápita era el 55% de Francia. Para mí no hay duda de que la causa de este atraso se debe a que perdemos los cuatro trenes antes mencionados[6].

Por centrarnos en nuestro tema, la educación obligatoria hasta los 14 años, instaurada en Prusia a mediados del siglo XVIII, llega a España en 1970 de la mano de Villar Palasí. La Universidad Humboldt –creada en Prusia en 1810 pero anticipada en Rusia en 1755 por Isabel I y Lomonosov– donde, debido al progreso continuo de la ciencia los profesores han de ser investigadores de primera línea, no llega a España antes de 1975.

**Educación y desarrollo**

Algunos siguen pensando que la riqueza de los Estados Unidos de América se basa en sus recursos naturales. Sin duda estos recursos jugaron un papel importante en las primeras décadas de desarrollo de esta nación. Pero ya hacia la mitad del siglo pasado, Leontief (y otros posteriormente) encontraron que las exportaciones americanas eran más intensas en trabajo que en capital. Esta es la llamada "Paradoja de Leontief", cuya resolución más convincente es que no se había tenido en cuenta el capital humano incorporado en tales exportaciones.

En España, el papel del capital humano ha sido más modesto. Como dicen Prados de la Escosura y Sánchez-Alonso (2023) "el crecimiento a largo plazo de la productividad laboral se explica, en proporciones semejantes, por la acumulación de capital físico y, en menor medida, humano, y las ganancias de eficiencia". Para mí este modesto papel de la educación puede ayudar a entender por qué nunca hemos logrado superar el 90% del PIB por habitante de las naciones desarrolladas (ver De La Fuente y Domenech, 2023). Porque una vez que se agotan las ganancias de la desregulación, es necesario innovar para continuar la escalada y ahí pinchamos en hueso.

---

[6] https://en.wikipedia.org/wiki/List_of_regions_by_past_GDP_(PPP)_per_capita. El papel de la "acumulación primitiva" europea en su desarrollo posterior es todavía un tema de debate ver https://en.wikipedia.org/wiki/Theories_of_imperialism. La descripción que hace Paul Baran (1957) de la entrada en Inglaterra de las riquezas coloniales es inolvidable.

## España hoy

Yo he visto cosas que vosotros no creeríais. Una España paupérrima que se ha transformado en un país del primer mundo. A pesar de que recibí la que posiblemente era una de las mejores educaciones universitarias que España podía ofrecer a finales de los 60 y principios de los 70, ninguno de mis profesores se acercó jamás a tener algo parecido a una publicación internacional. Hoy o tienes cuatro artículos en revistas internacionales de prestigio en los primeros siete años de tu carrera o no vas a ser profesor estable en ningún buen departamento de economía en España.

Y he visto nacer la ciencia económica transformándose desde una aplicación de las Matemáticas (micro) y una peli de buenos y malos (macro) en un conocimiento con sólidas bases teóricas, empíricas y experimentales. Y con aplicaciones que pueden mejorar drásticamente el nivel de vida de muchos terrícolas. De hecho, podría significar la nueva revolución por su impacto en nuestras condiciones de vida similar al que tuvieron y tienen la Ingeniería, la Medicina y las Ciencias Biológicas. Sólo hace falta que la sociedad despierte, se tome un poco de trabajo en aprender a juzgar a los académicos –que en todo caso requiere menos trabajo que el necesario para juzgar a los deportistas, cocineros, etc., a los que tanto apreciamos– los apoye y deje de escuchar a los charlatanes. Quizá sea este el momento apropiado de introducir a los autores del libro.

Antonio Cabrales, PhD por la Universidad de California ha sido catedrático de la Pompeu Fabra y el University College y lo es de la Carlos III. Jesús Carro es doctor por el CEMFI y profesor titular de la Universidad Carlos III de Madrid. Caterina Calsamiglia es PhD por la Universidad de Yale y catedrática de la Universidad Autónoma de Barcelona. José García Montalvo es PhD por la Universidad de Harvard y es catedrático de la Universidad Pompeu Fabra. Ismael Sanz es doctor por la Universidad Complutense y profesor titular de la Universidad Rey Juan Carlos. Todos poseen publicaciones en revistas de alto impacto internacional ampliamente citadas como puede comprobar el lector tecleando su nombre y a continuación la palabra "citations". Varias de esas publicaciones se refieren al tema educativo. Luego haga lo mismo con alguno de nuestros economistas que pasan por genios y ya verá lo que pasa...

## El futuro

A pesar de todos nuestros avances, la educación sigue siendo una de nuestras asignaturas pendientes como se demuestra cada vez que aparecen los resultados

PISA. Este libro puede ayudar a que de una vez nos encaminemos a la resolución de este secular problema.

El libro que ahora prologo, expone en primer lugar de manera amena en qué consisten los principales métodos empíricos que se usan en el seguimiento de las políticas educativas. Y a continuación se zambulle en el estudio científico de temas como el efecto del tamaño de la clase, la familia, los profesores y los compañeros sobre los resultados educativos. Y da sabrosos consejos sobre cómo mejorar la gestión escolar y administrativa. Luego, el libro se adentra en el estudio de temas más específicos como el bilingüismo y la importancia de las escuelas infantiles. Y lo que ocurre más allá del instituto, o sea la FP y la Universidad. Por último, se centra en temas candentes como el del género y el origen de las habilidades intelectuales. El libro da evidencias del éxito o fracaso de las políticas educativas que se han seguido en el mundo. Y sugiere la implementación de algunas de ellas. Es de resaltar que esa implementación no requiere ni una nueva ley de educación ni un coste exagerado.

Por sus conclusiones y sobre todo por el método usado para alcanzarlas este libro debería estar en la mesa de todos los gestores relacionados con la educación. Para que se eviten errores que pagan inocentes y se mejore lo que ya se conoce. Y que España no desperdicie más talento. Que así sea.

## BIBLIOGRAFÍA


Baran, Paul A. (1957). *The Political Economy of Growth*. New York: Monthly Review Press.

Broadberry, S. y R. Zhai (2023). *Innovation and the great divergence*. https://economics.yale.edu/sites/default/files/2023-12/Innovation%20and%20the%20Great%20Divergence%20v3.pdf

De la Croix, D. y O. Licandro (2015). *The longevity of famous people from Hammurabi to Einstein*. Journal of Economic Growth.

De la Fuente, A. y R. Domènech (2023). *Renta per cápita y productividad en la OCDE de 1960 a 2022*. Fedea, Documento de Trabajo 2023/09, noviembre de 2023.

Prados de la Escosura, L. y B. Sánchez Alonso (2023). *El crecimiento económico moderno y su distribución en España. Información Comercial Española*.

Octubre-Noviembre-Diciembre 2023. Nº 933 (2023). https://www.revistasice.com/index.php/ICE/article/view/7683/7761

Roca Barea, Elvira (2016). *Imperiofobia y leyenda negra: Roma, Rusia, Estados Unidos y el Imperio español.* Siruela biblioteca de ensayo.

Luis Corchón
Catedrático emérito
Universidad Carlos III de Madrid

# Capítulo 1
# INDICADORES DEL SISTEMA EDUCATIVO ESPAÑOL

Ismael Sanz
Departamento de Economía Aplicada I e Historia e Instituciones Económicas
Universidad Rey Juan Carlos
Department of Social Policy
London School of Economics
ORCID: https://orcid.org/0000-0003-1286-4124

**Resumen**

El presente capítulo de Indicadores del sistema educativo español tiene cuatro apartados, siguiendo la estructura de Indicadores del Sistema Estatal de Indicadores de la Educación del Ministerio de Educación y Formación Profesional. Se comienza el análisis por un primer apartado con la información relativa a alumnado, incluyendo las tasas de matriculación, y el profesorado, con información relativa a las condiciones profesionales de los docentes y el entorno de aprendizaje. En este primer apartado se centra la atención en el acceso a las etapas no obligatorias como Infantil, FP, Bachillerato y Universidad. En el segundo apartado se examinan los indicadores de financiación educativa española en relación con la UE y la OCDE, con subapartados dedicados a gasto total, gasto por alumno y becas y ayudas. El tercer apartado, analiza indicadores de resultados muy relevantes como las tasas de abandono temprano educativas, las de graduación y las de repetición. En esta misma sección también se revisan otros resultados del sistema educativo de nuestro país, como las puntuaciones logradas por el alumnado de Primaria en las pruebas internacionales de Primaria de Lectura (PIRLS), Matemáticas y Ciencias (TIMSS), así como nivel de los estudiantes de Secundaria en la evaluación educativa más conocida de PISA. Este capítulo finaliza con un último apartado dedicado la relación de la educación y el empleo, mostrando la asociación entre el nivel educativo alcanzado y la tasa de paro e ingresos salariales. Además, se explora con mayor detalle la fase de la transición de la formación al empleo.

# 1. INDICADORES DE ESCOLARIZACIÓN Y ENTORNO EDUCATIVO EN ESPAÑA

## 1.1. Alumnado, tasas de matriculación

### *1.1.1. Educación Infantil de primer y segundo ciclo*

La Educación Infantil y el cuidado de la primera infancia tiene una incidencia muy relevante en el desarrollo cognitivo y emocional, el aprendizaje y el bienestar del alumnado (Duncan et al., 2023). Los niños que participan desde una edad temprana en programas de aprendizaje de alta calidad tienen más probabilidades de obtener mejores resultados educativos de mayores. El impacto positivo de la asistencia a la Educación Infantil es particularmente importante para los niños provenientes de familias desfavorecidas, porque a menudo tienen menos oportunidades de desarrollar estas habilidades en sus entornos de aprendizaje en casa (véase el capítulo 4.a sobre Educación Infantil). Por lo tanto, ampliar el acceso a la Educación Infantil tiene el potencial de reducir la desigualdad de oportunidades en educación y formación. La extensión de la educación y cuidado en la primera infancia también facilita que los padres puedan trabajar y contribuir al bienestar y crecimiento económico.

La tabla 1 muestra que el 27% de los niños menores de tres años de la OCDE están matriculados en Educación Infantil de primer ciclo (0-2 años). Un 83% de los niños entre las edades de 3 y 5 años de la OCDE están matriculados en Educación Infantil temprana (segundo ciclo). La tasa de escolarización de España en el primer ciclo de Educación Infantil crece 26 puntos porcentuales desde el año 2005 y alcanza el 41,1%, superior a las medias de la OCDE (26,8%) y de la UE-22 (22,0%). España ha pasado del 34,0% en 2015 al 41,1% en 2020, lo que supone un crecimiento de 7,1 puntos porcentuales, mientras que la media de la OCDE crece 2,6 puntos (24,2% en 2015) y la media de la UE-22 solo 0,4 puntos (OCDE: Education at a Glance, 2022). De hecho, España es uno de los países en donde más ha crecido la matriculación en Educación Infantil de primer ciclo (0-2 años). La expansión de la Educación Infantil se generaliza de manera internacional para los años del segundo ciclo de Educación Infantil (3-5). En la OCDE, un 87,4% de los niños y de las niñas entre los 3 y los 5 años está escolarizado. Este porcentaje asciende al 90,5% en la UE-22. En España, entre los 3 y los 5 años, la escolarización es prácticamente total, pues supera el 97%.

Entre 2005 y 2020, la tasa de escolarización para las edades de 3-5 años no ha sufrido una variación significativa en España, pues ya en 2005 superaba el 97. En la OCDE, el crecimiento del segundo ciclo de Educación Infantil es de cerca de 9 puntos porcentuales y en la media de la UE-22, de casi 10 puntos porcentuales. Los países que se utilizan a lo largo del capítulo como referencia para España (Francia, Alemania, Reino Unido, Estados Unidos, promedio de OCDE y UE, y cuando hay información, Italia) son porque son los más cercanos a España o/y los más grandes de la UE u OCDE.

**Tabla 1. Tendencias en las tasas de matriculación en educación y cuidado de la primera infancia, por grupo de edad (2005, 2015 y 2020), en instituciones públicas y privadas.**

| | Edad en la que comienzan los servicios de Educación Infantil | Inicio típico de Educación Primaria | Inicio de la edu-cación obligatoria | Educación Infantil menos 0-2 años | | | Educación Infantil menos 3-5 años | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | 2005 | 2015 | 2020 | 2005 | 2015 | 2020 |
| | | 2020 | | 2005 | 2015 | 2020 | 2005 | 2015 | 2020 |
| Francia | 2-3 años | 6 | 3 | 9 | 4 | 4 | 100 | 100 | 100 |
| Alemania | 0 años | 6 | 6 | 17 | 37 | 39 | 87 | 96 | 94 |
| España | 0 años | 6 | 6 | 15 | 34 | 41 | 98 | 97 | 97 |
| Reino Unido | 0 años | 5 | 4-5 | – | – | 19 | – | 68 | 72 |
| Estados Unidos | | 6 | 4-6 | | | | 64 | 65 | 64 |
| **OCDE** | | | | **–** | **24** | **27** | **74** | **82** | **83** |
| **UE-22** | | | | **–** | **22** | **22** | **78** | **87** | **88** |

Nota: Educación Infantil temprana = ISCED 0. Los servicios educación temprana deben: 1) Objetivos educativos adecuados; 2) Proporcionarse en instituciones (escuelas o instituciones análogas para un grupo de niños); 3) Intensidad de, al menos, 2 horas por día de actividades educativas y 100 días al año; 4) Marco regulatorio o currículo reconocido por las autoridades nacionales; y 5) Contar con personal con acreditación pedagógica.
Fuente: Education at a Glance, 2022 (p. 163). https://stat.link/z5os1d

En la UE, la expansión de la Educación Infantil ha sido el resultado de un estímulo adicional de los objetivos establecidos en el Consejo Europeo de Barcelona en 2002 para proporcionar plazas subvencionadas a tiempo completo para un tercio de los niños menores de 3 años en 2010 (Comisión Europea:

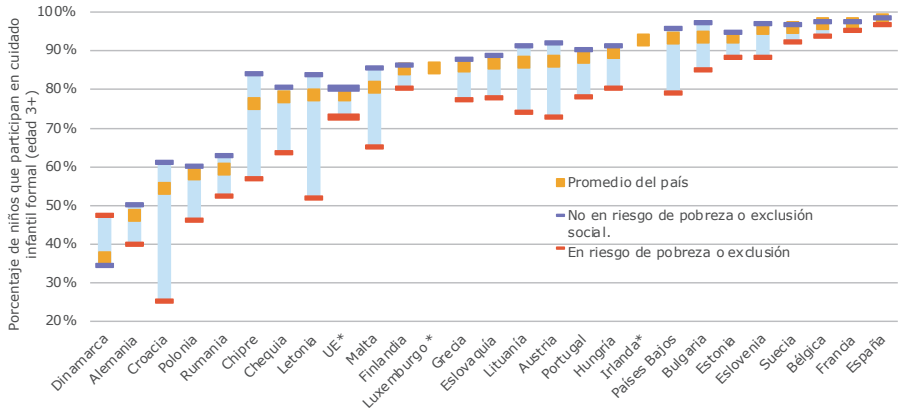
Education and Training Monitor, 2022). El objetivo a nivel de la UE es que, para 2030, al menos el 50% de los niños menores de 3 años participen en cuidado infantil formal; y al menos el 96% de los niños entre los 3 años y la edad de inicio de la Educación Primaria obligatoria participen en Educación Infantil. Para este objetivo a nivel de la UE, se considera que los niños participan en Infantil solo si están inscritos en programas que se consideran educativos según la Clasificación Internacional Normalizada de la Educación (nivel 0/Educación Infantil temprana), es decir, diseñados intencionadamente para apoyar el desarrollo cognitivo, físico y socioemocional de los niños (Comisión Europea: Education and Training Monitor, 2022). La propuesta de la Comisión de 2022 es aumentar la participación en Educación Infantil, en particular para los niños en situaciones vulnerables o de entornos desfavorecidos.

No cabe duda de que la calidad del programa es de suma importancia para su éxito. El Proyecto Preescolar Perry y el Programa Abecedarian en EE. UU., que se analizarán con detalle en los capítulos 3a y 4a sobre el impacto del tamaño de la clase y sobre Educación Infantil, por ejemplo, contrataron maestros con títulos universitarios y muy cualificados, con una ratio alumnos por personal bajo. Heckman et al. (2010) señalan que los programas de Educación Infantil temprana que toman como modelo programas como Perry pueden ofrecer un retorno de inversión anual del 7%-10%, lo cual es comparable al retorno sobre el capital después de la Segunda Guerra Mundial en las acciones de EE. UU.

A pesar de los esfuerzos por aumentar el acceso al Educación Infantil, la probabilidad de participación sigue dependiendo en gran medida del contexto familiar y del ingreso, especialmente en los servicios de desarrollo educativo temprano que dependen en gran medida de fuentes privadas de financiación. Los datos de la Encuesta de Estadísticas de Ingresos y Condiciones de Vida de la Unión Europea (EU-SILC) revelan que existen diferencias estadísticamente significativas en la participación en Educación Infantil en la mitad de todos los países de la UE para los niños que experimentan desventajas socioeconómicas en términos de ingresos del hogar, educación materna, riesgo de pobreza o exclusión social (Flisi y Blasko, 2019). En promedio, en los países europeos de la OCDE, los niños de 0-2 años procedentes de hogares de bajos ingresos tenían en 2017 aproximadamente un tercio menos de probabilidades de asistir a Educación Infantil (que incluye centros, hogares y cuidado infantil familiar organizado) en comparación con aquellos de hogares con altos ingresos (Comisión Europea: Education and Training Monitor de 2022 y OCDE 2020a). En casi todos los

países de la UE, la proporción de niños en riesgo de pobreza o exclusión social que están en cuidado infantil formal es consistentemente menor que la proporción entre los niños que no están en riesgo (gráfico 1). La brecha de participación es del 7,5% en promedio, y varía desde 1,5 puntos porcentuales en España hasta no menos de 35,7 puntos porcentuales en Croacia.

**Gráfico 1. Los niños en riesgo de pobreza o exclusión social tienen menos probabilidades de participar en cuidado infantil formal (3-5 años).**

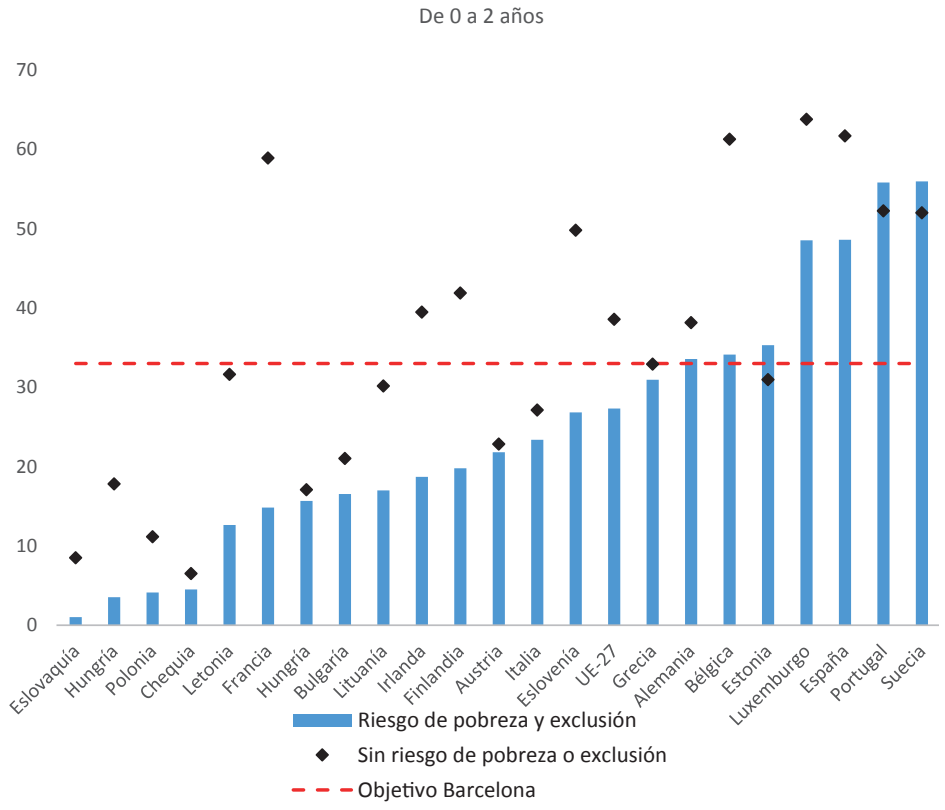



Fuente: Gráfico extraído de Comisión Europea (2022): Education and Training Monitor, 2022. Base de datos de Eurostat (EU-SILC 2020).
Nota: *Desglose no disponible para Luxemburgo e Irlanda; estimaciones no fiables para la categoría AROPE en Dinamarca, Chipre, Malta, Letonia y Países Bajos; no hay datos de EU-SILC sobre participación en cuidado infantil disponibles para Italia; Irlanda, Italia y Luxemburgo no están incluidos en el promedio de la UE; el límite superior del rango de edad está definido por la edad de inicio de la Educación Primaria obligatoria en cada país.

Las diferencias en España en la matriculación en Educación Infantil de 0-2 años entre niños con riesgo de pobreza o exclusión social si es significativa más baja que entre los que no tienen ese riesgo. En concreto, el 48,6% de los niños en riesgo de pobreza infantil asisten a Educación Infantil de primer ciclo por el 61,7% de los que no están en esa situación. En la UE, la diferencia es similar, 27,3% frente al 38,6%. Los datos correspondientes al gráfico 1 y gráfico 2 son más elevados que los de la tabla 1 porque en la recopilación de información de la OCDE se refleja la asistencia de los alumnos en programas con un componente educativo, mientras que la encuesta EU-SILC registra la participación en cuidado infantil formal, independientemente de un posible componente educativo en el programa. Además, dependiendo del momento de la recopilación de datos

EU-SILC 2020, los datos sobre la participación en cuidado infantil formal pueden verse afectados por los cierres de escuelas debido a Covid-19 en algunos países, una circunstancia que no ocurre en la información de la OCDE. En definitiva, los alumnos de entornos desfavorecidos, que son a los que más le beneficia la asistencia a Educación Infantil, son los que menos se matriculan tanto en España como en el conjunto de la UE.

**Gráfico 2. Los niños en riesgo de pobreza o exclusión social tienen menos probabilidades de participar en cuidado infantil formal (3-5 años).**




Fuente: Gráfico extraido de Comisión Europea: Employment and Social Developments in Europe, 2022.

### *1.1.2. El descenso del alumnado matriculado en segundo ciclo de Infantil y Primaria*

El descenso en la natalidad en España está comenzando a surtir efecto en las aulas educativas de nuestro país. La tabla 2 muestra que en España se está produciendo un descenso en el alumnado en el segundo ciclo de Educación Infantil desde el máximo alcanzado en el curso 2011-2012 de 1.470.717. Desde entonces, con la única excepción del curso 2019-2020, ha encadenado 11 cursos seguidos de disminución en matriculados hasta llegar a los 1.148.901 del curso 2022-2023, una caída de 321.816 alumnos (–21,9%). Esta disminución en el número de estudiantes se trasladó años más tarde al siguiente nivel educativo, el de Primaria. Desde el curso 2017-2018, el alumnado de Primaria en España ha bajado desde los 2.942.894 hasta los 2.783.606 del último curso. En total, se han perdido 159.288 estudiantes en Primaria en nuestro país en los últimos cinco años, un –5,4%. La cifra de 2022-2023 de matrícula en España es aún más elevada que el mínimo registrado en 2004-2005, cuando todavía no habían llegado a las aulas de Primaria el flujo de los hijos de inmigrantes que entraron en nuestro país. Por el contrario, en la ESO el número de estudiantes continúa subiendo en los últimos años en España, con un máximo histórico de 2.085.482 en el curso 2022-2023. En este nivel educativo de la ESO no ha llegado todavía el descenso de la natalidad que comenzó a afectar la matriculación en Primaria hace cinco años y que, por lo tanto, podría empezar a reducir el número de estudiantes en la ESO a partir de 2023-2024. Además, hay un segundo efecto que puede contrarrestar en alguna medida la posible futura caída de alumnado en la ESO, y es el hecho de que actualmente los estudiantes españoles abandonan menos los estudios que en el pasado, cuando algunos jóvenes dejaban Secundaria sin terminarla y sin llegar en algunos casos a 4º de la ESO. Ahora la inmensa mayoría de los alumnos españoles alcanza 4º de la ESO como se demuestra en la creciente tasa bruta de graduación que veremos a continuación. La continuación de la tendencia de reducción de alumnos que no finaliza los estudios obligatorios en España podría mitigar la disminución del alumnado matriculado en la ESO en los próximos años. De hecho, los últimos datos del Ministerio de Educación y FP muestran que en 2023-2024 continuará el descenso de alumnado, tanto en el segundo ciclo de infantil (–3,4%), así como en Primaria (–1,3%) (Datos y cifras. Curso escolar 2023-2024, Ministerio de Educación y FP, 2023b).

**Tabla 2. Evolución del alumnado matriculado en segundo ciclo de Educación Infantil, Primaria y ESO. España 1990-1991 hasta 2022-2023.**

| Curso | Segundo ciclo Infantil | Primaria | ESO |
|---|---|---|---|
| 2022-2023 (*) | 1.148.901 | 2.783.606 | 2.085.482 |
| 2021-2022 | 1.189.962 | 2.798.301 | 2.051.158 |
| 2020-2021 | 1.233.124 | 2.843.852 | 2.041.117 |
| 2019-2020 | 1.279.839 | 2.907.214 | 2.012.829 |
| 2018-2019 | 1.277.954 | 2.938.361 | 1.975.403 |
| 2017-2018 | 1.293.972 | 2.942.894 | 1.931.886 |
| 2016-2017 | 1.318.986 | 2.941.363 | 1.887.027 |
| 2015-2016 | 1.363.060 | 2.926.887 | 1.869.283 |
| 2014-2015 | 1.395.756 | 2.908.538 | 1.840.748 |
| 2013-2014 | 1.440.474 | 2.858.130 | 1.820.470 |
| 2012-2013 | 1.462.918 | 2.828.445 | 1.808.502 |
| 2011-2012 | 1.470.717 | 2.797.804 | 1.792.548 |
| 2010-2011 | 1.440.607 | 2.754.986 | 1.786.754 |
| 2009-2010 | 1.420.927 | 2.702.399 | 1.792.789 |
| 2008-2009 | 1.401.184 | 2.665.161 | 1.813.572 |
| 2007-2008 | 1.357.019 | 2.607.384 | 1.829.874 |
| 2006-2007 | 1.303.653 | 2.538.033 | 1.834.685 |
| 2005-2006 | 1.260.892 | 2.483.364 | 1.844.953 |
| 2004-2005 | 1.228.178 | 2.467.636 | 1.855.020 |
| 2003-2004 | 1.187.138 | 2.479.631 | 1.871.430 |
| 2002-2003 | 1.139.480 | 2.474.287 | 1.878.175 |
| 2001-2002 | 1.097.027 | 2.474.261 | 1.897.912 |
| 2000-2001 | 1.066.340 | 2.491.648 | 1.941.623 |
| 1999-2000 | 1.044.727 | 2.524.768 | 1.999.567 |
| 1998-1999 | 1.049.357 | 2.562.785 | 1.890.004 |
| 1997-1998 | 1.050.375 | 2.615.467 | 1.686.652 |
| 1996-1097 | 1.047.133 | 2.682.894 | 1.181.466 |
| 1995-1996 | 1.041.345 | 2.779.238 | 457.386 |
| 1994-1995 | 1.030.776 | 2.922.615 | 282.837 |
| 1993-1994 | 1.022.656 | 3.051.103 | 180.352 |
| 1992-1993 | 1.005.075 | 3.194.967 | 95.297 |
| 1991-1992 | 988.909 | 3.333.710 | – |
| 1990-1991 | 975.850 | 3.526.294 | – |

(*) Datos Avance.
Fuente: Estadística de las Enseñanzas no universitarias. Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional.
https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado/series.html

### *1.1.3. La expansión de la FP Media en España y la estabilidad de Bachillerato*

Al finalizar la ESO, los jóvenes pueden continuar con sus estudios realizando Bachillerato o eligiendo algún Ciclo Formativo de Grado Medio (CFGM). Desde el curso 2014-2015, también existe una nueva modalidad de FP, la FP Básica, que es un itinerario para los estudiantes de 2º o 3º de la ESO con problemas de seguimiento de los estudios por problemas de comportamiento o por ser alumnos con necesidades educativas especiales. La tabla 3 muestra que la matrícula de Bachillerato ha sido muy estable en los últimos 10 años, con un número total de estudiantes de algo menos de 700.000. También los CFGM han estado cinco años, desde 2013-2014 a 2018-2019, en el entorno de un alumnado de 350.000 estudiantes. La proporción, por tanto, de estudiantes que accede a Bachillerato era de aproximadamente dos tercios por un tercio en FP Media. Sin embargo, desde 2019-2020, los CFGM han experimentado un crecimiento muy importante en su matrícula, hasta alcanzar 437.558, con un crecimiento de casi 100.000 alumnos en cuatro años, algunos de ellos en la FP Media a distancia. Durante los cursos 2007-2008 hasta 2014-2015, hubo un notable aumento en las matrículas de CFGM. Si se analiza la información del número de matriculados en FP de Grado Medio en relación con las cohortes de población con las edades teóricas de realización de estos estudios (16 y 17 años), la tasa neta de matriculación se observaría que esta tasa ha estado estancada en un bajo 7% (Fundación Europea Sociedad y Educación, 2023). Sin embargo, desde los últimos años la tasa neta de matriculación ha experimentado un incremento de aproximadamente un punto por año, llegando al 12,5% en el curso 2021-2022, por el aumento de matriculación de jóvenes recién titulados en la ESO (véase capítulo 4b sobre Formación Profesional en este libro).

La tabla 3 también muestra que el aumento en la demanda de los CFGM se ha producido sin tener que disminuir la Bachillerato, que se ha mantenido estable. No hay un número excesivo de estudiantes en Bachillerato en España. La escasa matrícula que había en FP Media en España hasta ahora no era porque muchos jóvenes se decantaran por el programa más general, sino porque había un porcentaje importante de jóvenes que no finalizaba ESO o si lo terminaba no continuaba con sus estudios. El Informe anual de la OCDE de Education at a Glance, en su versión de 2021, muestra que el porcentaje de mujeres en FP Media en España es del 49,7%, cuando la media de la OCDE es del 45,4%. La proporción de mujeres en Bachillerato en España (55,5%) sí es similar al de la OCDE (54,5%). La conclusión es que faltan hombres en FP Media, los que dejan los estudios después de la ESO.

**Tabla 3. Evolución del alumnado matriculado en Bachillerato, FP Básica y FP Media. España 1990-1991 hasta 2022-2023.**

| Curso | Bachillerato | FP Básica | FP Media |
|---|---|---|---|
| 2022-2023 (*) | 687.511 | 78.371 | 437.558 |
| 2021-2022 | 690.481 | 75.276 | 420.227 |
| 2020-2021 | 687.084 | 75.952 | 401.066 |
| 2019-2020 | 673.740 | 76.440 | 368.359 |
| 2018-2019 | 667.287 | 73.810 | 350.220 |
| 2017-2018 | 676.311 | 72.180 | 344.266 |
| 2016-2017 | 687.595 | 69.528 | 343.920 |
| 2015-2016 | 695.557 | 61.909 | 349.631 |
| 2014-2015 | 690.738 | 39.867 | 355.120 |
| 2013-2014 | 696.648 | – | 350.250 |
| 2012-2013 | 692.098 | – | 332.495 |
| 2011-2012 | 684.176 | – | 312.016 |
| 2010-2011 | 672.213 | – | 297.877 |
| 2009-2010 | 650.563 | – | 277.925 |
| 2008-2009 | 629.247 | – | 253.516 |
| 2007-2008 | 622.133 | – | 239.559 |
| 2006-2007 | 630.349 | – | 235.224 |
| 2005-2006 | 640.028 | – | 232.149 |
| 2004-2005 | 646.174 | – | 233.513 |
| 2003-2004 | 657.400 | – | 230.688 |
| 2002-2003 | 679.773 | – | 225.479 |
| 2001-2002 | 714.390 | – | 211.622 |
| 2000-2001 | 766.518 | – | 192.011 |
| 1999-2000 | 797.924 | – | 158.906 |
| 1998-1999 | 964.163 | – | 119.809 |
| 1997-1998 | 1.195.500 | – | 75.916 |
| 1996-1097 | 1.309.797 | – | 48.812 |
| 1995-1996 | 1.445.555 | – | 30.200 |
| 1994-1995 | 1.555.638 | – | 22.587 |
| 1993-1994 | 1.593.833 | – | 15.229 |
| 1992-1993 | 1.580.552 | – | – |
| 1991-1992 | 1.633.432 | – | – |
| 1990-1991 | 1.592.330 | – | – |

(*) Datos Avance.
Fuente: Estadística de las Enseñanzas no universitarias. Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional.
https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado/series.html

En el recuadro 1 mostramos los niveles educativos y su correspondencia con el nivel de estudios, tal y como son considerados en las estadísticas oficiales del Ministerio de Educación y Formación Profesional y en los Informes internacionales de la OCDE.

**Recuadro 1**

| Nivel educativo | Estudios que comprende |
|---|---|
| Bajo | Hasta la ESO como mucho |
| Medio (Secundaria Superior) | Finalizado FP Media o Bachillerato (desde hace poco también FP Básica) |
| Superior (Terciaria) | Finalizado FP Superior o Grado, Máster o Doctorado universitario |

La tabla 4 muestra la distribución de los titulados en Secundaria Superior entre Bachillerato y FP Media en los principales países de la UE, además de Reino Unido y Estados Unidos (OCDE: Education at a Glance, 2022). Se puede observar que en España básicamente dos tercios (64,7%) de los jóvenes que se gradúan por primera vez en Secundaria Superior titulan en Bachillerato, mientras que un tercio (35,3%) lo hace en FP Media. La distribución de titulados es algo más elevada a favor del Programa General a la que veíamos en la tabla 3 por cuanto que la información de los datos de titulados tiene un retardo de dos años con respecto a los matriculados y también porque hay más estudiantes que finalizan los estudios en Bachillerato que en FP Media.

La tabla 4 muestra también que el porcentaje de titulados en FP Media con respecto al total en España es algo más baja que en el conjunto de la OCDE, en donde es dos puntos porcentuales superior (37,3%) sobre todo por el peso de Estados Unidos, en donde los estudios profesionales no tienen presencia. La diferencia es mucho mayor entre España y la UE-22, con más de ocho puntos porcentuales de distancia (35,3% frente a 44,7%). En definitiva, nuestro país necesita de más estudiantes en FP Media, un crecimiento que no se tiene que producir a costa de una disminución del alumnado que libremente elige Bachillerato. El aumento de la matrícula de FP Media, que ya se viene registrando en España, tendría que continuar a partir de una disminución de los jóvenes que no siguen sus estudios al finalizar la ESO o que ni siquiera terminan este nivel de estudios. La FP Media y Superior están mejorando significativamente su reputación en España, alejándose de la imagen que tuvo hace años en el sentido de ser el itinerario adecuado para los estudiantes sin éxito académico.

Uno de los problemas de la FP Media, que explica que haya menos proporción de titulados en FP Media con respecto a Bachillerato que en la distribución de matriculados, es el elevado porcentaje de estudiantes que comienza este nivel de estudios y no lo finaliza. En concreto el 30,7% del alumnado de nuevo ingreso en Ciclos Formativos de Grado Medio en 2016-2017 no ha logrado el título ni se encuentra matriculado en el sistema educativo en el curso 2020-2021 (Sistema Estatal de Indicadores Educativos 2023a). Una posibilidad para lograr que más estudiantes accedan a los Ciclos Formativos de Grado Medio y finalicen sus estudios en España sería impulsar la modalidad dual. La mayor inserción laboral de la FP Dual (Bentolila et al., 2023) y el hecho de que la mitad del tiempo de estudios se desarrolle en un centro de trabajo aplicando las competencias adquiridas pueden atraer y retener en FP Media a más jóvenes. La Formación Profesional Dual en España en estudios profesionales de Secundaria Superior alcanza sólo al 3% de los alumnos por el 38% en la UE-23 y el 34% de la OCDE (Education at a Glance, 2020a). Aunque los datos se refieren a 2020, desde entonces la situación no ha cambiado mucho en España, pues en el curso siguiente el porcentaje de estudiantes en modalidad dual sigue en registros muy similares (4,3% en FP Media y 1,2% para FP Básica, en 2021-2022 como muestran los últimos datos de Estadísticas del Alumnado de Formación Profesional del Ministerio de Educación y Formación Profesional). En los Ciclos de Formación de Grado Superior, la formación dual representa un 7,1% de la matrícula total, con la Comunidad Foral de Navarra como líder alcanzando un 21,5% (Fundación Europea Sociedad y Educación 2023: Indicadores comentados sobre el estado del sistema educativo en España https://www.sociedadyeducacion.org/core/wp-content/uploads/INDICADORES_-2023_WEB_09082023.pdf )

En España, el 22% de los jóvenes de 17 años están matriculados en formación profesional y el 65% en Bachillerato. En la media OCDE, el 31% está en formación profesional y el 55% en la segunda etapa de Educación Secundaria general. En la UE-22, aumentan hasta el 37% los jóvenes matriculados en formación profesional (Education at a Glance, OCDE, 2022). En cualquier caso, la diferenciación entre FP Superior y los Grados universitarios es en algunas ocasiones borrosa. El acceso a la Universidad desde la FP Superior es además una posibilidad que cada vez utilizan más jóvenes. La Ley Orgánica 3/2020, de Educación en su artículo 44, punto 6, establece que el título de Técnico Superior de Formación Profesional permitirá el acceso previa superación de un

procedimiento de admisión, a los estudios universitarios de Grado, así como las convalidaciones de los créditos universitarios que correspondan. También el Real Decreto 659/2023 establece que los titulados de Grado superior de FP que quieran acceder a un Grado universitario con mucha demanda pueden presentarse a la fase voluntaria de la EBAU para subir la nota media obtenida en el ciclo. De hecho, el 25,5% de los titulados en FP Superior en el curso 2018-2019 en España se matricula en estudios universitarios en alguno de los tres cursos siguientes (Ministerio de Educación y FP: Estadística de Seguimiento Educativo Posterior de los Graduados en Formación Profesional).

**Tabla 4. Distribución de titulados en 2020 entre Programa General (Bachillerato) y Formación Profesional (Ciclo Formativos de Grado Medio).**

| | Porcentaje de graduados por primera vez que obtuvieron un título general | Porcentaje de graduados por primera vez que obtuvieron una titulación profesional | Total |
|---|---|---|---|
| Francia | – | – | – |
| Alemania | 53,0 | 47,0 | 100,0 |
| Italia | 44,3 | 55,7 | 100,0 |
| España | 64,7 | 35,3 | 100,0 |
| Reino Unido | 35,0 | 65,0 | 100,0 |
| Estados Unidos | 100,0 | 0,0 | 100,0 |
| **OCDE** | **62,7** | **37,3** | **100,0** |
| **UE-22** | **56,3** | **43,7** | **100,0** |

Fuente: OCDE Education at a Glance, 2022.

### *1.1.4. Más población con estudios superiores en España que en la UE*

La diversidad en rutas educativas, junto con la creciente demanda de habilidades más especializadas en el ámbito laboral, ha llevado a un incremento global en la demanda de Educación Secundaria Superior y Terciaria. Se han establecido nuevos mecanismos de acceso a la educación superior a partir de estudios secundarios avanzados, como el Bachillerato o la FP de Grado Medio, brindando alternativas educativas que responden a las demandas de una sociedad en constante evolución. Las tasas de graduación terciaria reflejan la habilidad de un país para equipar a sus futuros profesionales con conocimientos y destrezas avanzadas. El acceso a este nivel educativo ha experimentado un notable

crecimiento, con la incorporación de distintas instituciones que presentan una variedad de opciones y modalidades de enseñanza.

La tabla 5 muestra el porcentaje de población con estudios superiores finalizados en España (41%) es más elevado que el del promedio de la UE-22 (38%) y está igualado al de la OCDE (41%). En comparación con los países de más tamaño, nuestro país está junto con Francia (también con el 41%) por debajo del Reino Unido y Estados Unidos, ambos con el 50% y por encima de Alemania (31%) e Italia (20%). El elevado porcentaje de población española con Ciclos Formativos de Grado Superior finalizados (12%) con respecto a la OCDE (7%) y UE-22 (5%) compensa el déficit de personas con el título de Grado o equivalente (11% frente al 19% de la OCDE y el 15% de la UE-22. La proporción de personas entre 25-64 años con Licenciatura, Máster o equivalente (16%) es similar a los promedios de OCDE y UE-22 y más elevada que los cinco países analizados (Francia, Alemania, Italia, Reino Unido y Estados Unidos).

**Tabla 5. Nivel educativo alcanzado de personas de 25 a 64 años (2021). Porcentaje de adultos con nivel de estudios superiores.**

| | Educación Superior/ Nivel terciario | | | | Total |
|---|---|---|---|---|---|
| | Terciario de ciclo corto (FP Superior) | Graduado o equivalente | Licenciatura, Máster o equivalente | Doctorado o equivalente | |
| | (1) | (2) | (3) | (4) | |
| Francia | 14 | 12 | 14 | 1 | 41 |
| Alemania | 1 | 18 | 11 | 2 | 31 |
| Italia | 0 | 5 | 14 | 1 | 20 |
| España | 12 | 11 | 16 | 1 | 41 |
| Reino Unido | 9 | 26 | 13 | 2 | 50 |
| Estados Unidos | 11 | 25 | 12 | 2 | 50 |
| **OCDE** | **7** | **19** | **14** | **1** | **41** |
| **UE-22** | **5** | **15** | **17** | **1** | **38** |

Fuente: OCDE Education at a Glance, 2022.

La tabla 6 muestra que, si nos centramos en la población más joven, el tramo de edad de los 25-34 años, se puede llegar a la conclusión de que los estudios superiores en España continúan al alza. Entre 2018 y 2022, la proporción de la población en ese tramo de edad ha crecido desde el 44,3% al 50,5%, superando ya

el objetivo europeo del 45% fijado para 2030. La diferencia entre España y la UE que veíamos para la población entre 25-64 años en la tabla 5 no sólo se mantiene, sino que se acrecienta entre los jóvenes, puesto que el promedio del tramo 25-34 años con estudios superiores en la UE-27 es del 42,0%. La desagregación por tipo de programas indica que tanto los estudios universitarios como los Ciclos Formativos de Grado Superior están contribuyendo al crecimiento de titulados superiores en España, con un aumento porcentual mayor en el caso de la FP Superior (20% frente al 12%). La buena evolución de la FP de Grado Superior en España comienza con un crecimiento paulatino y sostenido de la tasa bruta de matriculación en el curso 2007-2008 que propulsó la tasa desde el 22,3% al 43,1% del curso 2021-2022 (Fundación Europea Sociedad y Educación, 2023).

**Tabla 6: Evolución de la distribución de la población de 25-34 años con nivel de educación superior en España con respecto al tipo de programa.**

| | Estudios Superiores Universitarios | FP Superior | Total |
|---|---|---|---|
| 2018 | 31,0 | 13,2 | 44,3 |
| 2019 | 32,7 | 13,8 | 46,5 |
| 2020 | 33,3 | 14,0 | 47,4 |
| 2021 | 34,0 | 14,7 | 48,7 |
| 2022 | 34,7 | 15,8 | 50,5 |

Fuente: Ministerio de Educación y FP. Nivel de formación, Formación permanente y Abandono: Explotación de las variables educativas de la Encuesta de Población Activa https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/laborales/epa.html

En la tabla 7, se muestra la información de tasas de graduación de menos de 30 años en estudios superiores. Son datos complementarios a los de la titulación de la tabla 6 puesto que, al tratarse de tasas de graduación de menos de 30 años, es un indicador adelantado que muestra la proporción de los más jóvenes que han finalizado estudios universitarios o de FP Superior. La graduación por primera vez excluye a aquellos egresados que ya habían obtenido algún título en estudios superiores previamente, como puede ser el caso de una persona que ha finalizado FP de Grado Superior y después titula también en un Grado universitario o alguien que termina estos estudios y después continúa con Máster. La tasa de graduación por primera vez a los 30 años mide, por tanto, la relación entre graduados por primera vez de menos de 30 años entre la población de menos de 30 años. Podemos comprobar, aunque los últimos datos con comparación internacional sean de 2019, si la ventaja que tiene España en porcentaje de titulados superiores en la

población, en general, y entre los jóvenes, en particular, se va a consolidar en los próximos años. Las tasas de graduación en estudios superiores en España, la suma de la FP de Grado Superior y Universidad, han alcanzado los mejores registros de la OCDE. Así, por ejemplo, la tasa de graduación por primera vez de los menores de 30 años en España (excluyendo alumnos internacionales) se encontraba en el 54%, la primera de toda la OCDE y muy por encima del promedio del 38% o el 37% de la UE-22 (columna 4 de la tabla 7, Education at al Glance, 2021, p. 208). Las columnas 1-3 revelan la distribución de titulados superiores por primera vez entre FP Superior, Grado o Máster. Así en España, el 43% de los graduados superiores han finalizado Ciclos Formativos de Grado Superior, el 45% Grado y el 12% Máster, haciendo un total del 100%. El peso de la FP Superior en España es más del triple que en la UE (13%), sólo al nivel de Estados Unidos (40%). En definitiva, la tabla 7 confirma las conclusiones previas en el sentido de que una parte relevante de los buenos datos del acceso a estudios superiores de España lo compone la FP de Grado, en concreto en un 43% (columna 1).

**Tabla 7. Tasa de graduación y características de los graduados superiores por primera vez.**

| | % de graduados por 1ª vez por nivel educativo (2021) | | | Tasa de graduación por 1ª vez de menores de 30 años (2019) | | | |
|---|---|---|---|---|---|---|---|
| | FP Grado Superior (2-3 años) | Grado universi-tario | Máster o equiva-lente | Sin incluir alumnos internacionales | | | Total |
| | | | | Total | Hombres | Mujeres | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **OECD** | | | | | | | |
| Francia | – | – | – | – | – | – | – |
| Alemania | 1 | 86 | 12 | 37 | 34 | 41 | 39 |
| Italia | 2 | 83 | 15 | 35 | 29 | 43 | 36 |
| España | 43 | 45 | 12 | 54 | 47 | 60 | 56 |
| Reino Unido | 23 | 76 | 1 | 43 | 36 | 49 | 50 |
| Estados Unidos | 40 | 60 | – | – | – | – | – |
| **OCDE** | **16** | **76** | **8** | **38** | **31** | **46** | **42** |
| **UE-22*** | **13** | **78** | **10** | **37** | **30** | **45** | **38** |

* El dato de la UE-22 es el relativo al año 2019.
Fuente. OCDE, Education at a Glance, 2023, página 241 https://www.oecd.org/education/education-at-a-glance/ OCDE, Education at a Glance, 2021, página 208 https://www.oecd.org/education/education-at-a-glance/

## 1.2. Profesorado y entorno de aprendizaje

En este apartado se señalan dos de las condiciones laborales de los docentes españoles. Comenzamos por el salario anual medido en paridades de poder de compra del dólar, una medida que tiene en cuenta la diferencia en el coste de la vida entre países. En Educación Infantil y Primaria, los maestros españoles tienen un salario superior al de los franceses e italianos y, en el caso de Primaria, inferior al de los alemanes. Los salarios de los docentes españoles con 15 años de experiencia se encuentran algo por encima del promedio de la UE-25, tanto en Educación Infantil (+21,4%), como en Primaria (+10,2%) (Indicador D3.1 OECD, Education at a Glance, 2023). En algunos países, distintos de los mostrados en la tabla 8, la remuneración de los docentes de Primaria es algo superior a la de los de Educación Infantil. La variación de salarios en España a lo largo de la carrera profesional es más pequeña que en el conjunto de la UE-25. Los maestros europeos de Primaria, por ejemplo, comienzan su carrera con una remuneración de 35.175 dólares y tienen un máximo de 58.742 dólares, un 67% más. Por el contrario, en España el salario inicial es de 44.650 dólares y el máximo es de 63.910 dólares, un 43,1% de diferencia.

**Tabla 8. Salarios estatutarios de docentes de ESO, Bachillerato y FP Media, en Paridades de Poder de Compra del dólar en diferentes puntos de la carrera de los docentes (2022). Salarios anuales en instituciones públicas.**

| | Infantil | | | | Primaria | | | |
|---|---|---|---|---|---|---|---|---|
| | Salario inicial | Salario después de 10 años de experiencia | Salario después de 15 años de experiencia | Salario al tope de la escala | Salario inicial | Salario después de 10 años de experiencia | Salario después de 15 años de experiencia | Salario al tope de la escala |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Francia | 34.611 | 38.651 | 40.683 | 58.751 | 34.611 | 38.651 | 40.683 | 58.751 |
| Alemania | – | – | – | – | 70.419 | 80.750 | 85.699 | 91.713 |
| Italia | 32.981 | 36.160 | 39.584 | 48.015 | 32.981 | 36.160 | 39.584 | 48.015 |
| España | 44.650 | 48.516 | 51.715 | 63.910 | 44.650 | 48.516 | 51.715 | 63.910 |
| Inglaterra | 34.732 | – | 55.726 | 55.726 | 34.732 | – | 55.726 | 55.726 |
| Escocia | 43.895 | 55.096 | 55.096 | 55.096 | 43.895 | 55.096 | 55.096 | 55.096 |
| Estados Unidos | 45.931 | 50.953 | 68.905 | 76.985 | 44.992 | 61.054 | 66.251 | 78.190 |
| **OCDE** | **34.563** | **43.063** | **45.981** | **57.118** | **36.367** | **46.782** | **49.968** | **61.075** |
| **UE-25** | **32.541** | **39.061** | **42.586** | **53.626** | **35.175** | **42.682** | **46.894** | **58.742** |

Nota: La definición de las calificaciones más prevalentes de los docentes se basa en un concepto amplio, que incluye el típico nivel de logro ISCED y otros criterios.
Fuente: OCDE Education at a Glance (2023).

La tabla 9 muestra las remuneraciones de los docentes de la ESO (Educación Secundaria inferior) y Bachillerato y FP Media (Educación Secundaria Superior). El salario de los profesores españoles de la ESO, FP Media y Bachillerato es superior al de los de Francia, Italia, Inglaterra y Escocia e inferior al Estados Unidos y, sobre todo, Alemania. Las remuneraciones de los profesores españoles de la ESO con 15 años de experiencia son un 18,8% superiores al promedio comunitario y de un 15,3% en el caso de los de FP Media y Bachillerato. Como ocurriría también con los salarios de Infantil y Primaria, las posibilidades de promoción económica en nuestro país son inferiores al conjunto de la OCDE y la UE-25.

**Tabla 9. Salarios estatutarios de docentes de ESO, Bachillerato y FP Media, en Paridades de Poder de Compra del dólar en diferentes puntos de la carrera de los docentes (2022). Salarios anuales en instituciones públicas.**

| | ESO | | | | FP Media, Bachillerato | | | |
|---|---|---|---|---|---|---|---|---|
| | Salario inicial | Salario después de 10 años de expe-riencia | Salario después de 15 años de expe-riencia | Salario al tope de la escala | Salario inicial | Salario después de 10 años de expe-riencia | Salario después de 15 años de expe-riencia | Salario al tope de la escala |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Francia (euros)[1] | 37.720 | 41.760 | 43.792 | 62.169 | 37.720 | 41.760 | 43.792 | 62.169 |
| Alemania (euros) | 77.905 | 88.442 | 93.085 | 101.510 | 81.141 | 91.729 | 96.742 | 110.694 |
| Italia (euros) | 35.447 | 39.151 | 43.038 | 52.725 | 35.447 | 40.084 | 44.235 | 55.106 |
| España (euros) | 49.905 | 54.243 | 57.758 | 71.235 | 49.905 | 54.243 | 57.758 | 71.235 |
| Inglaterra (libras) | 34.732 | – | 55.726 | 55.726 | 34.732 | – | 55.726 | 55.726 |
| Escocia (libras) | 43.895 | 55.096 | 55.096 | 55.096 | 43.895 | 55.096 | 55.096 | 55.096 |
| Estados Unidos ($) | 46.018 | 64.196 | 69.439 | 79.031 | 48.187 | 63.026 | 69.641 | 75.988 |
| **OCDE** | **37.628** | **48.605** | **51.613** | **63.332** | **39.274** | **50.841** | **53.456** | **65.658** |
| **UE-25** | **36.446** | **44.444** | **48.606** | **61.090** | **36.985** | **45.986** | **50.082** | **63.474** |

[1] Incluye el promedio de bonificaciones fijas por horas extras para profesores de Secundaria inferior y superior.
Nota: La definición de las calificaciones más prevalentes de los docentes se basa en un concepto amplio, que incluye el típico nivel de logro ISCED y otros criterios.
Fuente: OCDE Education at a Glance, 2023.

La tabla 10 muestra que el número de días de enseñanza al año de los docentes españoles es similar a los de la UE-22 en ESO y Bachillerato, y algo más bajo en Primaria, FP Media y, sobre todo, Infantil. En cualquier caso, los datos del Informe Anual de la OCDE Education at a Glance podrían estar indicando el número de días de enseñanza en Educación Infantil de segundo ciclo en España (176 días al año como en el resto de las etapas educativas) mientras que en el caso del promedio de OCDE (196) y UE-22 (198) parece que incluye también el primer ciclo de infantil. Así, por ejemplo, Alemania e Italia proporcionan una información de días de clase al año diferente para Infantil y Primaria, una circunstancia que no ocurre en el caso de España, Francia, Inglaterra, Escocia o Estados Unidos. Las horas de enseñanza al año de los docentes de toda la Secundaria (ESO, FP Básica, Bachillerato y FP Media) en España son similares a las de la UE-22 y algo más bajas que las de la OCDE. En la etapa de Infantil, de nuevo, se registra una diferencia de hasta 200 horas menos de horas de enseñanza de los maestros españoles (871) con respecto a los de la UE-22 (1071), que podría deberse a que en España la información se circunscribe al segundo ciclo (3-5 años) y en la UE-22 a los dos ciclos (0-2 años y 3-5 años). En Primaria, por el contrario, los maestros españoles tienen más horas de clase (871) que los de UE-22 (740) y OCDE (784). En este nivel de enseñanza también hay menos días de clase en España (176) que en la UE-22 (180) y OCDE (184), lo que significa hay muchas más horas de clase por día en Primaria en nuestro país. Esta es una cuestión relevante, porque podría incidir en el hecho de muchos Colegios de Educación Infantil y Primaria (CEIP) en España estén eligiendo la jornada continua en lugar de la jornada partida. Como señalan Hospido et al. (2019), la jornada continua implica que la docencia se concentra durante la mañana (en general, en horario de 9 a 14 horas, con posibilidad de comedor posteriormente) frente a la jornada partida, que divide las horas de clase en mañana y tarde (habitualmente con horario de 9 a 16 horas con una pausa de dos horas en medio para comer). No se ha establecido aún una relación causal entre el tipo de jornada y el rendimiento educativo de los alumnos. La selección no-aleatoria de los centros entre la jornada continua y la partida dificulta el análisis del impacto de la estructura de la jornada escolar en el aprendizaje de los alumnos. De hecho, Hospido et al. (2019) muestran, que a medida que aumenta el nivel socioeconómico de los alumnos de los colegios de la Comunidad de Madrid, disminuye la proporción de centros con jornada continua.

**Tabla 10. Número de semanas reglamentarias de enseñanza, días de enseñanza y horas netas de enseñanza de los docentes en instituciones públicas durante el año escolar.**

| | Notas | Número de días de enseñanza al año | | | | | | Horas al año de enseñanza | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Infantil | Primaria | ESO | FP Básica | Bachillerato | FP Media | Infantil | Primaria | ESO | FP Básica | Bachillerato | FP Media |
| | | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) |
| Francia | 1 | 162 | 162 | – | – | – | – | 900 | 900 | 720 | – | 720 | 720 |
| Alemania | 1 | 225 | 193 | 193 | – | 193 | 193 | 1.755 | 691 | 641 | – | 610 | 626 |
| Italia | 3 | 189 | 169 | 169 | – | 169 | 169 | 945 | 744 | 608 | – | 608 | 608 |
| España | 1 | 176 | 176 | 176 | 176 | 176 | 176 | 871 | 871 | 665 | 665 | 665 | 665 |
| Inglaterra (Reino Unido) | | 190 | 190 | 190 | – | 190 | – | – | – | – | – | – | – |
| Escocia (Reino Unido) | 2 | 190 | 190 | 190 | – | 190 | – | 855 | 855 | 855 | – | 855 | – |
| Estados Unidos | 4 5 | 180 | 180 | 180 | – | 180 | – | 1.011 | 1.004 | 966 | – | 966 | – |
| **OCDE** | | **196** | **184** | **183** | **–** | **182** | **185** | **987** | **784** | **711** | **–** | **684** | **684** |
| **UE-22** | | **198** | **180** | **178** | **–** | **178** | **183** | **1.071** | **740** | **659** | **–** | **642** | **675** |

1. Tiempo de enseñanza típico (tiempo de enseñanza requerido para la mayoría de los docentes cuando no se aplican circunstancias específicas a los profesores); 2. Tiempo de enseñanza máximo; 3. Tiempo de enseñanza mínimo; 4. Tiempo de enseñanza real; 5. Año de referencia 2016.
Fuente: OCDE: Education at a Glance (2022).

## 2. INDICADORES DE FINANCIACIÓN EDUCATIVA

### 2.1. Gasto total en educación, público y privado

Los países invierten en educación para ayudar a fomentar el crecimiento económico, mejorar la productividad, contribuir al desarrollo personal y social y reducir la desigualdad social, entre otras razones. La financiación de la educación no responde de forma significativa a las fluctuaciones a corto plazo del PIB, sino que está más influenciada por tendencias en el crecimiento económico a largo plazo (OCDE: Education at a Glance, 2022). La tabla 11 muestra el gasto total en

instituciones educativas como porcentaje del PIB en 2019 (OCDE: Education at a Glance, 2022). El gasto en instituciones educativas incluye el gasto público, privado e internacional en entidades (colegios, institutos, universidades...) que proporcionan servicios de enseñanza o servicios relacionados con la educación a individuos. Se excluyen los gastos en bienes y servicios de enseñanza que ocurren fuera de las instituciones educativas. Por ejemplo, no se incluye el gasto de las familias en la compra de libros de texto y materiales o las tutorías privadas para sus hijos. En el nivel terciario, los costes de la vida de los estudiantes y los salarios no obtenidos también están excluidos de este indicador.

El gasto total en instituciones educativas en España en 2020 era del 5,8% del PIB, tres décimas por debajo del 6,1% de la OCDE y cuatro décimas por encima de del 5,4% de la UE-25. Los datos de 2020 están influidos por la crisis de la pandemia del Covid que redujo en PIB de la OCDE, incluso más en el caso de España. La recesión aumenta la ratio del gasto público en educación (que no disminuyó) en relación con el PIB, más aún en los países más afectados por la crisis. La información del gasto en educación en España en 2021, aún no recogidos por el Education at a Glance de la OCDE), muestran que nuestro país todavía no ha recuperado el gasto público en educación sobre el PIB que había en 2009 (Fundación Europea Sociedad y Educación, 2023). Educación Infantil y FP Media es el nivel educativo con una menor financiación relativa de España con respecto a la UE-25 (no sólo por el dato de 2020, sino también por la información de años previos), mientras que Bachillerato es la etapa con más recursos en nuestro país en relación con el conjunto comunitario (la diferencia en Primaria se ha producido sólo en 2020). Los demás niveles de educación tienen una financiación similar en España que en la UE-22. La Universidad española dispone de unos recursos similares a los de los países de nuestro contexto europeo (1,2% frente a 1,3%) aunque está lejos de los lideres en investigación como son el Reino Unido (2,1%) o Estados Unidos (2,5%, incluyendo en este caso tanto Universidad como FP Superior). El nivel de renta per cápita de España es algo inferior al conjunto de la OCDE mientras que nuestra población está algo más envejecida, dos factores que pueden influir en las estadísticas de financiación educativa cuando se miden en términos del porcentaje del PIB. A continuación, se analizarán los datos de gasto en educación por estudiante, que es un indicador si cabe más relevante.

El anuario estadístico de la OCDE, Education at a Glance, también ofrece información sobre la desagregación del gasto en instituciones educativas entre

**Tabla 11. Gasto total en instituciones educativas como porcentaje del PIB (2020).**

| | Infantil | Primaria | Secundaria | | | | | Postsecundaria no terciaria | Educación Superior/Terciaria | | | Total gasto en Educación (incluyendo I+D) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | ESO | Secundaria Superior | | | Total Secundaria | | FP Superior | Universidad | Total terciaria (incluyendo I+D) | |
| | | | | Bachillerato | FP Media | Total Secundaria Superior | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (1+2+7+8+11) |
| Francia | 0,7 | 1,3 | 1,3 | 0,8 | 0,5 | 1,3 | 2,6 | 0,0 | 0,4 | 1,3 | 1,6 | 6,2 |
| Alemania | 1,1 | 0,7 | 1,3 | 0,4 | 0,6 | 1,0 | 2,3 | 0,2 | 0,0 | 1,3 | 1,3 | 5,8 |
| Italia | 0,5 | 1,3 | 0,7 | – | – | 1,2 | 1,9 | – | 0,0 | 1,0 | 1,0 | 4,7 |
| España | 0,8 | 1,5 | 1,0 | 0,6 | 0,4 | 1,1 | 2,0 | – | 0,2 | 1,2 | 1,5 | 5,8 |
| Reino Unido | 0,4 | 1,9 | 1,0 | 1,0 | 0,4 | 1,3 | 2,4 | – | 0,1 | 1,9 | 2,1 | 6,7 |
| Estados Unidos | – | 1,7 | 0,9 | 1,0 | – | 1,0 | 1,9 | 0,0 | – | – | 2,5 | – |
| **OCDE** | **0,9** | **1,5** | **1,0** | **0,6** | **0,5** | **1,1** | **2,0** | **–** | **0,1** | **1,4** | **1,5** | **6,1** |
| **UE-25** | **0,9** | **1,3** | **1,0** | **0,4** | **0,5** | **1,0** | **1,9** | **0,1** | **0,1** | **1,3** | **1,3** | **5,4** |

Fuente: OCDE: Education at a Glance (2023), p. 292.

fuentes de financiación públicas y privadas (tabla C2.3 de Education at a Glance 2023, p. 294). La tabla 11 muestra que, si excluimos la Educación Infantil, para la que la OCDE no ofrece desagregación por fuentes de financiación, el gasto en instituciones educativas en Primaria, Secundaria y Terciaria en España suma el 5,0% (columna 2 + columna 7 + columna 8 + columna 11). Esa misma cifra de la suma de todos los niveles educativos excluyendo infantil es del 4,5% en la UE-22. Pues bien, en nuestro país el peso de la financiación privada es del 0,9% del PIB sobre ese total de 5,0% del PIB que se destina a todos los niveles educativos una vez excluida la etapa de infantil, mientras que en la UE-22 es del 0,4% del PIB (sobre el total del 4,5%). En definitiva, el peso de la financiación privada en España es el doble que en el conjunto de los países comunitarios.

Una de las conclusiones de la tabla 11 era que el gasto de España en Bachillerato era superior al de la UE-25 mientras que el de la FP Media resultaba ser inferior. Pero podría ser la financiación relativa de nuestro país con respecto al conjunto de la Unión fuera simplemente el resultado de que tenemos más estudiantes en Bachillerato y menos en FP Media, como vimos en la sección 1.1 que nuestros socios comunitarios. De hecho, el nivel de gasto en instituciones educativas está afectado primeramente por el tamaño de la población en edad escolar de un país además de las tasas de matrícula, los niveles de salarios de los docentes y la organización de los sistemas educativos. La tabla 12 muestra el gasto total en instituciones educativas, la misma definición que en la tabla 11, pero por estudiante equivalente a tiempo completo. Las conclusiones cambian de forma muy significativa. El gasto total en instituciones educativas en España se encuentra por detrás de la UE-22 en todos los niveles educativos de Primaria y Secundaria, exceptuando justamente la FP Media. El gasto por alumno en Primaria en España (9.077 dólares) es del 88% del promedio de la UE-22 (10.337 dólares), del 88% en ESO y del 96% en Bachillerato. El reducido número de estudiantes de FP Media en España con respecto al conjunto de la Unión puede explicar que no se hayan alcanzado economías de escala en nuestro país, de modo que el gasto por estudiante en esta etapa educativa sea más elevado. En estudios superiores, España tiene también un déficit de financiación por alumno. El gasto por estudiante en FP Superior es del 95% del conjunto de la UE-22 mientras que en la Universidad es del 83%. Fuera de la Unión, la financiación por alumno de educación superior de España (columna 11) no llega a la mitad de la de Reino Unido o Estados Unidos (líderes destacados en investigación).

**Tabla 12. Gasto total en instituciones educativas por estudiante equivalente a tiempo completo (2019). En USD equivalentes usando PPPs para el PIB.**

| | Primaria | Secundaria | | | | | Primaria y Secundaria | Terciaria | | | Primaria hasta Terciaria (incluyendo I+D) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ESO | Secundaria Superior | | | Total Secundaria | | FP Superior | Grado, Máster, Doctorado | Total Terciaria (incluyendo I+D) | |
| | | | Bachillerato | FP Media | Total Secundaria Superior | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (8) | (9) | (10) | (11) | (13) |
| Francia | 9 673 | 12 139 | 15 279 | 18 142 | 16 266 | 13 874 | 12 119 | 17.468 | 19.315 | 18.880 | 13.545 |
| Alemania | 11 587 | 14 197 | 15 681 | 20 394 | 18 098 | 15 614 | 14 343 | 7.981 | 20.788 | 20.760 | 15.767 |
| Italia | 12 008 | 9 760 | – | – | 11 059 | 10 569 | 11 096 | 4.697 | 12.746 | 12.663 | 11.439 |
| España | 9 077 | 10 658 | 10 482 | 14 188 | 11 668 | 11 159 | 10 173 | 10.770 | 15.354 | 14.361 | 11.123 |
| Reino Unido | 12 513 | 12 716 | 14 609 | 14 370 | 14 539 | 13 695 | 13 141 | 29.292 | 29.552 | 29.534 | 16.052 |
| Estados Unidos | 14 321 | 15 302 | 16 775 | | 16 775 | 16 018 | 15 186 | – | – | 36.172 | 19.973 |
| **OCDE** | **10 658** | **11 941** | **11 379** | **13 216** | **12 312** | **11 942** | **11 352** | **12.266** | **19.775** | **18.105** | **12.647** |
| **UE-25** | **10 337** | **12 111** | **10 967** | **13 072** | **11 664** | **11 696** | **11 123** | **11.317** | **18.571** | **17.578** | **12.275** |

Fuente: OCDE. Education at a Glance, 2023.

En suma, el gasto por alumno en España desde Primaria hasta estudios superiores (11.123 dólares) es un 9,4% más bajo que en el promedio de la UE-22 (12.275 dólares). Es sorprendente que, si el salario de los docentes españoles es algo superior que el del promedio de la UE, el gasto por estudiante en nuestro país sea inferior. Uno de los factores que explica en buena medida esta contradicción es la mayor ratio de alumnos por clase en España. Como se verá en el capítulo 3a sobre el impacto del tamaño de la clase en el rendimiento académico, nuestro país tiene un tamaño de clase promedio de 22 alumnos en Primaria (por los 19 en la UE-22) y 25 en ESO (por los 21 en la UE-22). En definitiva, los docentes españoles tienen algo más de salario que en la UE, una condición que más que se compensa con una mayor ratio de estudiantes por aula, para concluir con un gasto por alumno en España inferior.

Al examinar los datos a nivel regional en España, se puede comprobar que el gasto público por estudiante en educación no universitaria en instituciones públicas y concertadas en 2020 está liderado por el País Vasco (7.329 euros), Navarra (6.920 euros) y Galicia (6.714 euros) (Fundación Europea Sociedad y Educación, 2023). En relación con la tendencia general, se detecta un crecimiento sostenido hasta 2009. Sin embargo, este patrón ascendente cambió ese año, iniciando una disminución hasta 2013. A partir de ese momento, se ha logrado recuperar gran parte del gasto previo. Andalucía, Canarias, Castilla y León, y Extremadura son las únicas comunidades que ya han alcanzado o superado los niveles de gasto público por estudiante en instituciones públicas registrados en 2009 (Fundación Europea Sociedad y Educación, 2023).

### 2.2. Becas y ayudas públicas

La tabla 13 muestra el importante crecimiento del gasto público en becas y ayudas al estudio destinado a becas y ayudas. En 2021, la partida destinada a estos conceptos alcanzó los 2.491 millones de euros en España, después de un importante incremento del 19,4% entre los ejercicios 2020 y 2021. Dos tercios (67,7%) de las becas y ayudas las financia la administración general a través de los Ministerios de Educación y FP y el Ministerio de Universidades (Estadísticas de las becas y ayudas al estudio. Curso 2021-2022, Ministerio de Educación y FP, 2023c) mientras que el resto corresponde a las Comunidades Autónomas. Los niveles no-universitarios de Bachillerato, FP Media y FP Superior, excluidos los de enseñanza obligatoria que es totalmente gratuita, absorben el 46,8% del total de las

**Tabla 13. Gasto Público en becas y ayudas al estudio por niveles educativos y periodo (miles de euros).**

| | Total | Niveles no universitarios | Niveles universitarios excepto exención de precios académicos | Exención de precios académicos | Sin desglosar por niveles |
|---|---|---|---|---|---|
| 2021 | 2.491.256 | 1.156.489 | 932.529 | 380.591 | 21.647 |
| 2020 | 2.086.632 | 830.705 | 780.351 | 364.709 | 110.867 |
| 2019 | 2.080.173 | 827.025 | 765.451 | 386.177 | 101.520 |
| 2018 | 2.104.192 | 888.341 | 724.134 | 387.167 | 104.550 |
| 2017 | 2.084.715 | 949.035 | 677.596 | 361.724 | 96.360 |
| 2016 | 1.919.657 | 846.207 | 629.047 | 350.074 | 94.329 |
| 2015 | 1.875.086 | 791.145 | 658.777 | 334.101 | 91.063 |
| 2014 | 1.872.379 | 860.481 | 607.094 | 324.283 | 80.521 |
| 2013 | 1.530.114 | 672.937 | 557.512 | 228.302 | 71.363 |
| 2012 | 1.697.328 | 734.483 | 729.243 | 204.905 | 28.697 |
| 2011 | 1.799.499 | 829.847 | 688.220 | 241.702 | 39.730 |
| 2010 | 1.759.591 | 781.912 | 637.275 | 244.668 | 95.736 |
| 2009 | 1.770.540 | 753.917 | 628.425 | 235.381 | 152.818 |
| 2008 | 1.750.530 | 555.964 | 793.259 | 226.709 | 174.598 |
| 2007 | 1.465.721 | 375.449 | 574.232 | 209.571 | 306.469 |
| 2006 | 1.285.768 | 371.275 | 475.657 | 194.872 | 243.965 |
| 2005 | 1.152.289 | 357.509 | 455.575 | 185.356 | 153.849 |
| 2004 | 1.023.963 | 327.684 | 404.698 | 185.383 | 106.198 |
| 2003 | 944.836 | 292.804 | 382.116 | 188.676 | 81.240 |
| 2002 | 808.568 | 244.389 | 329.303 | 180.623 | 54.253 |
| 2001 | 793.886 | 210.451 | 283.747 | 237.290 | 62.398 |
| 2000 | 731.969 | 200.004 | 281.759 | 183.631 | 66.575 |
| 1999 | 707.301 | 196.745 | 250.496 | 199.990 | 60.070 |
| 1998 | 762.468 | 214.286 | 271.913 | 227.822 | 48.447 |
| 1997 | 622.140 | 175.182 | 267.228 | 139.792 | 39.938 |
| 1996 | 761.827 | 212.961 | 339.531 | 166.363 | 42.972 |
| 1995 | 578.000 | 203.319 | 247.884 | 92.925 | 33.872 |
| 1994 | 531.634 | 197.892 | 209.840 | 86.722 | 37.180 |
| 1993 | 528.617 | 182.637 | 242.148 | 78.569 | 25.263 |
| 1992 | 399.575 | 154.468 | 168.165 | 58.504 | 18.438 |

Fuente: Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional. Recursos económicos. Gasto público. Series
https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/economicas/gasto/series.html

becas y ayudas mientras que la Universidad representa el 53,2% restante. La proporción de beneficio privado que obtienen los individuos que realizan estudios universitarios (en forma de mejores empleos y salarios futuros como veremos después) es mayor que la rentabilidad que recibe la sociedad por tener un ciudadano más con formación superior, a través del pago de mayores impuestos y de tener una población con mayor nivel educativo y más capacidad de generar valor añadido (Levin et al., 2017). En los niveles inferiores de Secundaria superior (Bachillerato o FP Media) y FP Superior esa relación entre el beneficio que alcanza el propio titulado y el que tiene la sociedad es más equilibrado, por lo que una posibilidad en el futuro sería concentrar en mayor medida los aumentos de las becas y ayudas en los niveles educativos no universitarios. De hecho, entre 2020 y 2021, el aumento en el gasto en becas y ayudas en los niveles no universitarios ha sido del 39,2% por el 14,7% en la Universidad.

Nos centramos ahora en las becas y ayudas en la educación superior (incluyendo, por tanto, la FP Superior como la Universidad) para la que sí existe comparación internacional. La tabla 14 muestra los datos de becas y ayudas de España en comparación con los países más grandes de la UE-22 y la OCDE. El 78% de los estudiantes de FP Superior en centros públicos en España tienen exención de matrícula, una proporción similar a la de Francia. En la mayoría de las Comunidades Autónomas la matrícula es gratuita en los Institutos, mientras que en otras hay un pago de 400 euros. También hay un porcentaje elevado de becas entre los estudiantes de Grado universitario (53%) en comparación con los países para los que hay información (Francia, 42% e Italia, 39%). Por el contrario, la proporción de exentos de matrícula en España en el Máster (24%) es inferior a Francia e Italia. Parece que nuestro país apuesta por centrar las ayudas en Grado universitario en lugar de Máster, una política que está en línea con el hecho de que cuanto mayor es el nivel educativo más elevada es la proporción de rentabilidad privada con respecto a la social. Dado que hay un porcentaje más elevado de rentabilidad privada en los Máster que en los Grados universitarios, es una decisión acertada concentrar las ayudas en los segundos. Las columnas (5) y (7) recogen la existencia de préstamos privados o públicos, en solitario o combinadas con becas) para realizar estudios superiores. Este sistema de préstamos tiene una gran presencia en Reino Unido, donde un 96% de sus estudiantes recibe este tipo de préstamos, y también en Estados Unidos (53% sumando las dos

**Tabla 14. Criterios de elegibilidad para becas o subvenciones públicas para estudiantes de educación superior matriculados en instituciones públicas y privadas (2019-2020).**

| | | Estudiantes beneficiarios de exenciones en las tasas de matrícula (instituciones públicas) | | | | Estudiantes terciarios que recibieron (%): | | | | Montante anual promedio de becas públicas (USD convertidos usando PPPs) | Tipos de becas/subvenciones | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | FP Superior | Grado | Máster | Doctorado | Solo préstamos privados garantizados por el gobierno o préstamos públicos | Solo becas públicas | Becas públicas y préstamos privados garantizados por el gobierno o públicos. | Ni becas públicas ni préstamos privados garantizados por el gobierno o públicos. | | Universal | Basadas en la capacidad económica | Basadas en méritos | Dirigidas a grupos poblacionales desfavorecidos |
| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (10) | (11) | (12) | (13) | (14) |
| Francia | 1 | 76 | 42 | 31 | – | – | 34 | – | 66 | 2.589 | NO | SI | NO | NO |
| Alemania | 1, 2 | – | – | – | – | – | – | 22 | – | 4.193 | NO | SI | NO | NO |
| Italia | | – | 39 | 38 | 84 | 0 | 38 | 0 | 62 | 7.039 | NO | NO | SI | SI |
| España | | 78 | 53 | 24 | – | – | 44 | 0 | 56 | – | NO | SI | SI | NO |
| Inglaterra (Reino Unido) | 1, 3 | – | – | – | – | 96 | 0 | 0 | 4 | – | NO | SI | NO | SI |
| Estados Unidos | 1, 4 | – | – | – | – | 9 | 33 | 44 | 15 | 2.178 | | SI | SI | SI |

1. Año de referencia: año académico 2015/2016 para Alemania y Estados Unidos; año académico 2018/2019 para Inglaterra (Reino Unido), Francia y España; 2. Los datos no se refieren a estudiantes que reciben apoyo financiero público (BAföG) sino a aquellos elegibles para solicitarlo: 3. Instituciones privadas dependientes del gobierno en lugar de instituciones públicas; 4. La distribución de préstamos se refiere solo a ISCED 5 y 6.
Fuente: OCDE: Education at a Glance, 2022.

columnas) y Alemania (22%). Montalbán (2019) explica los mecanismos de este sistema de préstamos. Básicamente se trata de que los estudiantes piden un préstamo antes de empezar sus estudios universitarios para poder acceder a ellos y financiar su coste. Una vez han terminado de estudiar e incorporados al mercado laboral, deberán devolver el préstamo en función de sus ingresos. Si obtienen una mayor renta, devolverán una mayor cantidad de préstamo. Suele existir un nivel mínimo de renta que está exenta, y los pagos se establecen hasta un número máximo de años (Montalbán, 2019). Cabrales et al. (2019) establecen un laboratorio sobre Préstamos Contingentes al Ingreso en el que se analizan los efectos distributivos de diferentes sistemas de préstamos para financiar la educación universitaria a los costes actuales, así como para incrementar la financiación superior para mejorar su calidad. Con este objetivo, se utilizan para las simulaciones la información de salarios durante toda la vida de los graduados de los datos administrativos de la seguridad social. Cabrales et al. (2019) encuentran que el cuartil superior de la distribución de renta pagaría cerca del montante total de la matrícula, mientras que el 10% inferior no pagaría prácticamente ninguna matrícula. Además, la proporción de la financiación universitaria total subvencionada por el gobierno se reduciría entre 16 y 56 puntos porcentuales con respecto al sistema actual.

En España las becas universitarias se proporcionan tanto por rentas inferiores a umbrales determinados como por méritos. En concreto, el mantenimiento de las becas está condicionado a un rendimiento académico mínimo. Montalbán (2019) analiza la reforma nacional del sistema de becas introducida en 2013 por la que se elevaron los requisitos académicos con respecto a los preexistentes (comparables al de otros programas nacionales en el mundo). En cinco de los seis países analizados, las becas se basan en criterios de capacidad económica, concentrando las ayudas en aquellos estudiantes provenientes de hogares de menor renta, en donde además de los costes de asistencia a los estudios superiores (libros, materiales, manutención) hay que añadir el coste de oportunidad de no obtener un ingreso económico por trabajo.

Una alternativa para aumentar las becas y ayudas a los estudiantes provenientes de los hogares de menor renta, sin aumentar significativamente el coste presupuestario, sería aumentar la proporción del coste de la plaza universitaria que pagan el resto de las familias con renta media o alta. García-

Montalvo (2020) analiza los efectos del aumento en un 66,6% en los precios de matrícula de las universidades públicas catalanas para los alumnos de mayor nivel socioeconómico en el año académico 2012-2013. Este autor muestra que este aumento de precios públicos fue progresivo, es decir, hubo un subsidio diferencial en función de los ingresos familiares. En primer término, su estudio no encuentra que el aumento del precio de las matrículas de las universidades públicas catalanas aumentara la tasa de abandono de los estudiantes universitarios de nivel socioeconómico bajo. García-Montalvo (2020) muestra también los resultados de una encuesta representativa entre los estudiantes de universidades públicas catalanas sobre su percepción del nivel de los precios realizada en 2017. Un 36,6% de los estudiantes universitarios catalanes tienen la percepción de que el precio de la matrícula es muy alto, mientras el 41,5% indican que es alto. No obstante, sólo un 21,9% de los encuestados aciertan que el precio de la matrícula cubre menos del 25 por 100 del coste. Lo más importante de esta encuesta es que el 80,9% de los estudiantes encuestados se muestran bastante o totalmente de acuerdo con que el precio de la matrícula debe ser diferente en función de la renta. De hecho, García-Montalvo (2020) muestra que un 81,2% de los encuestados creen que la exención completa de matrícula no debería producirse por encima de los 30.000 euros de renta para familias de cuatro miembros, cuando esa exención es en realidad algo más elevada (hasta 38.831 euros).

## 3. INDICADORES DE RESULTADOS EDUCATIVOS

### 3.1 Tasa de abandono temprano educativo

La tasa de abandono es la media de 4 trimestres del año del porcentaje de población entre 18-24 años que no ha finalizado Bachillerato, FP Media o FP Básica y, que, no habiendo titulado en estos tres niveles educativos, no ha realizado un curso de formación en las últimas cuatro semanas. En 2002, nuestro país ya tenía una tasa de abandono del 30,9%, 14 puntos porcentuales más que la UE. De hecho, desde 2008 cuando la tasa de abandono educativo temprano era del 31,7, España había encadenado 13 años seguidos de disminución en este indicador hasta situarse en el 13,3% a 3,5 puntos porcentuales de distancia de la UE. Esa disminución del gap entre España y la UE se ha paralizado en 2022. Los datos oficiales del Ministerio de Educación y FP muestran que la tasa de abandono en España ha repuntado del 13,3% de 2021 al 13,9% en 2022.

España es en 2022 el segundo país de la UE con más jóvenes que no continúan sus estudios, tan solo mejor que Rumanía (15,6%) y a cierta distancia del antepenúltimo (Hungría, 12,4%). La media de la UE-27 en 2022 fue del 9,6%. La tasa de abandono del 13,9% en España en 2022 nos aleja del objetivo de la UE en Educación y Formación (Education and Training Monitor Report de la UE) para 2030 de situarse en el 9%.

**Tabla 15. Abandono temprano de la educación-formación en la UE.**

| | Unión Europea (27 países) | España |
|---|---|---|
| 2022 | 9,6 | 13,9 |
| 2021 | 9,8 | 13,3 |
| 2020 | 9,9 | 16,0 |
| 2019 | 10,2 | 17,3 |
| 2018 | 10,5 | 17,9 |
| 2017 | 10,5 | 18,3 |
| 2016 | 10,6 | 19,0 |
| 2015 | 11,0 | 20,0 |
| 2014 | 11,1 | 21,9 |
| 2013 | 11,8 | 23,6 |
| 2012 | 12,6 | 24,7 |
| 2011 | 13,2 | 26,3 |
| 2010 | 13,8 | 28,2 |
| 2009 | 14,0 | 30,9 |
| 2008 | 14,4 | 31,7 |
| 2007 | 14,7 | 30,8 |
| 2006 | 15,2 | 30,3 |
| 2005 | 15,6 | 31,0 |
| 2004 | 16,0 | 32,2 |
| 2003 | 16,4 | 31,7 |
| 2002 | 16,9 | 30,9 |

Fuente: Ministerio de Educación y FP. Nivel de formación, formación permanente y abandono: explotación de las variables educativas de la Encuesta de Población Activa https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/laborales/epa.html

La información del Ministerio de Educación y FP (Nivel de formación, formación permanente y abandono: explotación de las variables educativas de la Encuesta de Población Activa) muestra, además, que la tasa de abandono de jóvenes cuyas madres tienen estudios superiores es del 3,1%, 11 veces más baja que el 36,3% de la tasa de abandono de los jóvenes cuyas madres solo tiene estudios de Primaria o inferior. El capítulo 5b de este libro analiza con detalle la evolución de la desigualdad en el abandono educativo en España por nivel socioeconómico de los padres. El abandono temprano educativo en 2018, con datos de la Encuesta de Condiciones de Vida de ese año, era del 33,5% entre los jóvenes de hogares del último quintil de renta per cápita por el 6% de los jóvenes de hogares más acomodados (IVIE, y Fundación Europea Sociedad y Educación, 2021).

La resistencia a bajar el abandono desde este umbral del rango del 13%-14% va a requerir nuevas medidas educativas. Con la pandemia, los indicadores educativos de promociones mejoraron mucho y esa inercia inicial se ha perdido. Una potente oferta educativa podría evitar que estudiantes que, si no hubiera sido por la pandemia habrían dejado los estudios, abandonarán un poco más tarde. Es posible que los jóvenes a los que en mayor medida les afecta la tasa de abandono sean ya de colectivos de características singulares a los que sea más difícil llegar y a los que habrá que ayudar con medidas muy bien dirigidas, bien analizadas y evaluadas de forma constante. Para reducir la tasa de abandono en España, hay que centrarse en los colectivos a los que en mayor medida afecta el fracaso escolar, como alumnado inmigrante o de etnia gitana. Una medida educativa que ha tenido buenos resultados sería contratar jóvenes de esos colectivos con estudios superiores (ya sean de FP de Grado Superior o universitarios) para que sean tutores remunerados en centros educativos de Primaria y Secundaria con mucha presencia de estudiantes del mismo origen. Battaglia y Lebedinski (2022) muestran que los alumnos de bajo rendimiento pueden obtener mejores resultados académicos siguiendo los pasos de personas que puedan ser una referencia. En concreto, estos autores estudian el Programa de Asistentes de Enseñanza en Serbia para alumnos de etnia gitana de bajo rendimiento. Asigna una persona de esta etnia a cada centro participante para brindar apoyo a los alumnos seleccionados y establecer un puente con su comunidad. Battaglia y Lebedinski (2022) estiman efectos a medio plazo en los resultados educativos al final de la escuela primaria comparando estudiantes en centros educativos participantes y en centros que solicitaron, pero no fueron seleccionadas, antes y después de la introducción del programa. En el caso de EE. UU., Gershenson et al. (2022) utilizan el experimento de tamaño de clase STAR en Tennessee del que se habla en el capítulo 3a, y muestran que los estudiantes negros asignados al azar a al menos un profesor negro en el tercer curso de Primaria

tienen 9 puntos porcentuales más de probabilidades de graduarse en Secundaria y 6 puntos porcentuales más de probabilidades de acceder en la Universidad en comparación con sus compañeros negros que no tienen profesores negros. Los resultados de educación Postsecundaria se centran en las universidades de dos años y se concentran en varones desfavorecidos. Los programas de tutorías en pequeños grupos, de los que realiza una revisión en el capítulo 3a, podrían escalarse a más alumnos y por más cursos. En su metaanálisis, Nickow et al., (2020) muestran, a partir del estudio de 96 artículos de experimentos aleatorios, que el efecto de las tutorías en pequeños grupos es importante y significativo (37% de la desviación estándar).

El problema de la tasa de abandono en España proviene ya de la ESO. Un 26,5% de los jóvenes españoles entre 25-34 años había finalizado, como mucho, la ESO en 2022. Esta proporción es más del doble del 11,8% de la UE. Y aun así ese 26,5% es el porcentaje más bajo de toda la serie histórica, que comenzó en 2002 con un 40,3% de población entre los 25-34 años que no había titulado ni en FP Básica, Media o Bachillerato. No finalizar la ESO es el primer paso para terminar abandonando la formación de forma definitiva, porque cuando un joven comienza a desvincularse del sistema educativo resulta muy difícil volver a implicarle y atraerle de nuevo a los estudios reglados. El primer paso para reducir el abandono debe pasar por el incremento de los jóvenes que finalizan la ESO. Al mismo tiempo que aumenta el abandono educativo en nuestro país, se incrementa también la proporción de jóvenes españoles de 25-34 años con estudios superiores finalizados (del 48,7% de 2021 al 50,5% de 2022, confirmando que España en uno de los países lideres en formación terciaria). Es decir, los datos de 2022 muestran que tenemos más jóvenes con baja y alta formación y menos, por lo tanto, con formación media. No se trata de que haya un exceso de jóvenes con estudios superiores, pero sí de que a España le faltan jóvenes con formación media (con hasta Bachillerato o FP Media). Necesitamos disminuir la proporción de la población que solo alcanza la ESO, para aumentar los que acceden y terminan FP Media o Bachillerato. El capítulo 4b en este libro realiza un análisis en profundidad sobre la FP, y muestra que en los últimos años más jóvenes españoles terminan la ESO y que se está produciendo un incremento en el porcentaje de estudiantes que elige FP Media, aunque siga siendo minoritaria con respecto a los que escogen Bachillerato. La Educación Secundaria superior es el nivel de estudios mínimo que se requiere para afrontar con garantías la inserción laboral y para continuar con la educación superior. Aquellos que no completan este nivel de estudios medios tienen más probabilidades de no estar ni empleados ni en la enseñanza después (ninis) con un coste importante tanto para ellos como para la sociedad.

La tabla 16 muestra la evolución de la tasa de abandono temprano en la educación. formación por Comunidades Autónomas. Hay una gran heterogeneidad en el abandono, con País Vasco (5,6%), Navarra (5,7%) o Cantabria (8,9%) ya con registros por debajo del objetivo de la UE para 2030 del 9,0% y Galicia (9,9%) y Castilla León (9,9%) en camino de este propósito. Otras Comunidades como Murcia (18,7%), Baleares (18,2%) se encuentran más lejos del promedio de la UE del 9,6% e incluso de la media de España (13,9%). En todas las Comunidades Autónomas, el abandono temprano de la educación es más baja entre las mujeres que entre los hombres, en algunos casos incluso la mitad como en Aragón (6,9% por 15,5%), Extremadura (6,9% por 14,7%), Canarias (7,3% por 16,0%) o Navarra (3,5% por 7,9%). Tan sólo en Asturias (11,2% por 11,8%) y Cantabria (8,3% por 9,5%) la diferencia es de menos del 15%.

**Tabla 16. Abandono temprano de la educación por Comunidad Autónoma y sexo (2022).**

| | Total | Mujeres | Hombres |
|---|---|---|---|
| España | 13,9 | 11,2 | 16,5 |
| Andalucía | 15,3 | 12,0 | 18,4 |
| Aragón | 11,4 | 6,9 | 15,5 |
| Asturias, Principado de | 11,5 | 11,2 | 11,8 |
| Balears, Illes | 18,2 | 14,3 | 22,1 |
| Canarias | 11,7 | 7,3 | 16,0 |
| Cantabria | 8,9 | 8,3 | 9,5 |
| Castilla y León | 9,9 | 7,8 | 12,0 |
| Castilla-La Mancha | 15,1 | 12,5 | 17,5 |
| Cataluña | 16,9 | 14,2 | 19,5 |
| Comunitat Valenciana | 15,7 | 13,1 | 18,0 |
| Extremadura | 10,8 | 6,9 | 14,7 |
| Galicia | 9,9 | 8,4 | 11,4 |
| Madrid, Comunidad de | 13,2 | 11,8 | 14,7 |
| Murcia, Región de | 18,7 | 13,4 | 23,5 |
| Comunidad Foral Navarra | 5,7 | 3,5 | 7,9 |
| País Vasco | 5,6 | 4,2 | 7,0 |
| La Rioja | 11,3 | 7,0 | 15,7 |
| Ceuta | 15,6 | – | – |
| Melilla | 12,7 | – | – |

Fuente: Ministerio de Educación y FP: Nivel de formación, formación permanente y abandono: explotación de las variables educativas de la Encuesta de Población Activa https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/laborales/epa.html
Nota: Los datos de Ceuta, Melilla, Navarra, La Rioja y Cantabria deben ser tomados con precaución, pues los derivados de tamaños muestrales pequeños están afectados por fuertes errores de muestreo.

### 3.2. Tasas de graduación en ESO, FP y Bachillerato

España tiene una elevada tasa de abandono por la falta de jóvenes que hayan titulado Secundaria posobligatoria: FP Media o Bachillerato (y desde la LOMCE, la FP Básica). Como se muestra en la tabla 17, la tasa bruta de graduados en la ESO registró un incremento constante desde el curso 2006-2007 hasta alcanzar el 79,3% en el 2015-2016. La tasa bruta de graduación en ESO se define como la relación entre el alumnado que termina con éxito esta etapa educativa y es propuesto para el título de Graduado en ESO, independientemente de su edad, y el total de la población de la "edad teórica" de comienzo del último curso de esta etapa, que es 15 años (Sistema Estatal de Indicadores Educativos, 2023a).

Sin embargo, en el curso 2016-2017 se produjo una reducción importante en el porcentaje de jóvenes que se graduaba en la ESO, hasta retroceder al 75,5%. En 2017-2018 se recuperaron 2,3 puntos de ese retroceso de 3,8 puntos del curso previo. Entre las causas de ese importante retroceso en 2016-2017 se encontraba, en primer término, que fue el primer curso en el que la primera cohorte de alumnos del Programa de Mejora del Aprendizaje y Rendimiento (PMAR, un itinerario para alumnos rezagados que comienza en 2º de la ESO) se reincorporaron a sus grupos ordinarios. Después de haber estado dos cursos, los de 2º y 3º de la ESO en un itinerario diferente, los alumnos de PMAR se reincorporan a 4º de la ESO con el resto de sus compañeros. Hasta 2015-2016 existía un itinerario, denominado diversificación, que los estudiantes que lo seguían realizaban en 3º y 4º de la ESO. Esta es una de las diferencias entre el PMAR y la diversificación: en el primer caso, los alumnos se reincorporan con todos sus compañeros ya desde el inicio de 4º de la ESO, mientras que, en el segundo, continúan con su itinerario hasta el final de la ESO. En segundo término, el retroceso en la graduación en 4º ESO en 2016-2017 se debe a que, en los cursos previos (desde 1º de la ESO a 3º de la ESO), las asignaturas de libre configuración autonómica no contaron para determinar si un alumno promocionaba de curso. En 2016-2017 fue la primera vez en la que alumnos con varias asignaturas de libre configuración pendientes de 1º de la ESO a 3º de la ESO llegaban a 4º de la ESO. La diferencia es que, para obtener el título de la ESO al finalizar 4º, las asignaturas de libre configuración sí que contabilizaban. Así que muchos estudiantes que en otras circunstancias hubieran repetido un curso entre 1º y 3º de la ESO, lo hicieron en su lugar en 4º de la ESO. La tasa de graduación en la ESO de 2018-2019 continuó con la recuperación del terreno perdido hasta alcanza el 78,8%.

También se produjo una tendencia alcista de la tasa bruta de graduados en Bachillerato y FP Media desde 2006-2007 hasta el curso 2016-2017 año en el que, como en el caso de la ESO, se produjo una ruptura de la trayectoria de expansión. La tasa bruta de graduación en FP Media y Bachillerato se define como la relación entre el alumnado que termina con éxito esta etapa educativa, independientemente de su edad, y el total de la población de la "edad teórica" de comienzo del último curso de esta etapa, que es 17 años (Sistema Estatal de Indicadores Educativos, 2023a). El cálculo de la tasa bruta de titulación en los Ciclos Formativos de Grado Superior se computa sobre la población de 19 años. En los siguientes tres cursos hasta el 2018-2019, antes del surgimiento del Covid-19, los dos indicadores se mantuvieron estables en el entorno del 55% para el Bachillerato, y 23% de los Ciclos Formativos de Grado Medio. La FP de Grado Superior, también ha experimentado un crecimiento en la tasa bruta de graduación desde el curso 2006-2007 hasta alcanzar en el curso 2018-2019 el 29,3%. El número de titulados en FP Superior (independientemente de su edad) supone un 29,3% de la población de 19 años. FP Básica llegó a registrar el 4,1% de tasa bruta de graduación en 2018-2019, que fue justamente la misma tasa que registró su predecesor de Programas de Cualificación Profesional Inicial en su último año en vigor de 2014-2015.

En suma, antes del surgimiento del Covid, la tasa bruta de graduación de todos los niveles educativos experimentó un crecimiento desde año anterior al inicio de la crisis financiera (2006-2007) hasta el 2016-2017. A partir de ese año se produjo un estancamiento en FP Básica, FP Media, Bachillerato y FP Superior (con un vaivén en la ESO por motivos normativos) hasta que llegó el curso 2019-2020. La tasa bruta de graduados en la ESO aumentó entre 2018-2019 a 2019-2020 del 78,8% al 84,0% en 2020-2021, en alguna medida por la flexibilización acometida en el sistema educativo español, pero ha vuelto a retroceder a 82,1% para el curso 2020-2021 con los datos conocidos en noviembre de 2022. Las tasas brutas de graduación en FP Básica, Media y Superior no aumentaron el año que surgió el Covid, el 2019-2020, sino al siguiente en el 2020-2021. Los datos que se publiquen en noviembre de 2023 mostrarán si el importante aumento en la titulación se confirma durante el curso 2021-2022. Puede ser que una parte del incremento en el 2020-2021 se debiera a que algunos estudiantes retrasaron la realización de la Formación en Centros de Trabajo (FCT) durante el 2019-2020 por el surgimiento de la pandemia y la dificultad de realización de las prácticas (aunque se dio la oportunidad de sustituirlas por un trabajo). En cualquier caso, las tasas de graduación en FP están en máximos históricos y aunque haya algo de corrección en el curso 2021-2022, la tendencia es positiva. En

Bachillerato, sí se ha corregido todo el aumento circunstancial que se produjo en la tasa bruta de graduación desde el 55,1% en 2018-2019 al 61,0% en 2019-2020 para retornar de nuevo al 55,0% en el 2020-2021. La trayectoria de los graduados en la ESO a través de las ofertas de Educación de Adultos (Educación Secundaria para personas adultas y Pruebas libres), ha sido descendente, lo que se explica por el aumento de los titulados en la ESO en los Institutos. El rango de edades de estos titulados es bastante amplio, aunque el peso de los jóvenes que han dejado recientemente la etapa obligatoria sea importante (Ministerio de Educación y FP: Estadística de las Enseñanzas no universitarias Resultados académicos del curso 2020-2021).

**Tabla 17. Tasa bruta de graduación por niveles educativos.**

| | ESO | ESO Adultos | FP Básica | FP Media | Bachillerato | FP Superior |
|---|---|---|---|---|---|---|
| 2020-2021 | 82,2 | 6,7 | 4,6 | 27,0 | 55,0 | 35,3 |
| 2019-2020 | 84,0 | 7,5 | 4,3 | 23,6 | 61,0 | 27,3 |
| 2018-2019 | 78,8 | 7,1 | 4,1 | 23,3 | 55,1 | 29,3 |
| 2017-2018 | 77,8 | 7,6 | 4,1 | 22,8 | 55,4 | 29,3 |
| 2016-2017 | 75,5 | 7,9 | 3,8 | 23,5 | 55,7 | 26,7 |
| 2015-2016 | 79,3 | 9,3 | 3,1 | 24,7 | 57,4 | 27,8 |
| 2014-2015 | 77,6 | 11,3 | – | 24,7 | 55,0 | 28,5 |
| 2013-2014 | 76,8 | 12,1 | – | 23,9 | 54,8 | 27,3 |
| 2012-2013 | 75,4 | 13,2 | – | 21,7 | 53,4 | 24,0 |
| 2011-2012 | 75,1 | 12,7 | – | 22,4 | 52,2 | 22,6 |
| 2010-2011 | 74,7 | 13,0 | – | 19,9 | 49,7 | 21,2 |
| 2009-2010 | 74,6 | 11,1 | – | 18,3 | 47,2 | 18,8 |
| 2008-2009 | 73,0 | 8,4 | – | 17,3 | 45,4 | 17,4 |
| 2007-2008 | 69,4 | 5,9 | – | 16,7 | 44,5 | 16,1 |
| 2006-2007 | 69,0 | 4,7 | – | 16,4 | 44,3 | 16,3 |
| 2005-2006 | 69,2 | 4,3 | – | 16,7 | 44,8 | 16,8 |
| 2004-2005 | 69,8 | 3,6 | – | 16,4 | 44,5 | 16,9 |
| 2003-2004 | 71,3 | 3,5 | – | 16,3 | 45,2 | 17,0 |
| 2002-2003 | 71,9 | 2,7 | – | 15,4 | 46,0 | 15,7 |
| 2001-2002 | 71,9 | – | – | 13,5 | 45,3 | 17,2 |
| 2000-2001 | 73,4 | – | – | 10,9 | 46,8 | 15,7 |
| 1999-2000 | 73,4 | – | – | 9,1 | 45,1 | 17,6 |

Fuente: Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional. https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/resultados/series.html

### 3.3. Tasas de repetición en Primaria y Secundaria en España en comparación con OCDE y UE

Otro de los factores que podría ser un indicador adelantado de la tasa de abandono educativo es la repetición. Un estudiante que repite algún curso podría estar comenzando el camino que le conduzca a dejar los estudios años más tarde. La tabla 18 muestra la evolución de la tasa de repetición en Primaria y ESO en España desde 2007-2008 hasta 2021-2022. Esta tasa se define como el porcentaje de alumnos de una etapa educativa (o curso) que ha repetido el año anterior con respecto al total de matrícula. En concreto, la definición del Sistema Estatal de Indicadores Educativos (2023a) es "porcentaje de alumnado repetidor se calcula para cada curso de enseñanza relacionando porcentualmente su alumnado repetidor con su matrícula del año académico anterior". La tasa de repetición en Primaria en España ha mostrado una tendencia decreciente desde el 2,8% de 2007-2008 al 2,1% de 2021-2022. Ha habido dos rupturas en esta trayectoria. La primera en 2015-2016, el primer año en el que, de acuerdo con la LOMCE, los alumnos podían repetir también en los cursos impares de Primaria. Este impacto inicial se reabsorbió en los años siguientes. Es decir, el estudiante que repetía en 4º de Primaria, por ejemplo, antes de la LOMCE pasó a repetir en 3º de Primaria. La segunda ruptura en la serie se produjo como consecuencia de la pandemia del Covid cuando la flexibilización en el curso 2019-2020 produjo que la proporción de repetidores en el curso siguiente de 2020-2021 se redujera desde el 2,3% al 1,2%. Al año siguiente, la tasa de repetición retornó en buena medida a la senda original, con un 2,1%. Habrá que esperar hasta julio de 2024 para ver si se confirma la tendencia decreciente en los repetidores de Primaria o se estanca la caída de los últimos años. También la tasa de repetición en la ESO ha mantenido una trayectoria descendente desde el 14,7% del curso 2007-2008 hasta el 7,6% de 2021-2022. Además de la excepción del curso 2017-2018, que se corresponde la caída ya comentada de la tasa bruta de graduación el año anterior de 2016-2017, la repetición en la ESO también disminuyó el año siguiente a la irrupción de la pandemia del Covid, desde el 8,5% al 4,2%, una disminución que se ha corregido en buena medida en el 2021-2022. Los datos de julio de 2024 corroborarán si la tasa de repetición continúa con la corrección o se mantiene en el entorno del 7%. El indicador de la tasa de idoneidad a los 15 años refleja "el porcentaje de alumnado que se encuentra matriculado en el curso teórico correspondiente a su edad". La tasa de idoneidad a los 15 años está "relacionada con cuarto de Educación Secundaria Obligatoria, estando la edad de 15 años también relacionada

**Tabla 18. Tasa de repetición por niveles educativos.**

| | Tasa de repetición Primaria | Tasa de repetición ESO | Tasa de idoneidad 15 años |
|---|---|---|---|
| 2021-2022 | 2,1 | 7,6 | 75,5 |
| 2020-2021 | 1,2 | 4,2 | 75,5 |
| 2019-2020 | 2,3 | 8,5 | 71,4 |
| 2018-2019 | 2,4 | 8,7 | 70,8 |
| 2017-2018 | 2,5 | 9,5 | 69,6 |
| 2016-2017 | 2,5 | 8,8 | 68,6 |
| 2015-2016 | 3,0 | 10,2 | 67,5 |
| 2014-2015 | 2,2 | 10,9 | 65,8 |
| 2013-2014 | 2,3 | 11,2 | 63,6 |
| 2012-2013 | 2,5 | 11,8 | 62,5 |
| 2011-2012 | 2,4 | 12,4 | 61,7 |
| 2010-2011 | 2,6 | 12,7 | 61,2 |
| 2009-2010 | 2,7 | 13,3 | 59,6 |
| 2008-2009 | 2,8 | 13,7 | 58,7 |
| 2007-2008 | 2,8 | 14,7 | 57,7 |
| 2006-2007 | – | – | 57,4 |
| 2005-2006 | – | – | 57,7 |
| 2004-2005 | – | – | 58,4 |
| 2003-2004 | – | – | 59,4 |
| 2002-2003 | – | – | 60,5 |
| 2001-2002 | – | – | 62,0 |
| 2000-2001 | – | – | 63,2 |
| 1999-2000 | – | – | 63,8 |
| 1998-1999 | – | – | 64,6 |
| 1997-1998 | – | – | 61,9 |
| 1996-1997 | – | – | 60,2 |
| 1995-1996 | – | – | 58,6 |
| 1994-1995 | – | – | 58,0 |
| 1993-1994 | – | – | 58,4 |
| 1992-1993 | – | – | 60,6 |

Fuente: Estadística de las Enseñanzas no universitarias. Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional. https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado/series.html

con el inicio de la FP Básica" (Sistema Estatal de Indicadores Educativos, 2023a). La tasa de idoneidad a los 15 años ha encadenado 14 cursos consecutivos de descensos, desde el 57,4% de 2006-2007 hasta el 75,5% de 2020-2021. Además, este indicador no ha caído en el curso 2021-2022 con respecto a 2020-2021, quizá porque la tasa de idoneidad a los 15 años recoge el acumulado de lo ocurrido en la repetición de los años anteriores, y la caída previa de las tasas de repetición se compensó con la corrección en las medidas de flexibilización en la promoción de los alumnos en el curso del Covid de 2019-2020.

Maurin y McNally (2008, https://bit. ly/2IwBYfT) analizan el "experimento natural" del efecto de las movilizaciones de mayo de 1968 en Francia en el futuro de la cohorte de estudiantes que se preparaban para hacer el examen de baccalauréat. Los exámenes no tuvieron en ese año el mismo nivel de exigencia que habían tenido antes y tampoco el que tuvieron después. Como resultado, la tasa de aprobados aumentó significativamente permitiendo que una proporción mayor de estudiantes nacidos en 1948 y 1949 pudiera alcanzar más años de escolarización. Los alumnos que se encontraban en el margen, que accedieron a la Universidad en ese año y que no hubieran alcanzado los estudios superiores en otras circunstancias, terminaron en buena medida sus estudios y alcanzaron en su carrera profesional salarios y puestos de responsabilidad mejores que si no hubieran aprobado la Selectividad. Este mayor progreso educativo de los jóvenes de esas cohortes se ha transmitido también a sus hijos. El estudio de Maurin y McNally analiza la flexibilización del acceso a la Universidad en un solo año y que, además, como en el caso de la Covid-19 fue sobrevenida. Habría que comprobar si una reducción de los niveles de exigencia permanente en el tiempo y que los alumnos puedan anticipar pudiera tener efectos negativos en la dedicación y esfuerzo de los alumnos.

En términos de comparación internacional, España tenía en 2019, el último año con datos comparables un 8,7% de estudiantes repetidores en ESO, en comparación con el 1,9% y en promedio en la OCDE. Los chicos tienen más probabilidades de repetir un curso que las chicas (OCDE, 2021).

### 3.4. Resultados de los estudiantes españoles en pruebas internacionales de Primaria. PIRLS y TIMSS

PIRLS es una evaluación internacional de Lectura para estudiantes de cuarto de Primaria. Esta prueba se realiza cada cinco años. Los resultados publicados en

mayo de 2023 por la IEA, una organización internacional con base en Países Bajos dedicada a la educación y con participación de los gobiernos, indican que España ha disminuido la competencia lectora de los estudiantes de 4º de Primaria entre 2016 y 2021. Los promedios globales de los países participantes se expresan en una escala continua con un punto de referencia central de 500 puntos, que fue fijado en 1995, y desviación típica 100 (Ministerio de Educación y FP: Informe PIRLS español). El estudio PIRLS utiliza la teoría de respuesta al ítem (TRI) para asignar las puntuaciones a los estudiantes, que expresa el nivel de rendimiento en una escala común a la empleada para estimar la dificultad de las preguntas de la prueba. Los estudiantes de cuarto de Primaria de España obtuvieron 521 puntos en competencia lectora en 2021, siete puntos menos que los 528 obtenidos en 2016. La competencia lectora de los estudiantes españoles en 2021 está por debajo del promedio de la UE (528) y la OCDE (533). España no es el único país en donde los alumnos han bajado sus competencias en Lectura entre 2016 y 2021. De los 32 países para los que se dispone de datos comparables de los resultados de PIRLS en 2016 y 2021, hay 21 que también han experimentado una disminución en los resultados en competencia lectora. Esta caída generalizada en los resultados se debe en gran medida al cierre de las escuelas que ocurrió a partir de marzo de 2020 debido a la pandemia de la Covid en todo el mundo. Los estudiantes de muchos países no aprendieron tanto con la educación en línea como lo hubieran hecho en un curso ordinario con la enseñanza presencial. Patrinos et al. (2023) muestran que la mayoría de los 36 estudios robustos que han analizado, encuentran que durante los cierres de los centros educativos por la pandemia se produjeron pérdidas de aprendizaje que ascienden, en promedio, a 0,17 de una desviación estándar, equivalente a aproximadamente medio año escolar de aprendizaje.

Hay 11 de los 32 países con datos comparables que han logrado que sus estudiantes no se vean afectados en su competencia lectora por el cierre de las escuelas durante el Covid, aunque la evaluación en algunos de ellos como Inglaterra se realizó a la promoción siguiente un año más tarde, con la ventaja de que el efecto la pandemia podría haberse comenzado a diluir.

**Tabla 19. Resultados de PIRLS en 2006, 2011, 2016 y 2021 en España y la OCDE.**

| | 2006 | 2011 | 2016 | 2021 |
|---|---|---|---|---|
| España | 513 | 513 | 528 | 521 |
| OCDE | 533 | 538 | 540 | 533 |

Fuente: IEA's Progress in International Reading Literacy Study - PIRLS 2021. https://pirls2021.org/results

España es el país de todos los evaluados, donde menos diferencia hay en Lectura entre las alumnas y los alumnos. Sólo dos puntos separan en 2021 a las chicas de cuarto de Primaria españolas con respecto a los chicos españoles, cuando la diferencia era de 8 puntos en 2016. Y ese acercamiento en los resultados entre alumnas y alumnos no se ha producido por la mejora de los chicos, sino por una mayor disminución en la competencia lectora de las chicas. Las directoras y directores de los colegios educativos españoles creen que los alumnos empiezan Primaria con competencias lectoras (PIRLS 2021: Exhibit 6.1. Schools Where Students Begin the Primary Grades With Literacy Skills). El 51% de las directoras y directores de los centros educativos españoles señalan que más del 75% de los estudiantes comienza la nueva etapa educativa con competencias lectoras, el segundo país con más responsables con esa opinión (tan solo por detrás de Singapur) y muy por encima del promedio internacional de 26% de responsables de colegios que indica la mayoría de los alumnos tienen esas competencias. También más madres y padres españoles tienen la percepción de que los niños de 6 años terminan Educación Infantil con buenas competencias lectoras. Es decir, que parece que nuestros alumnos de Primaria comienzan Primaria con una ventaja en competencia lectora con respecto a otros países, pero esa ventaja inicial se pierde solo 3 o 4 años después (PIRLS 2021: Exhibit 5.2. Could Do Early Literacy Tasks When Beginning Primary School) Podría inferirse que España no aprovecha bien la extensión de su Educación Infantil.

Aunque España tiene unas elevadas tasas de matriculación tanto en Educación Infantil de primer ciclo (0-2 años) como de segundo ciclo (3-5 años), los resultados del alumnado de nuestro país en las pruebas internacionales de 4º de Primaria de PIRLS (Lectura) y TIMSS (Matemáticas y Ciencias) no son positivos. Los niños españoles que asisten a Educación Infantil de primer ciclo tienen mejores resultados que los que no se matriculan en esta etapa, incluso después de tener en cuenta el nivel socioeconómico de los padres, pero el impulso inicial de la educación española se pierde ya en 4º de Primaria como muestran los recientes resultados de PIRLS 2021. Uno de los aspectos que podrían explicar este sorprendente resultado es la calidad de los programas de Educación Infantil: los tipos de servicios de Educación Infantil disponibles, los procesos que se desarrollan, tamaño de los grupos, cualificación del personal y el número de horas por semana que los niños suelen asistir. La calidad del proceso se refiere a las interacciones diarias que ocurren entre los niños y su entorno,

incluyendo sus relaciones con sus compañeros, el personal, las familias, las comunidades y el entorno físico (Cadima et al., 2020 y OCDE Education at a Glance, 2022). Algunas áreas de aprendizaje suelen incluirse en los marcos curriculares diseñados para guiar el desarrollo de los niños que participan en Educación Infantil. OCDE (2021) muestra que en los marcos curriculares de educación infantil de 26 países de la OCDE revisados el respeto hacia los demás se especificaba como un área de aprendizaje en cada currículo, y la mayoría también incluía la alfabetización/lenguaje oral, cooperación, respeto por la diversidad, juego, expresión y apreciación artística, y bienestar físico. Identificar qué aspectos de los servicios de Educación Infantil constituyen una provisión de alta calidad es, por lo tanto, de gran interés para las políticas. En España, por cada educador en Educación Infantil hay 12,1 niños, situándose por encima de la media de la UE-22 (11,7 niños/educador), pero por debajo de la media de la OCDE (12,7 niños/educador). La investigación reciente muestra que unos entornos enriquecedores y estimulantes, además de una pedagogía de alta calidad, se fomentan a partir de un personal mejor cualificado (Araujo et al., 2016). A su vez, unas interacciones de mejor calidad entre los niños y el personal facilitan mejores resultados de aprendizaje. Por ello, la calidad de los programas de Educación Infantil se vincula más que en otras etapas a las ratios entre el personal de los centros y la cantidad de niños y niñas.

La disminución en los resultados en España se ha sentido de manera más intensa en el extremo inferior de la distribución. Del 3% de estudiantes españoles que se encontraban en el nivel muy bajo en PIRLS 2016, se ha pasado a un 5% en PIRLS 2021. El alumnado que se encuentra en el nivel bajo de competencia lectora también ha aumentado del 17% al 19% en España mientras que el se situaba en el nivel alto disminuía desde el 33% al 29%. Lo estudiantes en los niveles intermedios (41%) y avanzado (6%) han permanecido estables en España entre PIRLS 2016 y PIRLS 2021 (IEA, 2023; INEE, 2023). Alemania e Italia también han experimentado un mayor impacto del Covid en las competencias lectoras de los estudiantes rezagados, aunque en su caso también ha afectado a la proporción de alumnos excelentes. Francia ha mejorado sus resultados entre PIRLS 2016 y PIRLS 2021, mientras que en Inglaterra el Covid ha reducido el porcentaje de alumnos excelentes sin alterar en de los rezagados. Por lo demás, la tabla 20 muestra también que España tiene una proporción de estudiantes excelentes más reducida que el resto de los países, con la excepción de Francia, una característica que también se reproduce tanto en TIMSS como en PISA.

**Tabla 20. Tendencias en la Distribución de Estudiantes que Alcanzan los Puntos de Referencia Internacionales de PIRLS.**

| | Durante la pandemia del Covid-19 | | | | | Antes de la pandemia del Covid-19 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2021 | | | | | 2016 | | | | |
| | Avanzado (>625) | Alto (550-625) | Medio (475-550) | Bajo (400-475) | Muy Bajo (<400) | Avanzado (>625) | Alto (550-625) | Medio (475-550) | Bajo (400-475) | Muy Bajo (<400) |
| Francia | 5 | 27 | 40 | 22 | 6 | 4 | 26 | 42 | 22 | 6 |
| Alemania | 8 | 31 | 36 | 19 | 6 | 11 | 36 | 34 | 14 | 5 |
| Italia | 8 | 36 | 39 | 14 | 3 | 11 | 41 | 35 | 11 | 2 |
| España | 6 | 29 | 41 | 19 | 5 | 6 | 33 | 41 | 17 | 3 |
| Inglaterra (1) | 18 | 39 | 29 | 11 | 3 | 20 | 37 | 29 | 11 | 3 |

Fuente: IEA's Progress in International Reading Literacy Study - PIRLS 2021. https://pirls2021.org/results. 1. Inglaterra fue evaluado un año más tarde de lo originalmente programado: tendencia de seis años desde PIRLS 2016. Los datos están extraídos de https://pirls2021.org/results y son muy similares, aunque no totalmente coincidentes a los de PIRLS 2021. Informe Español (p. 57).

TIMSS es una evaluación también llevada a cabo por el organismo internacional de la IEA, con la diferencia de que se centra en Matemáticas y Ciencias. Analiza las competencias de los estudiantes de 4º grado (4º de Primaria en España) y 8º grado (2º de la ESO) aunque nuestro país solo participa en las pruebas de Primaria. La última prueba se desarrolló en abril-mayo de 2019 y los resultados se publicaron en diciembre de 2020. La próxima evaluación será la de TIMSS 2024 y los resultados se conocerán en diciembre de 2025. Los resultados de la prueba de TIMSS 2019 de 4º de Primaria muestran que Matemáticas es la competencia en la que peores resultados tiene España. En Matemáticas la puntuación de España ha sido de 502, similar a los 505 de 2015 y mejor que los 482 de 2011. En Ciencias, España ha obtenido 511 puntos, peor que los 518 de 2015 y mejor que los 505 de 2011. Es decir, que los resultados de nuestro país en Primaria en Ciencias y Matemáticas son algo peores que los de 2015 pero siguen siendo mejores que los registrados en 2011. En cuanto a la comparación por materias, nuestro país se encuentra mejor en términos relativos al promedio internacional en Lectura (521 puntos en Lectura como acabamos de ver en PIRLS 2021) y Ciencias (511, TIMSS 2019) que en Matemáticas (502, TIMSS 2019). De hecho, España está significativamente por debajo del promedio de la UE y la OCDE tanto en Ciencias como (sobre todo) en Matemáticas.

**Tabla 21. Resultados de TIMSS en 2011, 2015, 2019.**

| | TIMSS | 2011 | 2015 | 2019 |
|---|---|---|---|---|
| Matemáticas | España | 482 | 505 | 502 |
| | OCDE | 522 | 525 | 527 |
| Ciencias | España | 505 | 518 | 511 |
| | OCDE | 523 | 528 | 526 |

Fuente: Ministerio de Educación y FP. TIMSS 2019. Informe Español https://www.educacionyfp.gob.es/inee/evaluaciones-internacionales/timss/timss-2019.html

En el contexto internacional, los países de Asia (Singapur, China Taipéi, Corea, Japón y Hong Kong) consolidan su ventaja con respecto a Europa, América y Oceanía en Matemáticas y Ciencias ya desde Primaria. Singapur lidera Matemáticas, con el resto de los países asiáticos liderando con grandes diferencias con respecto al resto de participantes. Singapur, Taiwán, Corea y Japón también obtuvieron buenos resultados en Ciencias, aunque en este caso, la Federación Rusa y Finlandia, les siguen a menos distancia. Los países asiáticos tienen también muchos alumnos excelentes, que entienden y saben aplicar conocimientos complejos y razonarlos, a la vez que comienzan a estar familiarizados con los procesos científicos. Alumnos excelentes en Matemáticas y Ciencias que se necesitan cada vez más en un mundo donde dominará la robotización, la automatización y el big data. Uno de los problemas por los que España tiene peores resultados en Matemáticas puede ser por la falta de confianza en sí mismo en Matemáticas. El 27% de los alumnos españoles tiene mucha confianza en sí mismo para abordar las Matemáticas, el 43% tiene algo de confianza y el 30% no tiene confianza en sí mismo en Matemáticas. Entre los países participantes en TIMSS, el 32% tiene mucha confianza en sí mismo en Matemáticas, el 44% tiene algo de confianza y el 23% no tiene ninguna confianza. Los estudiantes españoles también tienen algo menos de confianza en sí mismos en Ciencias, aunque las diferencias son algo menores. El 35% de los alumnos españoles tiene mucha confianza en sí mismo para abordar las Ciencias, el 41% tiene algo de confianza y el 23% no tiene confianza en sí mismo en Matemáticas. Entre los países participantes en TIMSS, el 38% tiene mucha confianza en sí mismo en Matemáticas, el 43% tiene algo de confianza y el 19% no tiene ninguna confianza.

### 3.5. Resultados de los estudiantes españoles en Secundaria: PISA

La OCDE publica cada tres años los resultados de PISA, la prueba que evalúa las competencias de los alumnos de 15 años en Lectura, Matemáticas y Ciencias. Los últimos datos se corresponden con PISA 2018, que se realizó en

España en los meses de abril y mayo de ese año y cuyos principales resultados se publicaron en diciembre de 2019 y julio de 2020. PISA retrasó la siguiente evaluación de 2021 a 2022 por la pandemia y los resultados se conocerán en diciembre de 2023. El promedio de la puntuación de la OCDE en esta evaluación se fijó en 500 puntos en edición inicial de cada competencia (2000 Lectura, 2003 Matemáticas y 2006 Ciencias). Cada edición se centra en una de las tres competencias, lo que significa que cerca de dos tercios de las pruebas se centran en esa área y la estimación de sus resultados es más precisa que las de las otras dos materias secundarias. En 2018, la competencia principal fue Lectura, en 2022 será Matemáticas y en 2025 Ciencias. PISA 2022 incluirá también además pensamiento creativo y PISA 2025 las competencias en un idioma extranjero como señala el capítulo 3e dedicado a bilingüismo.

Los resultados de Lectura, la competencia principal en PISA 2018, de nuestro país se tienen que tomar con cautela porque la OCDE concluyó que "no se puede garantizar completamente la comparabilidad de los datos de PISA 2018 para España con los de evaluaciones anteriores de PISA" (OCDE, 2020b, https://www.oecd.org/pisa/PISA2018-AnnexA9-Spain.pdf). En concreto la OCDE concluye que se produjeron respuestas muy rápidas a algunas preguntas en la prueba de PISA en una parte de la muestra española, un patrón que se observó predominantemente en un pequeño número de centros educativos en algunas áreas de España. Estos centros, se encuentran en Comunidades Autónomas donde el período de evaluación PISA coincidió con los exámenes de los propios Institutos con repercusiones académicas. Entre PISA 2009 y 2015, nuestro país había mostrado una tendencia alza que le había llevado desde los 481 puntos hasta los 496 puntos. En relación con los resultados de Matemáticas, España se mantiene en un resultado similar a los que ha alcanzado desde los orígenes de PISA. En concreto, en 2018 España tiene 481 puntos, significativamente por debajo de la media de la OCDE de 489 puntos. Nuestro país tiene 24 países de la OCDE significativamente por encima, 6 países que tienen una puntuación estadísticamente similar (Italia, República de Eslovaquia, Luxemburgo, Lituania, Hungría y Estados Unidos), y otros 6 significativamente por debajo. En Ciencias, España ha registrado 483 puntos en PISA 2018, una disminución significativa de diez puntos con respecto a los 493 que tenía nuestro país en PISA 2015 y también inferior al promedio de la OCDE en 2018 que se sitúa en 489. Es una evolución preocupante por cuanto España se situaba en la media de la OCDE en 2015, mientras que en 2018 se encuentra por debajo. Hay 23 países significativamente por encima de

la media de España en Ciencias, 3 con puntuaciones similares a nuestro país (Letonia, Lituania y Hungría) y 10 significativamente por debajo. Aunque los problemas de la evaluación de PISA en 2018 en España afectaron sobre todo a Lectura, parece que también hubo un impacto algo menor en los resultados en Matemáticas y Ciencias.

Estonia con 530 puntos en Ciencias, el primero de toda la OCDE, y 523 puntos en Matemáticas (tercero, sólo superado por los asiáticos Japón y Corea) se convirtió en PISA 2018 en la "nueva Finlandia", el país europeo con mejor rendimiento académico y referencia en el Viejo Continente. Es decir, los alumnos de Estonia de 14 años tienen las mismas competencias de los alumnos españoles de 15. También Polonia continúa con su ascenso imparable. Comenzó PISA en el año 2000 con registros muy similares a España. En 2018, y tras un crecimiento sostenido desde entonces, se sitúa en las primeras posiciones con 511 puntos en Matemáticas y 516 en Ciencias. Por el contrario, dos países que habían recibido mucha atención entre la comunidad educativa española, Finlandia y Portugal, bajaron su rendimiento académico en PISA 2018 con respecto a ediciones anteriores.

**Tabla 22. Resultados de España y OCDE en PISA 2009, 2012, 2015 y 2018.**

| | | 2009 | 2012 | 2015 | 2018 |
|---|---|---|---|---|---|
| Matemáticas | España | 483 | 484 | 486 | 481 |
| | OCDE | 495 | 494 | 490 | 489 |
| Ciencias | España | 488 | 496 | 493 | 483 |
| | OCDE | 501 | 501 | 493 | 489 |
| Lectura | España | 481 | 488 | 496 | 477* |
| | OCDE | 493 | 496 | 493 | 487 |

PISA proporciona información sobre el porcentaje de estudiantes que se encuentra en los niveles bajos y muy bajos y entre los excelentes. Con la excepción de PISA 2018, la tabla 23 muestra una característica muy consistente del sistema educativo español y que en alguna medida ya se infería de PIRLS: tenemos menos alumnos rezagados que en la OCDE y también una proporción inferior de estudiantes excelentes. Es decir, que nuestro país tiene más porcentaje de los alumnos en la parte media de la distribución. La falta de alumnos excelentes en España en comparación a la OCDE (y la UE) podría suponer un déficit a medio y largo plazo que podría lastrar las posibilidades de desarrollo de I+D, y la necesaria apuesta por la especialización en sectores de alto valor añadido, intensivos en capital humano y tecnológico. Alemania y Reino Unido han logrado tener también una proporción de estudiantes rezagados más baja que en el promedio de

la OCDE, pero disponen de más excelentes que España y que el conjunto de países desarrollados. Italia, por su parte, tiene una proporción similar de estudiantes rezagados a la OCDE, incluso algo más elevada, y menos alumnos excelentes, con la excepción de Matemáticas. Estados Unidos tiene una distribución de sus alumnos buena en Lectura y Ciencias, con más estudiantes excelentes y menos rezagados que el promedio de la OCDE, y una preocupante situación en Matemáticas. En definitiva, el problema más acuciante para España en relación con los países más grandes y al promedio de la OCDE, es la escasa proporción de estudiantes de 15 años que tiene resultados excelentes (nivel 5 o nivel 6 de los seis existentes).

**Tabla 23. Distribución de los estudiantes españoles en PISA por niveles de rendimiento.**

| **España** | | **2009** | **2012** | **2015** | **2018** |
|---|---|---|---|---|---|
| Lectura | Nivel bajo | 20 | 18 | 16 | 23 |
| | Nivel alto | 3 | 6 | 6 | 5 |
| Matemáticas | Nivel bajo | 24 | 24 | 22 | 25 |
| | Nivel alto | 8 | 8 | 7 | 7 |
| Ciencias | Nivel bajo | 18 | 16 | 18 | 21 |
| | Nivel alto | 4 | 5 | 5 | 4 |
| **OCDE** | | **2009** | **2012** | **2015** | **2018** |
| Lectura | Nivel bajo | 20 | 19 | 20 | 22 |
| | Nivel alto | 7 | 8 | 8 | 8 |
| Matemáticas | Nivel bajo | 24 | 24 | 25 | 24 |
| | Nivel alto | 12 | 12 | 10 | 11 |
| Ciencias | Nivel bajo | 19 | 19 | 22 | 22 |
| | Nivel alto | 8 | 8 | 7 | 7 |
| **Francia** | | **2009** | **2012** | **2015** | **2018** |
| Lectura | Nivel bajo | 20 | 19 | 21 | 21 |
| | Nivel alto | 10 | 13 | 13 | 9 |
| Matemáticas | Nivel bajo | 22 | 22 | 23 | 21 |
| | Nivel alto | 14 | 13 | 11 | 11 |
| Ciencias | Nivel bajo | 19 | 19 | 22 | 21 |
| | Nivel alto | 8 | 8 | 8 | 7 |
| Alemania | | 2009 | 2012 | 2015 | 2018 |
| Lectura | Nivel bajo | 19 | 15 | 16 | 21 |
| | Nivel alto | 8 | 9 | 12 | 11 |
| Matemáticas | Nivel bajo | 19 | 18 | 17 | 21 |
| | Nivel alto | 18 | 18 | 13 | 13 |

| | | | | | |
|---|---|---|---|---|---|
| Ciencias | Nivel bajo | 15 | 12 | 17 | 20 |
| | Nivel alto | 13 | 12 | 11 | 10 |
| **Italia** | | **2009** | **2012** | **2015** | **2018** |
| Lectura | Nivel bajo | 21 | 20 | 21 | 23 |
| | Nivel alto | 6 | 7 | 6 | 5 |
| Matemáticas | Nivel bajo | 25 | 25 | 23 | 24 |
| | Nivel alto | 9 | 10 | 11 | 10 |
| Ciencias | Nivel bajo | 21 | 19 | 23 | 26 |
| | Nivel alto | 6 | 6 | 4 | 3 |
| **Reino Unido** | | **2009** | **2012** | **2015** | **2018** |
| Lectura | Nivel bajo | 18 | 17 | 18 | 17 |
| | Nivel alto | 8 | 9 | 9 | 12 |
| Matemáticas | Nivel bajo | 20 | 22 | 22 | 19 |
| | Nivel alto | 10 | 12 | 11 | 13 |
| Ciencias | Nivel bajo | 15 | 15 | 17 | 17 |
| | Nivel alto | 11 | 11 | 11 | 10 |
| **Estados Unidos** | | **2009** | **2012** | **2015** | **2018** |
| Lectura | Nivel bajo | 18 | 17 | 19 | 19 |
| | Nivel alto | 10 | 8 | 10 | 13 |
| Matemáticas | Nivel bajo | 23 | 26 | 29 | 27 |
| | Nivel alto | 10 | 9 | 6 | 8 |
| Ciencias | Nivel bajo | 18 | 18 | 20 | 19 |
| | Nivel alto | 9 | 8 | 9 | 9 |

Fuente: Ministerio de Educación y FP (Instituto Nacional de Evaluación Educativa, INEE). PISA 2018. Informe español (Volumen I y Lectura en España.) https://www.educacionyfp.gob.es/inee/evaluaciones-internacionales/pisa/pisa-2018/pisa-2018-informes-es.html

La tabla 24 muestra la puntuación medida en Lectura en PISA 2018. Como comentamos, el resultado de España presenta problemas de comparabilidad temporal y con otros países, por lo que la OCDE opta por no incluirlos en algunos de sus análisis. Estados Unidos y Reino Unido son los que tienen los mejores resultados de entre los países más grandes, coincidente con el buen resultado de Inglaterra en PIRLS 2021, aunque se traten de niveles educativos diferentes. El principal indicador que utiliza la OCDE para medir equidad es la fortaleza de la relación entre el índice de estatus económico, social y cultural y el rendimiento educativo (en este caso, en Lectura). Cuanto mayor sea el porcentaje de la varianza en el rendimiento en Lectura que es explicado por ese nivel socio-

económico, menor será la equidad. Lo ideal sería que no existiera ninguna relación entre el índice de estatus económico, social y cultural y el rendimiento educativo, y que fueran dos aspectos independientes. A este respecto, Francia y Alemania serían los países con menor equidad, mientras que Italia y Reino Unido tienen mayor igualdad de oportunidades. La siguiente columna muestra la diferencia en los resultados en Lectura entre el cuartil de mayor nivel socioeconómico y el cuartil inferior. Francia y Alemania, son de nuevo los países con más diferencias y, por tanto, con perores resultados en relación a la igualdad de oportunidades. El capítulo 5b incide en las diferencias señaladas en la columna 4 y muestra resultados para Matemáticas y Ciencias, así como con información por Comunidades Autónomas españolas. Italia y Reino Unido de nuevo destacan por disponer de una mayor proporción de estudiantes académicamente resilientes: alumnos desfavorecidos que obtienen puntuaciones en el cuartil superior de rendimiento en Lectura entre los estudiantes de su propio país.

**Tabla 24. Indicadores de disparidades socio-económicas y rendimiento académico.**

| | Puntuación media en Lectura en PISA 2018 | Índice de Cobertura: cobertura de la población de 15 años | Fortaleza: porcentaje de varianza en el rendimiento en Lectura explicado por ESCS ($R^2$) | Diferencia entre estudiantes favorecidos y desfavorecidos en Lectura | Porcentaje de estudiantes desfavorecidos que son académicamente resilientes |
|---|---|---|---|---|---|
| | (1) Media | (2) | (3) % | (4) Diferencias resultado | (5) % |
| Francia | 493 | 0,91 | 17,5 | **107** | 10 |
| Alemania | 498 | 0,99 | 17,2 | **113** | 10 |
| Italia | 476 | 0,85 | 8,9 | **75** | 12 |
| España | m | 0,92 | m | **m** | m |
| Reino Unido | 504 | 0,85 | 9,3 | **80** | 14 |
| Estados Unidos | 505 | 0,86 | 12,0 | **99** | 10 |

Notas: ESCS se refiere al índice PISA de estatus económico, social y cultural. Un estudiante socioeconómicamente favorecido (desfavorecido) es un estudiante en el cuarto superior (inferior) de ESCS en su propio país/economía. Los estudiantes académicamente resilientes son estudiantes desfavorecidos que obtuvieron puntuaciones en el cuartil superior de rendimiento en Lectura entre los estudiantes de su propio país.
Fuente: OECD, PISA 2018 Database, Tables I.B1.10, II.B1.2.1, II.B1.2.3 y Table II.B1.3.1. https://www.oecd-ilibrary.org/docserver/b5fd1b8f-en.pdf?expires=1692800634&id=id&accname=guest&checksum=DCC4C358CD83D9590D73F96FC063642B

PISA 2018 si proporciona para España el índice de inclusión social (PISA 2018 Results. Volume II, OCDE 2019, p. 23). Es un índice que mide la

variación en el estatus socioeconómico de los alumnos dentro de los centros educativos sobre el total de la variación en este status (incluyendo dentro y entre centros educativos). Por lo tanto, un valor cercano a 100 indicaría que hay mayor inclusión social en el sentido de que en todos los centros educativos se pueden encontrar estudiantes de nivel socioeconómico alto, medio y bajo, mientras que un indicador cercano a cero estaría señalando que los de estatus elevado están en unos institutos o colegios distintos a los de clase media, que a su vez sería diferentes de los de entornos desfavorecidos. España mejoró su inclusión social entre PISA 2015 y PISA 2018 pasando del 69,0 en PISA 2015 al 75,8, situándose ya cerca del promedio de la OCDE del 76,1. En cuanto a la inclusión académica, la variación en el rendimiento académico dentro de los centros dividido por la variación total en el rendimiento en Ciencias, nuestro país ha destacado siempre como de los más inclusivos con un indicador en PISA 2015 del 86,6, muy por encima del 69,9 de la OCDE. Es decir, en España se pueden encontrar estudiantes de rendimiento elevado, medio y rezagado en todos los centros educativos, al contrario que en Países Bajos, Alemania o Austria que al incorporar itinerarios desde edades tempranas tienen una elevada concentración (lo contrario de inclusión) de alumnado avanzado en los itinerarios más académicos y de alumnos medio en los más profesionales.

En resumen, las evaluaciones de dos organismos internacionales diferentes (IEA y OCDE), que analizan niveles educativos distintos (Primaria y Secundaria), con metodologías heterogéneas (PISA evalúa un concepto más amplio como el de las competencias mientras que TIMSS y PIRLS se centran más en el currículo y los conocimientos) llegan a la misma conclusión de que Matemáticas es la asignatura pendiente en España. La mejora de las competencias Matemáticas de los alumnos españoles es un objetivo alcanzable. La diferencia en los resultados en esta asignatura entre España y la OCDE, aunque permanece estable desde hace mucho tiempo, no es insalvable.

La competencia en Matemáticas es un determinante importante de los salarios tanto en la OCDE como en España. Hanushek et al. (2015) muestran, que las personas con más competencias Matemáticas tienen mayores salarios. Como se puede ver en PIAAC (una evaluación como PISA también de la OCDE, pero a la población entre 15 y 65 años), las personas con una desviación estándar más de dominio de las Matemáticas tienen un 17,8% más de salario (22,8% más en el caso de España). En esta misma línea, Hernández y Serrano (2013) encuentran,

utilizando los datos del mismo PIAAC para España, que los conocimientos del individuo tienen un efecto positivo sobre el salario. Las estimaciones que obtienen estos autores indican que cada punto adicional en PIAAC en comprensión lectora aumenta el salario un 0,37%, un efecto muy similar al de un punto adicional en competencia matemática (0,38%). Al considerar simultáneamente ambos tipos de conocimiento, los dos siguen teniendo efectos positivos y significativos sobre el salario, aunque los de las competencias Matemáticas son de mayor magnitud. Por cada punto PIAAC en Matemáticas habría un 0,27% más de salario, mientras que cada punto de comprensión lectora supondría un 0,13% adicional (columna 4). En la columna 7 se puede observar que, al incluir los años de estudio del individuo, el coeficiente de la competencia lectora deja de ser significativo (y de hecho, es cero), mientras que el de Matemáticas continúa siendo significativo. Es decir, que el dominio de las Matemáticas tiene un impacto distintivo en los salarios con respecto a la formación general.

Entre las razones que pueden explicar que las competencias de los alumnos españoles sean más bajas en Matemáticas se puede encontrar que es una asignatura que se les presenta a los niños como si fuera muy difícil, solo apta para los más inteligentes. De hecho, en PISA 2012 los alumnos españoles señalaron que Matemáticas genera más nerviosos e incluso ansiedad a los alumnos españoles que a los de la OCDE, y muy en particular a las alumnas (un aspecto que se trata el capítulo 5c). Puede haber, por tanto, una especie de profecía que se cumple a sí misma. Dado que muchos los alumnos españoles parten de la base de que Matemáticas es difícil y es sólo apta para los más inteligentes, a algunos de ellos se les termina dando mal. Hay mucho margen de mejora presentando las Matemáticas como una competencia que no es inaccesible, que es muy importante en la vida diaria y que cada vez lo va a ser más, en un mundo dominado por el big data y la robotización. De hecho, TIMSS muestra que, tanto en Matemáticas como en Ciencias, la variable que mayor impacto tiene sobre la resiliencia es la confianza que el niño o la niña tenga en el aprendizaje del área. En el caso de España el impacto es superior al del total UE por lo que cuanta mayor confianza se genere en los niños y niñas mejores resultados se pueden llegar a conseguir. La mentalidad del crecimiento que se analiza en el capítulo 5d y que hace hincapié en que la inteligencia no es fija, sino que puede crecer a base de esfuerzo, de aprender de los errores y de perseverancia puede ayudar a que una parte del alumnado no vea las Matemáticas como algo inaccesible.

En 2008, la IEA también realizó una evaluación de conocimientos y pedagogía en Matemáticas a los futuros maestros de Primaria (estudiantes del último curso de la carrera de Educación Primaria). Entre las conclusiones para España se encontraba que el 26,8% de los entonces estudiantes de Magisterio había evitado la asignatura de Matemáticas en Bachillerato. Y un 40,3% adicional no había impartido esta materia en 2º de Bachillerato. Este organismo internacional también encontró que sólo el 6,6% de los créditos de la carrera se dedicaba a Matemáticas. Es decir, que otro de los factores que puede explicar el déficit de las competencias en Matemáticas de nuestros alumnos podría encontrarse en la escasa relevancia que tiene esta asignatura en la formación inicial de los maestros en Primaria. Tanto en el estudio que realizó la IEA en 2008 (TEDS-M, INEE 2012) como ahora en TIMSS 2019 se muestra que los docentes españoles tienen más formación en el área de los procedimientos pedagógicos para las Matemáticas que en los contenidos matemáticos en sí mismos. Equilibrar ambos aspectos, muy importantes y que se refuerzan entre sí, puede ser también otro de los aspectos que podría mejorar la preparación de los docentes en Matemáticas en Primaria.

## 4. EDUCACIÓN Y EMPLEO

La tabla 25 muestra que, para todos los países, también en España y en el promedio de la OCDE y la UE, la tasa de desempleo es mayor cuanto más bajo es el nivel educativo. En España, la tasa de paro de los jóvenes entre 25-34 años que tenía sólo hasta la ESO era del 28% en 2021, del 18% para los que tienen terminado Bachillerato, FP Media o FP Básica y del 13% para los titulados en FP Superior o Universidad. Parece que en la UE-22 la rentabilidad de la educación superior es mayor en términos de tener menor probabilidad de estar en desempleo, un aspecto en el que se indaga también en el capítulo 4c. El paro del 6% es tres veces inferior entre los jóvenes con FP Superior o Universidad en la UE-22 que entre los que tienen sólo hasta estudios obligatorios (18%) mientras que en es en España es sólo la mitad (13% frente a 28%). Otro aspecto interesante también de la tabla 25 es que se muestra que la educación protege en alguna medida en el mercado de trabajo en años de crisis económica. La tasa de desempleo en España aumentó en cinco puntos porcentuales entre los jóvenes con estudios de sólo hasta la ESO, mientras que se incrementa en un punto porcentual tanto para los de estudios medios y superiores.

**Tabla 25. Tendencias en las tasas de desempleo de personas de 25 a 34 años, según el nivel de educación (de 2019 a 2021).**

| | Tasa de paro | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Hasta la ESO | | | Bachillerato, FP Media o FP Básica | | | FP Superior o Universidad | | |
| | 2019 | 2020 | 2021 | 2019 | 2020 | 2021 | 2019 | 2020 | 2021 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Francia | 24 | 20 | 20 | 11 | 11 | 10 | 6 | 7 | 6 |
| Alemania | 12 | 12 | 11 | 3 | 4 | 4 | 3 | 4 | 3 |
| Italia | 21 | 21 | 21 | 14 | 13 | 14 | 12 | 11 | 10 |
| España | 23 | 28 | 28 | 17 | 20 | 18 | 12 | 15 | 13 |
| Reino Unido | 7 | 7 | 9 | 3 | 4 | 4 | 2 | 3 | 3 |
| Estados Unidos | 10 | 9 | 11 | 6 | 8 | 10 | 2 | 3 | 4 |
| **OCDE** | **13** | **15** | **15** | **7** | **9** | **8** | **5** | **7** | **6** |
| **UE-22** | **16** | **18** | **18** | **7** | **9** | **8** | **5** | **6** | **6** |

Fuente: OCDE. Education at a Glance 2022 Página 78 https://www.oecd-ilibrary.org/docserver/3197152b-en.pdf?expires=1692871828&id=id&accname=guest&checksum=6BC4DDB093888CCD1A28C745E480F768

La rentabilidad de la educación superior también es más reducida en España en términos de ingresos de la población de 25-64 años (véase capítulo 4c). La tabla 26 revela que las personas con educación superior tienen un 41% más de ingresos laborales que los que tienen, como mucho, Bachillerato, FP Media o FP Básica, una prima salarial que en el promedio de la OCDE es del 55% y en la UE-22 del 50%. Las diferencias en rentabilidad entre España y la OCDE provienen tanto de la FP Superior, como del Grado Universitario como del Máster. En comparación con la UE-22, los ingresos laborales de los dos primeros niveles educativos también son inferiores, mientras que en los Máster se iguala. La diferencia de ingresos de los trabajadores entre las personas con estudios de sólo hasta la ESO con respecto a los estudios medios (Bachillerato, FP Media o FP Básica) en España es, por el contrario, más elevada que en la UE y, sobre todo, que en la OCDE. El capítulo 4c de este libro explora la evolución, causas y consecuencias de la prima salarial en España por niveles de

estudio. Una posible hipótesis que podría explicar la escasa prima salarial de la realización de Grados universitarios en España en relación con otros países sería que la formación que aporten estos estudios a nuestros jóvenes no sea de la misma calidad que la de los países con universidades más punteras. Es decir, que la cantidad de capital humano "real" contenida en una titulación superior (Felgueroso et al., 2011), sea menor en España que en el promedio de la OCDE.

**Tabla 26. Ingresos relativos de los trabajadores, según el nivel de educación (2020). Personas de 25-64 años con ingresos provenientes del empleo (trabajadores a tiempo completo durante todo el año); nivel de Educación Secundaria superior = 100.**

| | Hasta la ESO | Educación Superior | | | |
|---|---|---|---|---|---|
| | | FP Superior | Grado Universitario | Máster Universitario | Total Educación Superior |
| | (1) | (3) | (4) | (5) | (6) |
| Francia (1) | 90 | 122 | 142 | 181 | 149 |
| Alemania | 78 | 133 | 167 | 171 | 162 |
| Italia (1) | 83 | – | – | – | 137 |
| España (1) | 81 | 105 | 130 | 168 | 141 |
| Reino Unido | 73 | 121 | 140 | 168 | 145 |
| Estados Unidos | 74 | 112 | 163 | 219 | 171 |
| **OCDE** | **83** | **120** | **144** | **188** | **155** |
| **UE-22** | **86** | **122** | **136** | **168** | **150** |

El año de referencia para Francia e Italia es 2018 y para España es 2019.
Fuente: OCDE. Education at a Glance 2022 Página 92 https://www.oecd-ilibrary.org/docserver/3197152b-en.pdf?expires=1692871828&id=id&accname=guest&checksum=6BC4DDB093888CCD1A28C745E480F768

PISA 2018 incluyó en su evaluación, junto con Lectura, Matemáticas y Ciencias, la "Competencia Global" (OECD Education at a Glance, 2022). Esta competencia incluye la actitud de los estudiantes hacia los inmigrantes a través de sus respuestas a las siguientes afirmaciones: "Los niños inmigrantes deberían tener las mismas oportunidades educativas que otros niños en el país", "Los inmigrantes que viven en un país durante varios años deberían tener la oportunidad de votar en las elecciones", "Los inmigrantes deberían tener la oportunidad de continuar con sus propias costumbres y estilo de vida", y "Los inmigrantes deberían tener todos los mismos derechos que

todos los demás en el país". Las respuestas se proporcionaron en una escala de cuatro puntos: "totalmente en desacuerdo", "en desacuerdo", "de acuerdo" y "totalmente de acuerdo". Las respuestas de los alumnos se reflejan en un índice que mide la actitud general hacia los inmigrantes se, donde un valor positivo indica que los estudiantes tienen actitudes más positivas hacia los inmigrantes que el estudiante promedio en países de la OCDE. La tabla 27 muestra la información del Anuario estadístico de la OCDE, Education at a Glance 2022, revela que en la OCDE y la UE-22 el índice de la actitud hacia los inmigrantes es más positivo cuanto mayor es el nivel educativo de la madre del estudiante encuestado, aunque la asociación es menos concluyente al examinar algunos de los componentes como los relativos a la identidad o derechos políticos. En el caso de España, el índice de la actitud hacia los inmigrantes es más positivo que en el conjunto de la OCDE y UE-22.

La Competencia Global de PISA 2018 preguntó también a los estudiantes hasta qué punto están de acuerdo ("totalmente en desacuerdo", "en desacuerdo", "de acuerdo" o "totalmente de acuerdo") con las siguientes afirmaciones: "Me considero un ciudadano del mundo", "Cuando veo las condiciones precarias en las que viven algunas personas, siento la responsabilidad de hacer algo al respecto", "Creo que mi comportamiento puede afectar a personas de otros países", "Es correcto boicotear a las empresas que se sabe que ofrecen malas condiciones de trabajo a sus empleados", "Puedo hacer algo acerca de los problemas del mundo" y "Cuidar del medio ambiente global es importante para mí" (OCDE, Education at a Glance 2022). Los valores positivos en este índice indican que los estudiantes tienen un mayor sentido de mentalidad global que los estudiantes promedio en los países de la OCDE. En todos los países, excepto en Italia, los estudiantes con madres con educación superior tienen un mayor sentido de mentalidad global. El índice de la mentalidad global de los estudiantes españoles de 15 años (+0,24) es superior al primerio de la OCDE y la UE-22, y es también más positivo cuánto más elevado es el nivel educativo de la madre.

La tabla 28 termina con la transición entre la educación y el mercado de trabajo. Un 48% de las personas de las personas con ESO como mucho están empleados en la franja de edad de 25-29 años y otro 48% no está en el mercado de trabajo, bien porque está desempleado (27,9%) o porque está inactivo (20,2%). Entre los jóvenes de 25 a 29 años con Bachillerato, FP Media o FP Básica finalizado, hay un 22% en el sistema educativo (por el 4%

**Tabla 27. Tolerancia social de los estudiantes, según el nivel educativo de la madre (2018) PISA; basado en respuestas de estudiantes de 15 años.**

| | Actitud en relación con la inmigración | | | | Mentalidad global | | | |
|---|---|---|---|---|---|---|---|---|
| | Nivel educativo de la madre | | | | Nivel educativo de la madre | | | |
| | Hasta ESO | Bachillerato, FP Media, FP Básica | Educación Superior | Total | Hasta ESO | Bachille-rato, FP Media, FP Básica | Educación Superior | Total |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Francia | – | – | – | – | – | – | – | – |
| Alemania | 0,09 | 0,05 | 0,21 | 0,12 | –0,35 | –0,32 | –0,10 | –0,27 |
| Italia | –0,20 | –0,22 | –0,23 | –0,22 | –0,13 | –0,09 | –0,11 | –0,10 |
| España | 0,37 | 0,44 | 0,40 | 0,39 | 0,15 | 0,20 | 0,29 | 0,24 |
| Escocia (Reino Unido) | 0,18 | 0,27 | 0,42 | 0,34 | –0,03 | –0,08 | 0,01 | –0,05 |
| Estados Unidos | – | – | – | – | – | – | – | – |
| **OCDE** | **–0,02** | **0,00** | **0,07** | **0,03** | **–0,10** | **–0,03** | **0,08** | **0,01** |
| **UE-22** | **–0,15** | **–0,15** | **–0,08** | **–0,12** | **–0,20** | **–0,12** | **–0,01** | **–0,08** |

Fuente: PISA 2018 Results (Volume VI) https://www.oecd.org/pisa/publications/pisa-2018-results-volume-vi-d5f68679-en.htm

de los de ESO), 51% de empleados (similar al 48% de entre los de ESO) y un 27,3% que ni estudia ni trabaja (mucho menos que el 48,1% de entre los de la ESO). También hay un 22% de entre los jóvenes españoles de 25-29 años con estudios superiores que continúan en el sistema educativo, pero hay más empleados (59% frente al 51% de los que tienen estudios medios) y menos que ni estudian ni trabajan (18,4% versus 27,3%).

**Tabla 28. Porcentaje de jóvenes de 25 a 29 años en educación/no en educación, según nivel educativo y situación laboral (2021, datos trimestrales).**

| | ESO | | | | | | Bachillerato, FP Media o FP Básica | | | | | | FP Superior, Universidad | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | En educación | No en educación | | | | | En educación | Empleados | No en educación | | | | En educación | No en educación | | | | |
| | | Empleados | Ni en educación ni en el mercado de trabajo | | | Total | | | Ni en educación ni en el mercado de trabajo | | | Total | | Empleados | Ni en educación ni en el mercado de trabajo | | | Total |
| | | | Desempleo | Inactivo | Total | | | | Desempleo | Inactivo | Total | | | | Desempleo | Inactivo | Total | |
| | (1) | (2) | (3) | (4) | (5) = (3) + (4) | (6) = (2) + (5) | (7) | (8) | (9) | (10) | (11) = (9) + (10) | (12) = (8) + (11) | (13) | (14) | (15) | (16) | (17) = (15) + (16) | (18) = (14) + (17) |
| Francia | 3 | 46 | 13,0 | 37,5 | 50,5 | 97 | 9 | 68 | 10,7 | 12,1 | 22,8 | 91 | 12 | 76 | 6,9 | 4,7 | 11,6 | 88 |
| Alemania | 10 | 51 | 8,4 | 30,8 | 39,2 | 90 | 21 | 68 | 3,7 | 7,6 | 11,3 | 79 | 26 | 68 | 2,2 | 4,5 | 6,7 | 74 |
| Italia | 3 | 41 | 18,5 | 37,5 | 55,9 | 97 | 16 | 50 | 11,0 | 22,9 | 33,9 | 84 | 27 | 51 | 7,7 | 14,1 | 21,9 | 73 |
| España | 4 | 48 | 27,9 | 20,2 | 48,1 | 96 | 22 | 51 | 18,9 | 8,4 | 27,3 | 78 | 22 | 59 | 11,8 | 6,6 | 18,4 | 78 |
| Reino Unido | 5 | 57 | 6,1 | 32,1 | 38,3 | 95 | 11 | 76 | 4,2 | 9,4 | 13,6 | 89 | 15 | 79 | 2,7 | 3,7 | 6,4 | 85 |
| Estados Unidos | 7 | 50 | 7,8 | 35,1 | 43,0 | 93 | 10 | 63 | 8,2 | 19,2 | 27,4 | 90 | 14 | 75 | 3,8 | 7,2 | 11,0 | 86 |
| **OCDE** | **10** | **49** | **11,0** | **31,8** | **42,2** | **92** | **17** | **63** | **7,7** | **12,5** | **20,0** | **83** | **18** | **70** | **6,0** | **6,1** | **12,1** | **82** |
| **UE-22** | **10** | **48** | **12,4** | **32,7** | **44,4** | **92** | **16** | **66** | **7,7** | **11,2** | **18,9** | **84** | **18** | **70** | **5,9** | **5,5** | **11,4** | **82** |

Fuente: OCDE. Education at a Glance 2022 Página 59 https://www.oecd-ilibrary.org/docserver/3197152b-en.pdf?expires=1692871828&id=id&accname=guest&checksum=6BC4DDB093888CCD1A28C745E480F768

## BIBLIOGRAFÍA


Araujo, M. C., Carneiro, P. Cruz-Aguayo, Y. y Schady, N. (2016). Teacher Quality and Learning Outcomes in Kindergarten. *Quarterly Journal of Economics, 131*(3), 1415–53. DOI:10.1093/qje/qjw016

Battaglia, M. y Lebedinski, L. (2022). With a little help from my friends: Medium-Term effects of a remedial education program targeting Roma minority. *Economics of Education Review 86*, 102196. https://doi.org/10.1016/j.econedurev.2021.102196

Bentolila, S., Cabrales, A. y Jansen, M. (2023). The Impact of Dual Vocational Education on the Labor Market Insertion of Youth: Evidence from Madrid. *FEDEA,* (mimeo). https://www.fedea.net/nsaw/descargas/NSAW05en_dual_vocational_madrid.pdf

Cabrales, A., Güell, M., Madera, R. y Viola, A. (2019). Income contingent university loans: Policy design and an application to Spain, *Economic Policy*. 34, 99, 479-521. https://doi.org/10.1093/epolic/eiz010

Cadima, J., Nata, G., Barros, S., Coelho, V. y Barata, C. (2020). Literature review on early childhood education and care for children under the age of 3. *OECD Education Working Papers, 243*. OECD, Paris. https://doi.org/10.1787/a9cef727-en.

Comisión Europea (2022). *Education and Training Monitor 2022*. Comisión Europea, Luxemburgo. https://op.europa.eu/webpub/eac/education-and-training-monitor-2022/downloads/comparative-report/Education-and-Training-Monitor-Comparative-Report.pdf

Comisión Europea. *Employment and Social Developments in Europe 2022*. Comisión Europea, Luxemburgo. https://ec.europa.eu/social/main.jsp?langId=en&catId=89&newsId=10330

Duncan, G., Kalil, A., Mogstad, M. y Rege, M. (2022). Investing in Early Childhood Development in Preschool and at Home, en Machin, S.J., Woessmann, L. y Hanushek, E.A. (eds.), *Handbook of the Economics of Education,. Volume 7*. North Holland, Amsterdam (en prensa).

Felgueroso, F., Garicano, L. y Jiménez, S. (2011). *¿Vale la pena estudiar? I El subempleo y las ganancias salariales de la educación en España*. Apuntes Fedea

Educación, 01. https://www.fedea.net/apuntes-fedea/apuntes/educacion/apunte_educacion01_valelapenaestudiar1.pdf

Flisi, S. y Blasko, Z. (2019). A note on early childhood education and care participation by socio-economic background. *JRC Research Reports,* JRC117663. Joint Research Centre, Sevilla. https://publications.jrc.ec.europa.eu/repository/handle/JRC117663

Fundación Europea Sociedad y Educación 2023 (2023). *Indicadores comentados sobre el estado del sistema educativo en España*. Fundación Europea Sociedad y Educación, Madrid. https://www.sociedadyeducacion.org/core/wp-content/uploads/INDICADORES_-2023_WEB_09082023.pdf

García-Montalvo, J. (2020). El impacto de la introducción de subvenciones progresivas en la educación superior el caso de Cataluña. *Papeles de Economía Española*. 166, 123-142.

Gershenson, S., Hart, C. M. D., Hyman, J., Lindsay, C.A. y Papageorge, N.W. (2022). The Long-Run Impacts of Same-Race *Teachers. American Economic Journal: Economic Policy*. 14 (4), 300-342. DOI: 10.1257/pol.20190573

Hanushek, E. A., Schwerdt, G., Wiederhold, S. y Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73(C), 103-130. DOI: 10.1016/j.euroecorev.2014.10.006

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*. 312 (5782), 1900-1902. https://doi.org/10.1126/science.1128898

Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A. y Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics, 94*, 1-2, 114-128. https://doi.org/10.1016/j.jpubeco.2009.11.001.

Hospido, L., Crespo, L., Fernández, M. y Montalbán, J. (2019). ¿Qué sabemos sobre el efecto del tipo de jornada escolar sobre el rendimiento educativo?, en *Indicadores comentados sobre el estado del sistema educativo español*. Fundación Europea Sociedad y Educación, Madrid, pp. 86-90.

IEA (2023). *PIRLS 2021*. IEA: Amsterdam. https://www.iea.nl/studies/iea/pirls/2021

Instituto Nacional de Evaluación Educativa (INEE) (2012). *TEDS-M Informe Español. Estudio internacional sobre la formación inicial en matemáticas de los maestros*. Ministerio de Educación y FP, Madrid. https://www.educacionyfp.gob.es/dctm/inee/internacional/teds-mlinea.pdf?documentId=0901e72b8143866e

Instituto Nacional de Evaluación Educativa (2023). *PIRLS 2021*. Ministerio de Educación y FP, Madrid. https://www.educacionyfp.gob.es/inee/evaluaciones-internacionales/pirls/pirls-2021.html

Levin, H., McEwan, P., Belfield, C., Bowden, A. y Shand, R. (2017). *Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis*, SAGE, Thousand Oaks.

Maurin, E. y McNally, S. (2008). Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry Students. *Journal of Labor Economics, 26*(1), 1-33. DOI: 10.1086/522071

Ministerio de Educación y FP (2023a). *Sistema Estatal de Indicadores de la Educación*. Ministerio de Educación y FP, Madrid. https://www.educacionyfp.gob.es/dam/jcr:e63eb14e-5c21-49ca-be81-90cedfc55144/seie-2023.pdf

Ministerio de Educación y FP (2023b). *Datos y cifras. Curso escolar 2023-2024*. Ministerio de Educación y FP, Madrid. https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/indicadores/datos-cifras.html

Ministerio de Educación y FP (2023c). *Recursos económicos. Becas y ayudas al estudio. Curso 2021-2022.* Ministerio de Educación y FP, Madrid. https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/economicas/becas/2021-2022.html

Montalbán, J. (2019). *Becas universitarias y requisitos académicos: ¿un trade-off entre eficiencia y equidad? Nada es gratis.* https://nadaesgratis.es/jose-montalban/becas-universitarias-y-requisitos-academicos-un-trade-off-entre-eficiencia-y-equidad (accedido 16 octubre 2023).

Nickow, A., Oreopoulos, P. y Quan, V. (2020). The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence. *National Bureau of Economic Research Working Paper, 27476*.

OCDE (2019). *PISA 2018 Results (Volume II): Where All Students Can Succeed*. OCDE, Paris. https://www.oecd.org/publications/pisa-2018-results-volume-ii-b5fd1b8f-en.htm

OCDE (2020a). *Education at a Glance 2020: OECD Indicators*. OCDE, Paris. https://doi.org/10.1787/69096873-en

OCDE (2020b). *PISA 2018 Results (Volume I): What Students Know and Can Do, Annex A9. A note about Spain in PISA 2018: Further analysis of Spain's data by testing date* (updated on 23 July 2020). PISA, OCDE, Paris. https://www.oecd.org/pisa/PISA2018-AnnexA9-Spain.pdf

OCDE (2021). *Education at a Glance 2021: OECD Indicators*. OCDE, Paris. https://doi.org/10.1787/b35a14e5-en

OCDE (2022). *Education at a Glance 2022: OECD Indicators*. OCDE, Paris. https://doi.org/10.1787/3197152b-en

OCDE (2023). *Education at a Glance 2023: OECD Indicators*. OCDE, Paris. https://www.oecd.org/education/education-at-a-glance/

Patrinos, H., Vegas, E. y Carter-Rau, R. (2023). An Analysis of Covid-19 Student Learning Loss. *Oxford Encyclopedia Research of Economics and Finance* https://doi.org/10.1093/acrefore/9780190625979.013.893

# Capítulo 2
# MÉTODOS EMPÍRICOS: BREVE INTRODUCCIÓN A LA INFERENCIA CAUSAL

Jesús M. Carro
Departamento de Economía
Universidad Carlos III de Madrid
ORCID: https://orcid.org/0000-0003-2952-7319

## 1. INTRODUCCIÓN

Este capítulo es meramente instrumental dentro de este libro. Pretende ayudar a aquellos no formados en Econometría en general y, en particular, en las técnicas econométricas para la evaluación de impacto y la estimación de efectos causales. Se trata de poder leer los otros capítulos entendiendo cómo se han identificado los efectos de los que se hablan e incluso, discutir sobre el alcance y limitaciones de los mismos. He intentado centrarme en discutir a nivel más bien intuitivo, pero a la vez riguroso, cómo se identifican esos efectos y qué posibles sesgos puede haber al usar los datos, sin entrar en detalles técnicos. Tampoco he intentado explicar las partes más estadísticas, como los errores estándar, que, aun siendo necesarias en el trabajo científico del que salen los resultados, no son la parte central de la discusión respecto a la validez e interpretación de los resultados y potenciales sesgos. Dado esto, este capítulo en absoluto sustituye a un buen curso general de Econometría y a otro curso específico sobre todas estas técnicas econométricas. No obstante, espero que haga más fácil (o si se quiere más profunda) la lectura del resto de los capítulos por parte de aquellos que carecen de esa formación.

Algunas particularidades de este capítulo:

- He intentado reducir la notación matemática y las ecuaciones al mínimo. Por eso casi no se encuentra nada en la mayoría del capítulo. Pero, no lo he evitado del todo, porque un poco de notación a veces ayuda mucho a explicar las cosas. Para compensar esta falta de expresiones matemáticas, hay una penúltima sección que pretende ser un resumen de todo el capítulo, pero escrito sólo con expresiones matemáticas.

- Dado que pretende ser formativo y no una revisión de la literatura, he optado por no incluir referencias bibliográficas a cada paso, ni andar constantemente diciendo dónde se pueden encontrar más detalles de cada cosa, ni de dónde se demostró por primera vez. Para compensar, y también ayudar al lector que quiera saber más, o que quiera consultar alguna cosa específica, la última sección tiene entidad propia y no es una mera lista por orden alfabético con la bibliografía. En esa última sección se ofrece bibliografía agrupada por temática: desde los libros de texto a los que poder acudir, hasta algunos artículos recientes con los últimos avances en alguno de los estimadores.

Agradezco a Antonio Cabrales y a Ismael Sanz por la iniciativa para hacer este libro e invitarme a participar. También tengo que agradecer a José Montalbán haberse leído un primer borrador y haber proporcionado comentarios y materiales que han ayudado a mejorar este capítulo, a Jennifer Ruiz-Valenzuela por su sugerencia para algunas de las ilustraciones que uso, y a los participantes en el *Workshop* en el que presentamos los capítulos de este libro. Por supuesto, cualquier error o limitación que siga existiendo es de mi entera responsabilidad.

## 2. ¿QUÉ EFECTO QUEREMOS CONOCER?

### 2.1. Definiciones previas

#### *2.1.1. Aquello cuyo efecto queremos conocer*

Las primeras dos cosas que hay que definir es de qué queremos conocer el efecto y sobre qué variable. Por ejemplo, queremos conocer el efecto de reducir el tamaño de clase sobre el rendimiento académico de los estudiantes. De qué queremos conocer el efecto suele, a priori, estar relativamente claro (para dicho ejemplo, "reducir el tamaño de clase"). Pero, puede haber dificultades a la hora de medirlo u observarlo. Si lo que queremos es conocer el efecto de una política educativa, tal y como esta ha sido implementada, puede estar más claro lo que tenemos que observar. Tendremos que saber dónde y cuándo se ha implementado y en qué ha consistido la política. Esta política puede consistir en varias intervenciones, o afectar a varios aspectos que determinan un resultado educativo. Un ejemplo es conocer el efecto de una política educativa que implante un sistema de educación bilingüe en colegios. La política puede implicar cambios pedagógicos a la hora de dar las clases, cambios en los maestros que dan esas clases y la formación

de los mismos, cambios en los recursos de los que dispone el colegio, etc. Si el objetivo es conocer el efecto de esta política, podemos hacerlo sin necesidad de tener una medida separada de cada una de esas variables. Algo similar sucede si lo que se quiere es conocer el efecto de estudiar en un centro o tipo de centros educativos concreto frente a hacerlo en otros centros educativos, como por ejemplo en un centro público frente a uno privado.

Si, sin embargo, queremos saber el efecto de cambiar la magnitud de una variable que representa un insumo en la producción de educación –véanse los siguientes capítulos del libro para más detalles de la función de produción de educación–, como el efecto de reducir el tamaño de clase, tenemos que definir la medida que queremos usar; o elegir entre las que tengamos disponibles la que mejor aproxime nuestra variable de interés. En el ejemplo del tamaño de clase, se puede usar el número de alumnos que hay en una clase (grupo de alumnos) cuando estos están construidos de forma estable a lo largo del curso y día escolar. Si los grupos de alumnos, en concreto su tamaño, son distintos para distintas asignaturas, puede resultar más conveniente usar la ratio alumnos/profesor. Ambas están relacionadas, pero no son iguales (véanse los capítulos 1 y 3a de Ismael Sanz en este libro para más detalles sobre las diferencias y relación entre ambas medidas). En caso de disponer de ambas medidas habrá que decidir cuál es la que mejor refleja lo que queremos conocer.

Otras variables sobre las que nos gustaría conocer el efecto pueden ser inobservables, como la calidad docente de cada maestro, y tendremos que intentar buscar *proxis*, o, mejor, usar métodos que permitan identificar (de forma indirecta) esto que no observamos/medimos de forma directa.

Llamaremos $D$ a la variable cuyo efecto queremos conocer. Queremos conocer el efecto de cambiar el valor de $D$ de un valor $D_0$ a un valor $D_1$. En el ejemplo del tamaño de clase ($D$), podríamos querer conocer el efecto de reducirlo de 30 (este sería el $D_0$) a 25 (este sería el $D_1$) alumnos por clase.

### *2.1.2. La variable de resultado*

Definir sobre qué variable o medida de resultado queremos conocer el efecto, puede ser todavía más difícil en Educación. ¿Cómo se mide lo que se produce en Educación?

¿Cómo se mide la calidad de la educación recibida? En la literatura suelen usarse medidas de dos tipos:

- Medidas internas al sistema educativo: resultados en exámenes estandarizados que realizan todos los alumnos, progreso académico, etc.
- Valoración que el mercado de trabajo, o la sociedad en general, hace del capital humano adquirido a través de la educación: salarios, probabilidad de desempleo, criminalidad, etc.

Llamaremos $Y$ a la medida de resultado sobre la que queremos conocer el efecto de $D$. Un ejemplo de esta $Y$ es el resultado en un examen que tienen que realizar todos los alumnos de una región al terminar una etapa educativa.

## 2.2. Efecto causal

Una relación o efecto causal nos dice cuál sería el valor de $Y$ en un mundo alternativo, o contrafáctico, en el que se cambia el valor de $D$. Ese mundo contrafáctico, está formado por los mismos individuos que el fáctico, y sucede en el mismo momento, es decir, en las mismas circunstancias, pero el valor de $D$ se ha cambiado de $D_0$ a $D_1$. El efecto causal es la diferencia entre el valor de $Y$ cuando $D$ toma el valor $D_1$ y el valor de $Y$ cuando $D$ toma el valor $D_0$. Esta es la diferencia debida solamente a ese cambio del valor de $D$. Este es el efecto que nos dice qué pasaría si implementamos una intervención que nos permita cambiar el valor de $D$ en ese mundo.

### *2.2.1. Efecto total o de forma reducida, efecto directo o estructural y mecanismos que producen esos efectos*

Debido a que suele haber más variables que afectan a la $Y$, y a que puede haber relaciones de causalidad o efectos cruzados entre ellas y la $D$, entre los efectos causales de interés podemos distinguir dos tipos: a) efecto causal definido en el contexto de un modelo estructural; y b) efecto causal total o de forma reducida. La principal diferencia es si estamos interesados en conocer una relación causal definida a través de un mecanismo concreto, dentro de un modelo económico, o si lo que queremos es conocer el efecto total sobre $Y$ de cambiar $D$, independientemente de los mecanismos a través de los que el cambio en $D$ lleve a ese cambio en $Y$.

El primero suele centrarse en un único mecanismo directo, dejando el resto de las vías por las cuales *D* puede afectar a *Y* invariantes. Esto significa que este efecto nos dice el cambio en *Y* cuando dejamos constante cualquier otra cosa que pueda afectar a *Y*. El segundo efecto, incluirá también cualquier otro cambio en otra variable que afecte a la *Y* que se haya producido como consecuencia (y sólo por esta razón) del cambio en *D*.

Por ejemplo, si queremos conocer el efecto de reducir el número de alumnos en una clase (esta sería la *D*) en los resultados en un examen que hacen los alumnos al final del curso (la *Y*), pudiera ser que el efecto total viniera dado no sólo por el mecanismo por el cual cada alumno puede recibir más atención individualizada del profesor si hay menos alumnos, sino por cómo reaccionan los padres de esos alumnos ante esa nueva situación. El tiempo y recursos que los padres destinan a ayudar a los alumnos con sus estudios (*X*) puede cambiar como consecuencia de que ahora reciben más atención del profesor, de forma que el cambio en *D* tiene un efecto (positivo) directo sobre la *Y*. Pero, a la vez, el cambio en *D* puede causar un cambio en otro de los determinantes de *Y*, el tiempo que los padres dedican. De esta forma el efecto total de cambiar la *D* (el efecto de forma reducida) sería la suma del efecto directo de *D* sobre *Y* más el efecto indirecto a través de la reducción en *X* que el cambio en *D* causa. Este efecto total, y sin separar ambos mecanismos, puede ser el efecto causal de interés si lo que queremos es conocer el impacto que tendría una política educativa de reducción del tamaño de clase. Interesaría conocer el efecto total, sin necesidad de distinguir los distintos mecanismos, para evaluar la efectividad de la política sobre el resultado.

El efecto causal directo de *D*, separado del efecto de *X* (y por lo tanto del efecto de *D* a través de la *X*), es lo que nos interesa conocer cuando queremos caracterizar la contribución de cada uno de los factores que determinan la *Y* y/o los mecanismos por los cuales dicha política de cambiar *D* produce su efecto total. La utilidad científica de conocer y separar los distintos mecanismos que dan lugar al efecto causal total es mayor que la de conocer sólo el efecto de forma reducida. Esto incluye que de los primeros se puede obtener este último, pero no al revés. Incluso aunque conocer el mecanismo por el cual se dan los efectos de distintas políticas no sea el objetivo final, sólo conociéndolos podremos diseñar políticas alternativas que puedan funcionar mejor que la que hemos evaluado.

## 3. PRINCIPALES DIFICULTADES PARA CONOCER EL EFECTO CAUSAL

### 3.1. No se observa la situación contrafáctica

La primera y gran dificultad para conocer el efecto causal es insoslayable y casi obvia: no podemos observar a la vez a un mismo individuo con un valor $D_0$ y con un valor $D_1$. En un mismo momento del tiempo un individuo sólo puede tener uno de esos dos valores de $D$. Solo podemos observar el mundo fáctico; nunca el contrafáctico (o contrafactual), como sus nombres indican. Dada esta imposibilidad física, el reto pasa a ser encontrar un grupo de individuos que observemos teniendo el valor alternativo de $D$, $D_0$, y que sea comparable (igual en todo excepto en el valor de $D$) a los individuos que tienen $D$ igual a $D_1$.

**Resultados potenciales y observados.** Para todo individuo existen a priori dos resultados o valores de la $Y$ potenciales: 1) el que obtendrá si $D = D_0$ que llamamos $Y(D_0)$; y 2) el que obtendrá si $D = D_1$ que llamamos $Y(D_1)$. El efecto causal para este individuo es $Y(D_1) - Y(D_0)$. El problema es que para un mismo individuo sólo se observa uno de esos resultados potenciales. Dependiendo del valor de $D$ para este individuo, la $Y$ que observamos será igual, o bien a $Y(D_1)$, o bien a $Y(D_0)$. Pero no podemos observar los dos a la vez para un mismo individuo.

Observar la $Y$ para este mismo individuo con valores distintos de $D$ en distintos momentos del tiempo, tampoco resuelve esta dificultad, porque puede haber factores que determinan la $Y$, distintos de $D$, que cambien de un periodo a otro. Eso haría que la diferencia entre esos dos resultados observados no fuera el efecto causal de la $D$, sino que también incluyera el efecto de otras cosas que pueden cambiar en el tiempo.

Tomando la terminología de la literatura de evaluación de los tratamientos médicos, al grupo con $D_1$ se le llama grupo de los tratados, y al grupo con $D_0$ grupo de control. Usando esta terminología, dada la imposibilidad física de ser tratado y control a la vez, la dificultad pasa por encontrar un buen grupo de control que nos pueda decir cuál habría sido el valor de $Y$ de no haber sido tratados. O, de forma análoga, que nos permita decir que el valor de $Y$ que habría tenido el grupo de control de haber sido tratado es el valor de $Y$ que observamos para el grupo de los tratados.

La búsqueda de estos dos grupos comparables, es decir, la búsqueda de un buen grupo de control nos lleva a la siguiente dificultad importante, y que será la que los métodos que vamos a revisar tratan de solucionar: el sesgo de selección.

### 3.2. Sesgo de selección

Se produce un sesgo de selección cuando el grupo de individuos con $D_1$ tiene características (observables e inobservables) distintas del grupo de individuos con $D_0$. Importante, estas características tienen que ser relevantes para determinar (afectar a) la $Y$. No hay ningún problema en que ambos grupos sean distintos en aspectos irrelevantes para la $Y$. Por ejemplo, si la $Y$ es una medida de rendimiento académico, como la nota en un examen estandarizado, la edad del alumno es una variable relevante. No queremos comparar tratados y controles que tengan una edad distinta, porque atribuiríamos erróneamente al tratamiento lo que se debe a una diferencia de edad. Pero la intensidad con la que se veían las estrellas de alfa centauri en el cielo el día del nacimiento del alumno no es una variable relevante y no importa que sea distinta entre tratados y controles.

Nótese que, al hablar y comparar grupos, en lugar de un único individuo, estaremos interesados en una medida informativa del efecto causal para todo el grupo. La más frecuente es el efecto causal medio. Aunque existen otras posibles medidas que describan el efecto causal en la población, por simplicidad voy a referirme la mayor parte del tiempo al efecto medio. Dado esto, otra matización importante es que no necesitamos que cada individuo tratado tenga a alguien exactamente igual (en características relevantes) en el grupo de control. El sesgo de selección surge cuando ambos grupos son distintos en media, y seguirían siendo distintos, aunque hubieran tenido el mismo valor de $D$. En la sección 11.1 se definen matemáticamente los efectos causales medios y el problema del sesgo de selección.

Ejemplo 1. *Supongamos que queremos conocer la calidad de un centro educativo frente a otro, entendido como saber cómo cambia el resultado de los alumnos en el examen de acceso a la Universidad (Y) si estudian en un centro educativo o estudian en otro. Es muy frecuente observar que para decidir la calidad de un centro educativo frente a otro se compare la nota media de los alumnos de esos dos centros. Es más, suelen elaborarse rankings de centros ordenados por esa nota media de los alumnos de cada centro, interpretando las diferencias en notas medias como debidas (causadas) por la distinta calidad educativa aportada por cada centro a sus alumnos. Bajo esa interpretación, esa diferencia sería lo que mejo-*

*raría un mismo alumno en media si pasara de estudiar en un centro a estudiar en otro con mejor nota media de sus alumnos en el examen de acceso a la Universidad. Sin embargo, claramente no hay ninguna garantía de que eso nos dé el efecto causal y sufre un potencial problema de sesgo de selección. Los alumnos de un centro y de otro son distintos, con caracterísitcas distintas, y contextos socioeconómicos distintos. Las diferencias en Y serán en su mayor parte debidas a características y condicionantes del alumno que no cambiarían, aunque fuera a otro centro. Pudiera ser que incluso, el efecto causal fuera del signo contrario a la diferencia en los resultados medios observados. Pudiera suceder que centros que tienen que atender, por su situación geográfica, por ejemplo, a alumnos con mayores dificultades educativas cuenten con equipos docentes mejores que aporten un mayor valor añadido al aprendizaje del alumno. Pero en media sus alumnos obtienen resultados algo peores. Aunque el otro centro es de peor calidad (aporta poco valor añadido al aprendizaje de sus alumnos), las caracterísitcas de los alumnos del otro centro más que compensan las carencias de su centro con otros insumos del aprendizaje. Ejemplos de esos otros insumos son el esfuerzo de los alumnos, o los refuerzos y apoyos educativos que sus familias les pueden dar al margen del centro educativo. Debido a este sesgo de selección en la composición del alumnado, la comparación de resultados observados podría llevarnos, ya no solo a estimar un efecto de cuantía distinta, sino a estimar un efecto causal del centro educativo de signo contrario al que realmente tiene, infiriendo incorrectamente que el centro con mejor nota media es el de mas calidad de los dos.*

**El sesgo de selección como parte del problema de endogeneidad:** el sesgo de selección es parte de un problema que en términos econométricos se conoce como endogeneidad. Hay endogeneidad cuando a la vez que cambia *D*, cambian otras cosas que también afectan a la *Y* y por las que no estamos controlando. Las tres fuentes de endogeneidad son:

1) Variables (características) omitidas que afectan a la *Y* y están relacionas con *D*. El ejemplo 1 sería un caso de sesgo de variables omitidas. En la comparación entre centros se han omitido caracteristicas de los alumnos que hacen parecer a un centro mejor que otro, cuando en realidad se debe a las distintas caracterísitcas de los alumnos y sus familias.

2) Simultaneidad: tanto *Y* como otros factores que no observamos se determinan a la vez y se afectan mutuamente y están relacionados o incluso afectados por la *D*. Un ejemplo de simultaneidad se daría cuando queremos determinar el efecto de una intervención *D* sobre el precio *Y* de un mer-

cado. Entre los determinantes de $Y$ estaría la cantidad demandada en este mercado. Pero precio y cantidad vendida se determinan simultáneamente (conjuntamente), y la intervención $D$ que puede afectar al precio también afectaría a la cantidad vendida. Tendríamos un sesgo en la estimación de los efectos. Otro ejemplo de simultaneidad surge cuando se quieren estudiar los efectos de los compañeros o *peer-effects*. La $Y$ de cada estudiante afecta y se determina simultáneamente con las $Y$ de los compañeros, cuyo efecto sobre este estudiante se querría conocer (véase el capítulo 4 de este libro para más detalles sobre esta situación).

3) Error de medida en $D$: la variable que observamos es una medida imprecisa de $D$. La diferencia entre esa medida y el verdadero valor de $D$ que determina la $Y$ es el error de medida. Incluso en el caso en el que el error de medida es algo completamente exógeno esto va a generar un sesgo. Una parte de la variación en la $D$ que observamos viene dado por variación en el error de medida. Pero esa variación no va a implicar un cambio en $Y$ porque esta viene determida por el verdadero valor de $D$, no por el error de medida. Eso significa que asociaremos cambios en $D$ con que la $Y$ no cambie, estimando menor efecto de la $D$ del que realmente haya. Es lo que se conoce como un sesgo de atenuación. Si, además, el error de medida estuviera correlacionado con algún factor que afecte a la $Y$ el sesgo podría ir en cualquier dirección.

## 4. ENSAYOS O EXPERIMENTOS ALEATORIOS CONTROLADOS (*RANDOMIZED CONTROL TRIALS)*

La primera solución al sesgo de selección es impedir que los individuos se autoseleccionen o los seleccionen en base a cosas que afectan al resultado. En este método la asignación se hará de manera aleatoria e independiente de cualquier característica (relevante) y resultado potencial de los individuos. Esto nos garantiza por construcción que el grupo de tratados y de control sean comparables. Y con comparables no me refiero a que sean exactamente iguales, sino a que la distribución de características sea la misma (en un sentido estadístico). Esto incluye, entre otras cosas, que ambos grupos sean iguales en media. En el ejemplo de comparar dos colegios para conocer su calidad, el nivel de renta de las familias será en media (estadísitcamente) el mismo si los alumnos se han asignado a cada centro de forma aleatoria. Si se observa alguna diferencia en la nota media de los alumnos de cada centro en el examen que haces después de haber estado en ese centro, esta no puede de-

berse a que tengan distinta renta familiar, puesto que esta característica es la misma en ambos. La aleatorización garantiza la comparabilidad entre tratados y controles, y que podamos interpretar la diferencia de resultados medios entre un centro y otro, como debidas al centro educativo.

Este método es considerado el *gold standard* en la evaluación de políticas públicas. Ensayar una polítca, implementándola primero mediante un experimento aleatorio controlado, o RCT por su acrónimo en inglés, nos permitirá conocer el efecto causal de esa política, separado de otros factores cuyo efecto pudiéramos atribuir erróneamente al resultado, solucionando, por tanto, el problema del sesgo de selección. Pero tiene sus limitaciones, que describimos en las siguientes subsecciones.

### 4.1. Limitaciones en la implementación práctica

Estas limitaciones pueden surgir ante cualquier experimento y pueden llevar a situaciones en las que esta solución para conocer el efecto causal no se pueda implementar:

- Dificultades para implementarlos por consideraciones éticas y/o físicas. Por ejemplo, no es viable ni ética ni físicamente, decidir aleatoriamente quién tendrá estudios Universitarios y quién no, como forma de conocer el impacto causal que los estudios Universitarios tienen sobre la renta laboral de las personas. Primero, no se consideraría éticamente aceptable impedir por un experimento que personas, que están academicamente preparadas y que quieren ir a la Universidad, fueran. El coste para toda la vida de estas personas de ese experimento sería demasiado grande. Tampoco sería muy fácil obligar a alguien que no quiere estudiar a conseguir tener estudios universitarios. Es por esto, que la (importante) literatura que intenta conocer la rentabilidad (o impacto sobre la productividad) de los estudios, utiliza otros métodos.
- Alto coste del experimento. Requiere poner muchos medios económicos nuevos, no solo por la política a ensayar, sino por los requerimientos para implementar y seguir el experimento. Esto a veces se traduce en que sólo se implementa en una muestra demasiado pequeña, lo que lleva a que sus resultados no sean fiables o sean demasiado imprecisos.

- Dificultad para cumplir con la implementación aleatoria. Autoselección posterior a la aleatorización. Dos posibles desviaciones de la asignación aleatoria: a) algunos de los individuos asignados al grupo de tratamiento no quieren ser tratados y no lo son; y b) algunos de los controles, consiguen recibir el tratamiento, aunque se supone que no lo pueden recibir. La desviación respecto a la asignación aleatoria rompe la comparabilidad entre tratados y controles porque puede ser debida a autoselección. Aún así, si se observa a todos, se conoce bien quién ha sido realmente tratado y su asignación inicial, estos datos pueden seguir usándose para obtener un efecto causal. Pero, habrá que usar otro método para comparar tratados y controles y estimar el efecto causal. Esto lo veremos en la solución de variables instrumentales. Nuestra capacidad para estimar efectos causales de interés será mayor cuando sólo se dé una de esas dos desviaciones, y cuanto menor sea el número de individuos que se desvían.
- Dificultad para tener información sobre los individuos a lo largo del tiempo (*attrition*). El ensayo suele requerir implementar la política a lo largo de un periodo de tiempo, y hay que seguir y observar a tratados y controles antes, durante y después. En particular, hay que recoger información sobre la variable de resultados para todos lo individuos después de haber implementado la política. Desde la selección aleatoria de tratados y controles, algunos de ellos han podido perder interés en seguir colaborando con el experimento, o mudarse y dejar de estar localizables. En la medida en que esto no afecte a una proporción grande, y, sobre todo, se produzca de forma independiente a los resultados potenciales al experimento, esto no supone una amenaza para la validez del método. Pero siempre que haya *attrition*, habrá que intentar comprobar que seguimos teniendo grupos de tratados y controles que son comparables, representativos, y de tamaño suficiente para hacer estimaciones precisas. El ensayo no dará resultados válidos si, por ejemplo, son los individuos de determinado nivel de renta los que más frecuentemente dejan de cooperar con el ensayo.

### 4.2. Problemas de validez externa del experimento

Si el ensayo aleatorio se ha implementado correctamente nos dirá el efecto causal del tratamiento que se ha ensayado, sobre el grupo objeto del ensayo, y en el momento del ensayo. A esto se le llama validez interna de la estimación. Y no hay método que nos vaya a dar una mayor validez interna. Pero, nuestro objetivo no

es ese. Nuestro objetivo es conocer el efecto causal de una política que vamos a aplicar sobre toda la población de interés, y conocerlo no sólo en el momento del ensayo, sino en otros periodos también. Pensemos en un ensayo que es un proyecto piloto de una política que se aplica a un grupo pequeño de personas para evaluar su extensión a toda la población en años siguientes. Además, muchas veces, pretendemos utilizar resultados de experimentos llevados a cabo en el pasado en otro lugar del mundo, para saber el efecto causal de una política similar en otros lugares y momentos.

Todos esos objetivos, se refieren a la llamada validez externa del ensayo, y esta no está garantizada por mucho que se den las circunstancia que garantizan la validez interna de la estimación del efecto causal. Sin supuestos adicionales, sólo se puede conocer el efecto del programa tal cual fue implementado, y sólo en el momento en el que el experimento fue implementado. Otro programa, en otro momento, o para otra población podría dar resultados distintos.

Las principales amenazas a la validez externa son:

1) El experimento no es suficientemente similar a la política que se va a aplicar y cuyo efecto queremos conocer.

2) La muestra sobre la que se aplica el experimento es distinta (no representativa) de la población sobre la que se quiere extender. Los efectos causales son potencialmente heterogéneos, distintos entre individuos, y, por lo tanto, la política tendrá unos efectos causales distintos de los estimados en el grupo de individuos del ensayo.

3) Existencia de externalidades (en este contexto, efectos del tratamiento para otras personas distintas de las tratadas). En esta situación, la comparación de $Y$ entre tratados y controles no nos dirá todo el efecto causal del tratamiento sobre la población, puesto que no tendrá en cuenta estos efectos externos. Si, además, estos efectos externos afectan también al grupo de control, este grupo no será válido para decirnos lo que les habría ocurrido a los tratados en ausencia de tratamiento. Por lo tanto, el ensayo no estimará correctamente el efecto causal, ni siquiera sobre los tratados. No se darían las condiciones ni para tener validez interna de la estimación.

4) Si hay efectos de equilibrio general una vez se pasa del experimento a la universalización de la política. Efectos de equilibrio general se dan cuando, como consecuencia de la implementación del programa a gran escala, los

equilibrios de los mercados se ven afectados, porque se está alterando la oferta y/o demanda de algunos bienes. Por explicarlo mejor con un ejemplo, pensemos en que queremos conocer el efecto total sobre los resultados académicos de reducir el tamaño de clase. Hacemos un ensayo en algunos grupos de algunos centros educativos y estimamos un efecto positivo. Si después aplicamos esta reducción del tamaño de grupo como política general a todos los grupos y centros el país o región, encontraremos que necesitamos contratar un gran número de profesores nuevos para cubrir los nuevos grupos que se generan en todos los centros por la reducción del número de alumnos en cada grupo. Necesitar contratar unos pocos profesores en unos pocos centros en el ensayo no altera el mercado de profesores. Pero al extenderlo a todo el país, se necesitan muchos nuevos, y los requisitos y calidad exigida a los profesores que se contratan en el nuevo equilibrio de mercado podría ser inferior. Esto generaría un efecto negativo de la política que no se podría detectar en el ensayo.

La validez externa está todavía más en cuestión cuando lo que se quiere es aprender sobre los efectos de una nueva política que tenga los mismos objetivos que tenía la ensayada porque se ha visto que la ensayada no funciona. Este método estima de manera fiable el efecto causal total "interno" al experimento. Pero, al no aprender nada sobre los mecanismos que producen el resultado observado, no podemos usarlo para pensar qué otra política podría aplicarse que fuera más efectiva, o conocer la razón del resultado (fallido) observado.

Algunas de estas limitaciones son comunes a muchos de los métodos que veremos a continuación, puesto que son inherentes a la estimación de un efecto causal total o de forma reducida, pero otras son específicas de los ensayos aleatorios. Los siguientes métodos van a seguir intentando solucionar el problema del sesgo de selección en situaciones en las que no tenemos o no podemos hacer un ensayo para conocer el efecto causal de una variable.

## 5. CONTROLAR POR VARIABLES OBSERVABLES: REGRESIÓN CON CONTROLES, MÉTODO DE EMPAREJAMIENTO (*PROPENSITY SCORE MATCHING* )

El objetivo sigue siendo tener un grupo de control que nos diga lo que les habría pasado a los tratados en ausencia de tratamiento, solucionando el sesgo de selección que hace que tratados y controles difieran en más dimensiones (rele-

vantes) que en haber recibido el tratamiento. Si no tenemos un experimento que nos garantice esto y observamos que ambos grupos tienen características distintas, una posible solución es controlar por esas características observables. Si el sesgo de selección al comparar la *Y* media de dos centros educativos se debe a que en uno los alumnos tienen padres con mayores estudios que en el otro, en lugar de comparar a todos los alumnos de un centro con todos los alumnos del otro centro (que implica que comparamos alumnos con distintas características familiares), podemos comparar alumnos de uno y otro centro que tengan padres con el mismo nivel educativo. De esta forma estamos emparejando tratados con controles que tienen exactamente las mismas características observables, controlando por las diferencias que en el conjunto de los alumnos varían por los diferentes valores de esta variable entre un centro y otro.

Cuando se tengan varias característcas por las que controlar y/o estas tomen muchos posibles valores, este emparejamiento de grupos con exactamente las mismas características no será posible. Pero la misma idea se puede aplicar con dos métodos econométricos: el método de regresión, y el de emparejamiento o *matching* mediante la probabilidad de participar en el tratamiento o *propensity score*. Las dos subsecciones siguientes los resume, y la tercera los relaciona.

**Supuesto clave:** controlar por variables observables, *X*, nos va a dar una estimación del efecto causal si la selección, que hace diferentes a tratados y controles, se basa exclusivamente en esos observables. Dicho de otra forma, si ambos grupos son iguales (en media) en características inobservables. A efectos prácticos, los inobservables incluye todo aquello que por lo que no controlamos y que determina la *Y*. En la sección 11.2 se encuentra la expresión formal de este supuesto, así como el papel que juega en los métodos de estimación que se describen a continuación.

### 5.1. Regresión con controles

La regresión lineal es uno de los métodos mas usados para encontrar relaciones entre variables. Tiene varias propiedades interesantes, incluyendo la sencillez en su aplicación. El problema está en ser capaces de interpretar esas relaciones como relaciones causales. El supuesto clave para tener un efecto causal es el que se ha dicho arriba, de no selección en inobservables. Pero, además de eso, se hace un supuesto sobre la forma (funcional) en la que los controles *X* afectan a la *Y*. Este supuesto de forma funcional hace, de forma implícita, una extrapolación del

efecto de esas $X$ sobre la $Y$ para todos los posibles valores de las $X$. Esto nos puede dar una falsa idea de comparabilidad entre tratados y controles. Por ejemplo, si una de esas $X$ es renta de los padres, la regresión lineal va a suponer una misma relación con la $Y$ para todos los valores de $X$, como

$$Y = \beta_0 + \beta_1 \, ln(X) + \tau D + \varepsilon \tag{1}$$

y estimarla por Mínimos Cuadrados Ordinarios usando los datos disponibles. El problema puede surgir cuando los valores de $X$ en el grupo de tratados sean muy distintos a los valores de $X$ en el grupo de control. En concreto, haya niveles de renta solo presentes en uno de los dos grupos. En el ejemplo, puede ocurrir que no haya nadie con niveles de renta más altos en el grupo de control, pero sí en el de tratados. En esta situación estamos suponiendo que la relación en (1) nos da también lo que les ocurriría a los tratados con un nivel de renta ($X$) más alto en ausencia de tratamiento, a pesar de que no observamos a nadie con esos niveles que no haya sido tratado. Se está extrapolando el efecto para unos valores de $X$ que no están en el grupo de control. Si la extrapolación es correcta, no hay ningún problema. Pero si no lo es, estaremos estimando mal el efecto causal. En la última subsección se discute cómo se puede evitar hacer esta extrapolación.

**Endogeneidad en la regresión:** aunque nuestro interés está en el efecto de $D$, al introducir $X$ tenemos que preocuparnos también porque estas no sean endógenas y porque estas no capturen parte del efecto de la $D$. En caso de que alguna sea endógena o capture parte del efecto de interés, habrá que utilizar otro método de estimación para resolver esos problemas. Esto significa que estoy en una situación en la que por un lado es necesario incluir la $X$ por la selección en $X$ entre tratados y controles, y por otro lado al incluirla tengo un nuevo problema. La conclusión es que no puedo simplemente usar este método para estimar el efecto causal. En algunos contextos a este tipo de regresores se les ha denominado "malos controles" (*bad controls*), pero esta terminologíia induce a pensar que la solución es excluirlos del grupo de controles. Eso raramente será la solución. Sólo en un contexto en el que los controles no sean necesarios para obtener un efecto causal (aunque sea de forma reducida), podremos optar por no incluirlos porque estaríamos peor. Pero cuando se recurre al uso de controles porque tenemos sesgo de selección (en observables), no se puede prescindir de esos controles. Si estos son endógenos, o si capturan parte del efecto de interés, habrá que buscar la forma de, además, solucionar esos problemas para estimar el efecto causal de interés. Pero excluirlos no será la solución.

### 5.2. Método de emparejamiento basado en la probabilidad de tratamiento (*propensity score matching* )

El método de emparejamiento *(matching)* va a emparejar tratados con controles que puedan informarnos sobre el valor de $Y$ que habríian tenido los tratados en ausencia de tratamiento. La idea es que tratados y controles emparejados tengan los mismos valores de $X$. Cuando $X$ toma un conjunto grande de posibles valores (por ejemplo, por tener varias $X$), usamos el resultado siguiente: Si emparejar con base en las variables observables $X$ nos permite estimar un efecto causal, entonces el emparejamiento con base en $P(X)$ también lo hace. $P(X)$ nos da la probabilidad de que un individuo con caracteríisticas $X$ sea tratado, es decir, $P(X) = Pr\ (D = 1|X)$, denominado en inglés *propensity score*.

La intuición detrás de este resultado es que el problema viene porque esas variables $X$ hacen a tratados y controles no comparables en la medida en que afectan a quien es tratado y a quien es control. Si no hubiera selección en $X$ (y por lo tanto no necesitara emparejar por valor de $X$), la probabilidad de ser tratado sería la misma para cualquier valor de $X$. Esto es, $Pr\ (D = 1|X) = Pr\ (D = 1)$. Esto ayuda a entender porque si un individuo en el grupo de control tiene caracteríisticas tales que su Pr $(D = 1|X)$ es la misma la de otro individuo en el grupo de los tratados, estos van a ser un buen emparejamiento, sin necesidad de exigir que las $X$ sean todas exactamente las mismas.

Además de lo anterior, en la práctica habrá que decidir qué significa tener la misma $Pr\ (D = 1|X)$. Los detalles de la implementación se escapan de lo que este capíitulo puede cubrir. Baste con decir que existen varios criterios y métodos estudiados y disponibles en distintos programas informáticos, para estimar $Pr\ (D = 1|X)$, para decidir cómo emparejar en función de la cercaníia del valor de $Pr\ (D = 1|X)$, y para, finalmente, estimar el efecto medio del tratamiento desde ese emparejamiento.

Un paso previo al emparejamiento y la estimación del efecto causal en el que la literatura del *matching* ha puesto mucho énfasis y que es importante para evitar el problema de la extrapolación que mencionaba al presentar la regresión, es la comprobación/imposición de la llamada condición de soporte común. Hay que comprobar que el rango de valores de $Pr\ (D = 1|X)$ es el mismo en tratados y controles. Y si no lo es, debemos limitar nuestra estimación a usar sólo aquellos tratados y controles que estén dentro del mismo rango de valores. Por supuesto, el efecto causal medio que estimemos no será el efecto causal medio para toda

la población ni para todos los tratados. Será sólo para aquellos que cumplen la condición de soporte común. Si en el grupo de los tratados tenemos personas de renta alta y en el de control no tenemos a nadie, no podemos conocer el efecto de $D$ para los individuos de renta alta porque no tenemos ningún control que nos diga su contrafáctico (sin hacer supuestos sobre cómo cambia el efecto del tratamiento con la renta). Pero reconocer que no podemos obtener el efecto causal para ese grupo, e imponer la condición de soporte común, al menos sí nos va a permitir conocer el efecto causal para el resto.

### 5.3. Comparación entre la regresión y el método de emparejamiento

Ambos comparten lo más importante: decidir las variables por las que tengo que controlar y suponer que dadas esas $X$ no hay selección en inobservables. En base a este supuesto fundamental para la identificación de un efecto causal ninguno de los dos métodos representa una ventaja sobre el otro. Las ventajas vienen por otros supuestos, y son, en buena parte prácticas.

- En los dos hay que hacer algunos supuestos de forma funcional. En la regresión se supone la relación de $Y$ con la $X$ y la $D$. En el emparejamiento, esa relación no se especifica ni restringe, pero se supone una forma funcional sobre $Pr\ (D = 1|X)$.
- En el emparejamiento hay que tomar una decisión arbitraria sobre el criterio de emparejamiento.
- La estimación y cálculo de los errores estándar correctos es mucho más difíicil en el emparejamiento que en la regresión.
- La gran desventaja de la regresión con respecto al emparejamiento es que tradicionalmente la regresión ha ignorado la condición de soporte común y ha extrapolado a valores de las $X$ no contenidos en ambos grupos cuando esta no se cumple. Pero no hay ninguna razón para que la condición de soporte común no se pueda imponer de forma previa a la estimación de la regresión al igual que se hace en el emparejamiento.

Dado la anterior, estimar una regresión lineal en la muestra que cumpla la condición de soporte común, y usando una forma funcional flexible respecto a como entran la $D$ y las $X$ (polinomios, interacciones, etc.), puede ser igual de bueno que el emparejamiento y mucho más fácil de implementar.

### 5.4. Cuando no se cumple el supuesto de selección sólo en observables

Cuando nos preocupa que pueda haber selección en inobservables, y no sólo en observables, sabemos que los estimadores de esta sección no nos van a dar una buena estimación del efecto causal. Aún así, podemos usarlos para hacernos una idea de la gravedad del problema de selección. Comparando cómo cambia el efecto estimado de $D$ cuando se incluyen y se excluyen las $X$, podemos tener una indicación de la gravedad del problema de selección, y, según el caso, acotar el efecto. Por ejemplo, si tenemos un conjunto rico de $X$ y la estimación del efecto apenas cambia al estimarlas, es razonable suponer que el problema de selección no es muy grande y que estamos cerca de tener una buena estimación. Véase la bibliografía para más detalles sobre esto.

## 6. MÉTODO DE DIFERENCIAS EN DIFERENCIAS (DiD)

El principal problema con el método anterior es que es difícil pensar situaciones en las que el problema de selección se deba solo a variables observables. A no ser que observemos un conjunto muy amplio de características relevantes, lo normal será que si hay diferencias observables entre tratados y controles también las habrá en características que no observamos. Y cuanto mayor sean las diferencias, mayor será la necesidad de solucionar el sesgo de selección y más difícil será solucionarlo solo con los observables. El método de esta sección permite controlar por diferencias entre el grupo de control y el de tratamiento, sean estas observadas o inobservadas, siempre que en ausencia de tratamiento las diferencias se hubieran mantenido constantes, o tratados y controles hubieran evolucionado de forma paralela.

Si hay sesgo de selección, eso significa que no podemos comparar en un mismo momento del tiempo los resultados de tratados con los de los controles para saber el efecto del tratamiento. Si observamos los resultados de todos también en un periodo anterior, además de en un periodo posterior, al tratamiento, podríamos pensar en comparar el resultado de los tratados antes y después del tratamiento. Al tratarse de los mismos individuos es más razonable suponer que sus características son las mismas. El problema es que puede haber cosas que afecten al resultado y que cambien en el tiempo. Y al tomar esta diferencia temporal en resultados, podemos atribuir al tratamiento lo que se debe al cambio de un periodo a otro de otros factores que afectan a $Y$. El método que se expone a continuación intenta solucionar esto.

### 6.1. DiD básico y supuesto de identificación

Para solucionar los dos problemas que nos impiden estimar haciendo una sola diferencia (que tratados y controles son distintos, y que hay más cosas que cambian de un periodo a otro), podemos hacer una doble diferencia. Esta doble diferencia, y el supuesto de identificación que se requiere para identificar el efecto causal con DiD, se puede expresar de dos formas (que, por la propiedad conmutativa de la suma, son equivalentes):

1) [(*Y* media de tratados después) – (*Y* media de controles después)] – [(*Y* media de tratados antes) – (*Y* media de controles antes)]. Aquí, la segunda diferencia se ocupa de aquellas cosas que hacen a los tratados diferentes de los controles, y, por lo tanto, provoca el sesgo de selección. El supuesto de identificación en esta forma de escribirlo significa suponer que las diferencias preexistentes entre tratados y controles se habrían mantenido constantes en el periodo posterior de no haber existido tratamiento. Suponemos que cualquier cosa que haya cambiado esa diferencia pre-existente se debe al tratamiento.

2) [(*Y* media de tratados después) – (*Y* media de tratados antes)] – [(*Y* media de controles después) – (*Y* media de controles antes)]. Aquí, la segunda diferencia se ocupa de aquellas cosas que cambian de un periodo a otro. El supuesto de identificación en esta forma de escribirlo significa suponer que las cosas que habrían cambiado de un periodo a otro para los tratados, si no hubieran sido tratados, serían iguales a lo que han cambiado para los controles. Usamos a los controles, para descontar del cambio de un periodo a otro lo que no se debe al tratamiento. En la seccion 11.3 se escribe de manera formal este supuesto de identificación y lo que implica para el estimador DiD.

En sentido estricto, el supuesto de identificación no se puede contrastar, y, por lo tanto, solo se puede suponer y argumentar y dar evidencia a favor o en contra de su validez. Parte de la evidencia que es necesario aportar a favor o en contra de este supuesto, es el comprobar que el supuesto se cumple en periodos en los que nadie ha recibido el tratamiento. Esto se puede hacer cuando se observan varios periodos anteriores al tratamiento. Si en esos periodos no se ve qué tratados y controles evolucionan de forma paralela, eso haría que no fuera nada creíble el supuesto de que justo en el periodo en el que se recibe el tratamiento sí que habrían evolucionado de forma paralela. Esto, que a veces se le llama test de tendencias paralelas, no es un contraste formal de la hipótesis. Que el supuesto se

cumpla en periodos anteriores, no significa que en el periodo en el que se recibe el tratamiento se tenga que cumplir también. Pero, sí se exige ese test como requisito para considerar esos resultados como válidos.

Una de las situaciones en las que el supuesto de identificación es más plausible es cuando el cambio en $D$ se ha producido por un evento natural, exógeno a la voluntad de los individuos. Por ejemplo, un terremoto puede generar un cambio en el patrimonio de los individuos y ser usado para conocer el impacto de un cambio en el patrimonio sobre una variable de resultado de interés. Cualquier grupo de control que busquemos va a ser probablemente distinto a los afectados, pero, si presentan tendencias paralelas anteriores, pasa a ser más creíble que en ausencia del terremoto la evolución hubiera seguido siendo paralela. Por este motivo al método de DiD también se le conoce como experimentos naturales o cuasiexperimentos.

La forma más sencilla de obtener el estimador de DiD, $\widehat{\tau_{DiD}}$ , es mediante la siguiente regresión

$$Y = \beta_0 + \beta_1 D + \beta_2 Post + \tau_{DiD}\,(D * Post) + \varepsilon \qquad (2)$$

donde $D$ indica si es una observación del grupo que va a ser tratado, *Post* indica si es una observación del periodo posterior al tratamiento, y, por lo tanto, la interacción entre las dos, $(D * Post)$, indica si es una observación que ha recibido el tratamiento.

## 6.2. Extensiones del método de DiD básico

### *6.2.1. Varios periodos: efectos distintos en el tiempo y distintas tendencias*

Cuando se dispone de varios periodos posteriores al tratamiento se puede estudiar cómo el efecto del tratamiento cambia a lo largo del tiempo. Si se dispone de varios periodos anteriores al tratamiento, además de contrastar si hay tendencias paralelas, se podría permitir, si se considera más realista que, en lugar de seguir tendencias paralelas, tuvieran una velocidad de crecimiento paralela, haciendo, por ejemplo, una triple diferencia, u otras formas más paramétricas de permitir algunas diferencias en las tendencias entre tratados y controles.

### *6.2.2. DiD controlando además por diferencias observables*

Se puede combinar el DiD con los métodos que controlan por diferencias entre tratados y controles que observamos. El supuesto de tendencias paralelas no se tendría que cumplir en la comparación de todos los tratados con todos los

controles. Bastaría con que se cumpla condicionado a tener el mismo valor de los observables. Es posible que los alumnos con padres con distintos niveles educativos, no solamente tengan un nivel distinto, sino que sigan una evolución distinta en ausencia de tratamiento. Esta distinta evolución invalidaría el DiD si la distribución de niveles educativos de los padres es distinta entre tratados y controles. Pero, pudiera ser que sí hubiera evolución paralela (en ausencia de tratamiento) si comparamos alumnos cuyos padres tienen el mismo nivel educativo.

Dado que el estimador de DiD se puede obtener a través de la regresión (2), la forma más sencilla de controlar por observables es incluyendo las variables observables, $X$, como regresores adicionales en (2). Un inconveniente de esto es que impone restricciones de forma funcional y necesitamos un supuesto de tendencias paralelas más fuerte para que su validez no dependa de esa forma funcional (ver las referencias que se indican al final del capítulo para más detalles).

Otra forma de controlar por observables es combinar el DiD con el método de emparejamiento: dobles diferencias emparejadas. Se aplica el método de emparejamiento, no a las $Y$ de los dos grupos, sino a las dobles diferencias, de forma que se hace una doble diferencia para cada valor de los observables, o de $P(X)$ si se usa la probabilidad de tratamiento para emparejar.

### 6.3. Importancia de las pruebas de falsificación y placebos

Dado que el supuesto de identificación en el DiD no es contrastable, y que muchas otras cosas (además de la implementación del tratamiento) pueden cambiar de un periodo a otro y afectar de forma diferente a tratados y controles, es importante realizar pruebas de falsificación. Que se superen estas pruebas no nos va a demostrar que el supuesto se cumple, pero si no las superan, eso sí demostraría que el supuesto no se cumple. También se puede y conviene hacer estas pruebas de falsificación en otros de los métodos mencionados en este capítulo, como los estimadores basados en la selección sólo en observables.

Hay dos tipos de pruebas de falsificación:

- Se usa un placebo como indicador de tratamiento. Se consideraría un tratamiento que nunca ha existido, como si lo hubiera. Por ejemplo, si estamos evaluando el efecto de un programa educativo bilingüe sobre el rendimiento educativo, esta prueba de falsificación tomaría a un grupo de alumnos y centros que no han sido tratados y consideraría (falsamente)

que alguno de ellos sí ha sido tratado y tendrían $D = 1$. Si en esta muestra en los que nadie es de verdad tratado se encontrara un efecto de esta falsa $D$, eso indicaríia que nuestro método de estimación está recogiendo el efecto de otras cosas distintas del tratamiento.

- Se usa como resultado de interés, $Y$, uno que **no** ha podido ser afectado por el tratamiento. Por ejemplo, los años de estudio de los padres, no debería verse afectado por el tamaño de clase de los centros educativos de sus hijos. Si en vez de usar como $Y$ el resultado académico de los alumnos, usamos como $Y$ los años de educación de los padres no deberíamos encontrar ningún efecto de $D$. Si hay efecto, es que nuestro método de estimación le está atribuyendo incorrectamente a $D$ lo que es un efecto debido a otras cosas (como por ejemplo al diferente nivel educativo de los padres que existe entre tratados y controles).

## 7. VARIABLES INSTRUMENTALES

En los casos en los que hay sesgo de selección en variables que no observamos y, bien porque no tenemos los datos necesarios para hacer el DiD, bien porque el supuesto de tendencias paralelas o de selección en observables no sea creíble, no podemos aplicar ninguno de los métodos anteriores. Una solución alternativa es el uso de variables instrumentales, si se dispone de una variable, a la que llamamos instrumento y denotamos con $Z$, que cumpla los siguientes supuestos, cuya expresión formal puede encontrarse en la sección 11.4:

**Supuesto 1 (S.1. Independencia del instrumento).** *El instrumento es independiente de los resultados potenciales y de las asignaciones a tratamiento potenciales.*

**Supuesto 2 (S.2. Restricción de exclusión).** *El resultado Y, dado un valor del indicador del tratamiento, es sólo función de este y no lo es del instrumento. Esto es, dada una situación de tratamiento (tratado o no tratado), el valor de Y no depende del valor del instrumento. El instrumento está relacionado con el resultado Y sólo y exclusivamente a través del efecto que Z tiene sobre ser tratado.*

**Supuesto 3 (S.3. Condición de rango o de relevancia del instrumento).** *El instrumento afecta a la proporción de tratados. Esto es, la proporción de tratados cambia con los valores del instrumento.*

**Supuesto 4 (S.4. Supuesto de monotonicidad).** *El instrumento afecta a la probabilidad de ser tratado en la misma dirección a todos los individuos. Bajo este supuesto, puede que algunos individuos no vean afectada su decisión (o situación) de ser tratados por el instrumento, pero a aquellos a los que les afecta, lo hace para todos en la misma dirección Pr (D = 1|Z) aumenta con Z para todos, o disminuye con Z para todos, pero no se da la situación en la que para unos aumenta y para otros disminuye).*

Para discutir sobre la plausibilidad de estos supuestos, y sobre el efecto causal que identificamos, voy a usar un ejemplo. Supongamos que queremos evaluar el efecto de las academias de idiomas ($D$) sobre el aprendizaje de inglés ($Y$). No hemos realizado ningún experimento para decidir aleatoriamente quién va a esas academias y quién no, por lo que la comparación entre tratados y controles nos dará el efecto sobre $Y$ no solo de las academias, sino de las muchas características que hacen que el grupo de los que van a las academias sean distintos de los que no. Pero lo que si hemos hecho es dar unas becas para pagar la totalidad de lo que cuesta pagar esas academias ($Z$) y que sólo se pueden gastar en el pago de las academias. Un año después medimos el nivel de inglés de una muestra representativa de la población, lo que será nuestra medida de $Y$, además de saber quién de ellos ha ido a la academia (valor de $D$) y quién ha recibido la beca ($Z$), haya hecho uso de ella o no.

Si la beca la asignamos en función de la renta (por ejemplo, se la damos solo a familias por debajo de un determinado umbral de renta), no se va a cumplir el supuesto S.1. puesto que recibir la beca ($Z = 1$) no va a ser independiente de los valores potenciales de $Y$. Una forma de ver esto es darse cuenta de que, si comparamos la $Y$ de los que reciben la beca con la $Y$ de los que no la reciben, estamos comparando el nivel de inglés de familias con menos renta con el nivel de inglés de familias con más renta, que pueden tener ese nivel de inglés, no porque no hayan recibido la beca, sino porque tienen más recursos para destinarlos al aprendizaje de los idiomas. Mediríamos el efecto de la renta y no sólo el efecto de las academias. El supuesto S.1. sí se cumpliría si, por ejemplo, decidimos la asignación de una beca mediante un sorteo. En este caso, perceptores y no perceptores tendrían en media las mismas características, y, como beneficio colateral, podemos conocer el efecto causal de la beca sobre el aprendizaje de idiomas. De hecho, a ese efecto de $Z$ sobre $Y$ se le conoce como ITT por su acrónimo en inglés (*Intention To Treat Effect* ). Pero nuestro objetivo no es conocer el impacto de la beca sobre el

aprendizaje de idiomas. Nuestro objetivo es conocer el efecto de las academias, y, para este objetivo, no es suficiente con el supuesto S.1.

El supuesto S.2. implica en este ejemplo que la beca no tenga un impacto directo en el aprendizaje de idiomas (lo cual es bastante creíble, dado que la mera concesión de la beca no cambia el nivel de inglés de los alumnos), y que tampoco afecte a ninguna otra variable que afecte a $Y$, lo que ya no está tan claro. El supuesto S.2. requiere que la única relación entre $Z$ e $Y$ se dé únicamente a través del impacto que $Z$ pueda tener en $D$, operando como único efecto causal hacia $Y$ el que $D$ pueda tener. Pero si la concesión de la beca ($Z = 1$), hace que, por ejemplo, algunas familias destinen la renta que le ha quedado libre (dado que habría incurrido en el gasto en academias sin la beca) a cursos de idiomas en el extranjero, entonces al comparar el nivel de inglés entre los que reciben la beca y los que no, el efecto que encontramos no se debe solamente a, o no opera solamente a través de, las academias de idiomas. Se deberá, al menos en parte, al efecto que los cursos de idiomas en el extranjero puedan tener. Somos capaces de estimar bien el efecto total de $Z$ sobre $Y$ porque se cumple S.1., pero no podemos estimar el efecto de $D$ (aislado del efecto de otras variables) porque no se cumple S.2. Para estimar el efecto de S.2. necesitamos suponer que el mecanismo que he dicho no se da, y que el único efecto que la beca puede tener consiste en aumentar la probabilidad de $D = 1$.

La condición de relevancia, no es un supuesto que tengamos que hacer, puesto que se puede comprobar. Es una condición que necesitamos, pero podemos comprobar en los datos si se cumple o no. Si la beca no hace mas probable que la gente vaya a las academias, entonces no va a haber ninguna diferencia en el tratamiento, y no podemos usar esto como instrumento para conocer el efecto de $D$. El instrumento tiene que ser una variable que haga que haya variación en $D$, para poder conocer ese efecto.

El supuesto de monotonicidad lo que implica en este caso es suponer que no hay nadie al que recibir la beca le hace no apuntarse a las academias. A todos, recibir la beca hace que, frente a no haberla recibido, o bien toma la misma decisión respecto a apuntarse a la academia, o bien hace más probable que decida apuntarse. Dado que este estimador va a comparar la $Y$ de los que reciben la beca con la de los que no, y a corregir ese dato por cómo la beca ha hecho que aumente la proporción de los que van a la academia (por la condición S.3. ese aumento no puede ser cero), lo que nos va a dar es una media ponderada del efecto causal de $D$

para aquellos individuos que han cambiado su *D* como consecuencia de haber recibido la beca. Necesitamos que recibir la beca tenga un efecto monótono sobre *D*, para que ese estimador no nos de una diferencia de efectos causales, sino un efecto causal medio.

Una limitación del método de variables instrumentales es que, con estos 4 supuestos, el efecto causal que estima no es el efecto medio para toda la población, ni el efecto medio para los tratados (los que van a las academias), que suele ser el que más interesa; sino que nos va a dar el efecto medio solo para aquellos a los que el instrumento ha cambiado su valor de *D*, ponderando en esa media más a aquellos para los que el instrumento tiene más efecto sobre *D*. Esto hace que este efecto no sea relevante para todos aquellos que habrían acudido a la academia, aunque no hubieran recibido la beca, ni que tampoco les podamos decir a los que no van (reciban o no reciban la beca) el efecto sobre *Y* que están dejando de obtener. Para conocer esas dos cosas necesitaríamos estimar el efecto medio del tratamiento para los tratados, o el efecto medio del tratamiento para todos. Lo que se estima bajo los supuestos S.1. a S.4. es un efecto causal medio local del tratamiento, conocido como LATE, su acrónimo en inglés (*Local Average Treatment Effect* ). Un problema añadido es que no sabemos quiénes son los que se ven afectados por el instrumento en nuestra población. La decisión contrafactual respecto a *D* no la observamos. La ventaja es que, al menos, tenemos la estimación de un efecto causal libre del sesgo de selección que tanto ha confundido la estimación del efecto de políticas educativas y comparaciones entre centros educativos.

Hay una versión más fuerte del supuesto S.4. que es creíble en determinados contextos y que nos permitiría conocer el efecto medio del tratamiento sobre los tratados (ATT). Supongamos que, en lugar de la beca, lo que hacemos es decidir aleatoriamente qué estudiantes de un municipio van a ir a la única academia de idiomas del municipio y quiénes no. La asignación significa la gratuidad, por lo que en términos de dinero podría ser equivalente a la beca. Con este experimento se perseguiría conocer directamente el efecto de *D*, puesto que es *D* lo que se intenta asignar, y utilizaríamos el método descrito en la sección 4. El problema es que no podemos obligar a ningún estudiante que no quiera apuntarse a la academia, por lo que la decisión de ir o no a la academia no se implementa tal cual ha sido determinado por la aleatorización. Hay estudiantes que han sido asignados al grupo de los tratados pero que deciden no ir a la academia y están en realidad en el grupo de control, rompiendo la aleatorización y la comparabilidad. Pero lo que sí

somos capaces de implementar es que a ningún alumno asignado al grupo de control se le admite en la academia, aunque lo solicite. En esta situación la asignación, que será la $Z$, puede ser un buen instrumento de $D$ (haber ido realmente a la academia). Cumpliría S.1., S.2. y S.3., y, además, cumpliría una versión más fuerte de S.4. Con el supuesto de monotonicidad es posible que haya estudiantes para los que $Z$ no tenga ningún efecto sobre $D$ porque van a hacer $D = 1$ igualmente. Esa situación está excluida aquí, porque no se permite que los que no han sido seleccionados sean admitidos en la academia. En este caso, el estimador de variables instrumentales nos dará el efecto causal medio de ir a la academia para los que van (los tratados). Lo que es más importante, es que, incluso en casos en los que la implementación de un ensayo estricto no es posible, la experimentación aleatoria de algunas políticas permite generar, al menos, instrumentos válidos para estimar el efecto causal de las políticas para algunos grupos de la población.

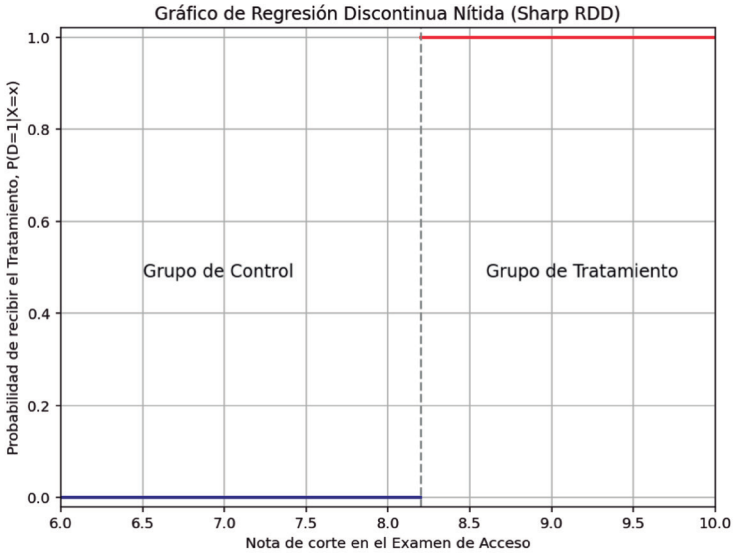
## 8. MÉTODO DE REGRESIÓN DISCONTINUA (RDD)

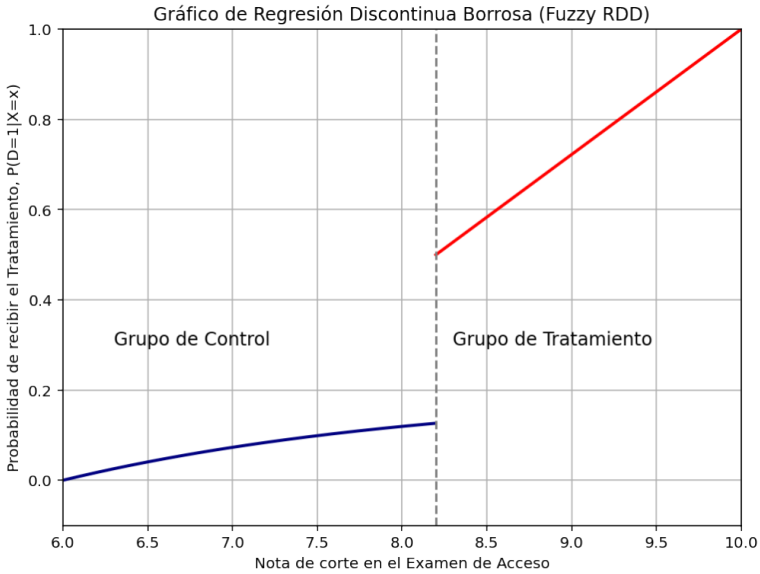
Pensemos en un Instituto de Bachillerato considerado de excelencia que recibe muchas solicitudes y admite a sus alumnos basándose en la nota que estos obtienen en un examen de acceso. Si queremos saber la calidad de este centro educativo, esto es, el efecto causal de estudiar aquí ($D$) sobre el aprendizaje de los alumnos, medido por el resultado de esto en los examenes de acceso a la Universidad ($Y$) que realizan al terminar el Bachillerato, no podemos simplemente comparar los resultados de los que han sido admitidos en ese instituto con el resultado de los que no. Como ya se ha explicado, habría un sesgo de selección en esa comparación porque esos alumnos han sido seleccionados precisamente por sus mejores resultados de partida. No sabríamos si las diferencias se deben a la calidad del instituto o a que los alumnos ya tenían más conocimientos antes de ir a ese centro. Sin embargo, este sistema de admisión, que hace no comparables a tratados y controles, genera una discontinuidad que este método va a explotar para detectar a dos grupos de tratados y controles que sí pueden ser comparables.

Por seguir con el mismo ejemplo, supongamos que el umbral, la nota de corte en el examen de acceso, que determina quienes son admitidos es 8,20 (en una escala de 0 a 10). Aquellos admitidos (tratados) que tienen un 9,4 son muy distintos de aquellos controles que no son admitidos porque han sacado un 6,1. Por eso no los podemos comparar. Pero aquellos que tienen entre un 8,20 y un 8,30 tienen un nivel de conocimientos previo (casi) igual a aquellos que tienen entre

un 8,10 y 8,19. Esa diferencia de una décima respecto al umbral que ha hecho que unos sean admitidos y otros no puede deberse mas al azar que ha hecho que ese día unos lo hicieran un épsilon peor o mejor que otros, dado un nivel de conocimientos igual. La diferencia de conocimientos no existe (o es tan pequeña que no es relevante), pero sí hay una gran diferencia en el valor de su *D*: unos son tratados y otros no. Por lo tanto, ya tenemos un grupo de tratados que es comparable con un grupo de control y cuyas diferencias en *Y* se las podemos atribuir al efecto causal de *D*, y no a la diferencia en conocimientos previos. Si hay un salto en *Y* al pasar de un lado al otro del umbral, esto se debe al tratamiento.

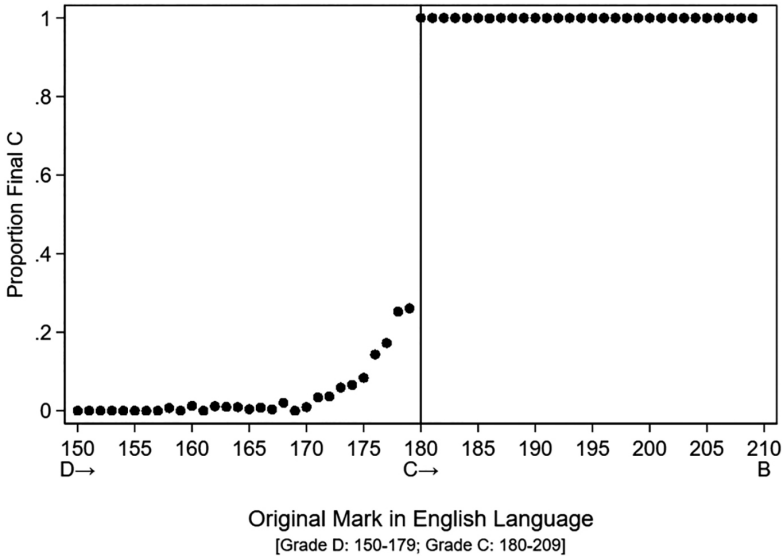
En el caso en el que todos los que están a un lado del umbral reciben el tratamiento y todos los que están al otro lado no lo reciben, se habla de Regresión Discontinua *nítida* (*sharp* RDD). En el caso en el que estar a un lado del umbral no determina por completo la asignación al tratamiento, sino que sólo sirve para cambiar la probabilidad de ser tratado, se habla de Regresión Discontinua *borrosa* (*fuzzy* RDD). Esto puede suceder, por ejemplo, porque la nota de acceso es solo uno de los criterios que se tiene en cuenta, y solo sirve para aumentar la probabilidad de ser admitido. Pero, algunos de los que tienen una nota de corte algo inferior son finalmente admitidos, y algunos de los que están por encima de la nota de corte, aún siendo admitidos, deciden no ir a ese centro. En los dos gráficos adjuntos se representan estas dos situaciones, el caso nítido y el caso borroso.





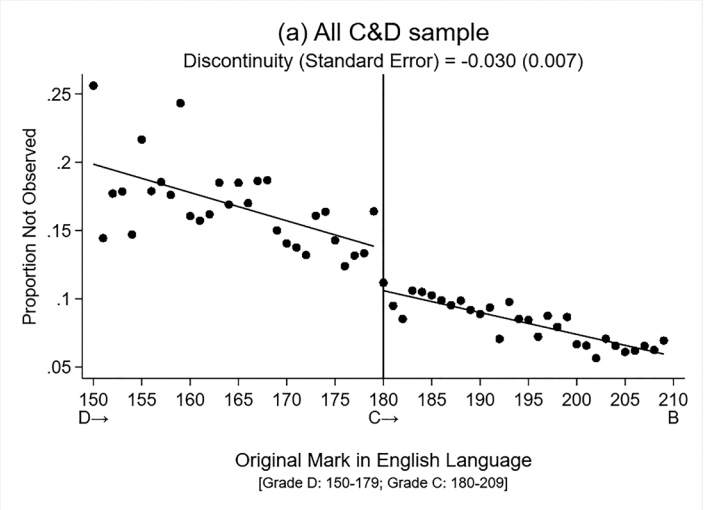

En el caso borroso el estimador es algo tan intuitivo como comparar el resultado medio de aquellos que están muy cerca del umbral a ambos lados del umbral. En el caso borroso el estimador se puede ver como un estimador de variables instrumentales en el que el instrumento es un indicador de estar por encima del umbral, en donde se va, no sólo a comparar el resultado en intervalos cercanos a ambos lados del umbral, sino a tener en cuenta que una parte tiene un valor de $D$ que no se corresponde estrictamente con estar a un lado del umbral. Para explicar esto gráficamente, voy a usar una aplicación real del método. En concreto el artículo de Machin, McNally y Ruiz-Valenzuela (2020) que estudia el efecto de aprobar (frente a no aprobarlo) en un conjunto de exámenes que tiene que haber realizado los estudiantes ingleses al completar la educación obligatoria a los 16 años de edad. El equivalente a completar al ESO en España. A efectos prácticos, aprobar requiere, entre otras cosas, sacar una nota C en lengua inglesa. Es el superar esta materia en este examen de lo que los autores van a analizar su efecto (variable $D$). Una de las variables de resultado ($Y$) que utilizan es si a los 18 años de edad siguen estudiando (sea en el nivel y tipo de estudios que sea) o no. Como variable que determina la asignación al tratamiento (superar el examen) utilizan la nota original en la prueba de inglés. En el primer gráfico puesto a continuación, que es la figura 4 en Machin, McNally y Ruiz-Valenzuela (2020), se presenta la proporción de alumnos que aprueban (obtienen una C), para cada puntuación obtenida en inglés. Como puede verse, la asignación es borrosa porque

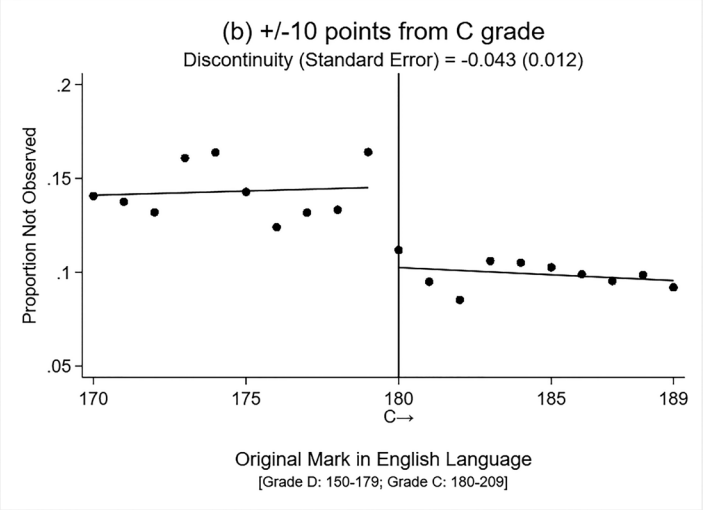
algunos de los que no llegan a 180 puntos (corte para obtener una C), consiguen finalmente superar el examen. Algunos de los que se quedan muy cerca, pueden solicitar la revisión de la nota original en algunos de los varios aspectos que la componen. Los detalles de todo esto pueden encontrarse en el artículo de Machin, McNally y Ruiz-Valenzuela (2020). Lo que es importante es que hay una discontinuidad en la proporción de aprobar. La tasa de aprobados no asciende de forma continua con la nota original. Salta de una tasa de alrededor el 30% para los que tienen una puntuación casi igual, pero inferior, a 180, a una tasa del 100% para aquellos que tienen 180 o más puntos en el examen.




Los dos gráficos puestos a continuación de este párrafo, que son parte de la figura 6 en Machin, McNally y Ruiz-Valenzuela (2020), presentan cómo evoluciona la media de *Y*, esto es, la proporcion de personas que a los 18 años no está estudiando para cada nivel de puntuación en inglés. Es muy interesante observar en este gráfico que hay una correlación negativa entre tener más nota y la proporción de los que a los 18 ya no estudian nada. Es por eso, que no podemos comparar a todos los que han aprobado con todos los que no han aprobado para conocer el efecto de aprobar. Esto alumnos son muy distintos en su nivel de conocimientos al finalizar la etapa obligatoria. El método de RD lo que compara es a aquellos que tienen prácticamente la misma puntuación, pero por muy poco, unos se quedan por debajo y otros por encima del aprobado. Si la proporción de *Y* evoluciona de la misma forma alrededor de ese valor, como lo hace alrededor de cualquier otro

valor, diremos que no hay efecto del hecho de aprobar. Pero como se ve en el gráfico, justo en ese valor, 180, se presenta un salto, una discontinuidad en la evolución de *Y*. La estimación del efecto vendrá dada por la comparación de los niveles de *Y* alrededor de 180. Según esa comparación, puede verse que hay un salto de un nivel de alrededor del 14% a un nivel de alrededor del 11%, a lo que habrá que descontar el que algunos por debajo de 180 finalmente aprueban. Por supuesto, la estimación concreta no se obtiene mirando un gráfico. Hay procedimientos numéricos y estimadores que hacen esto y dan los errores estándar de esa estimación. Estos se pueden ver en las tablas del artículo. Pero los gráficos sirven muy bien para entender cómo identifica RDD el efecto causal.

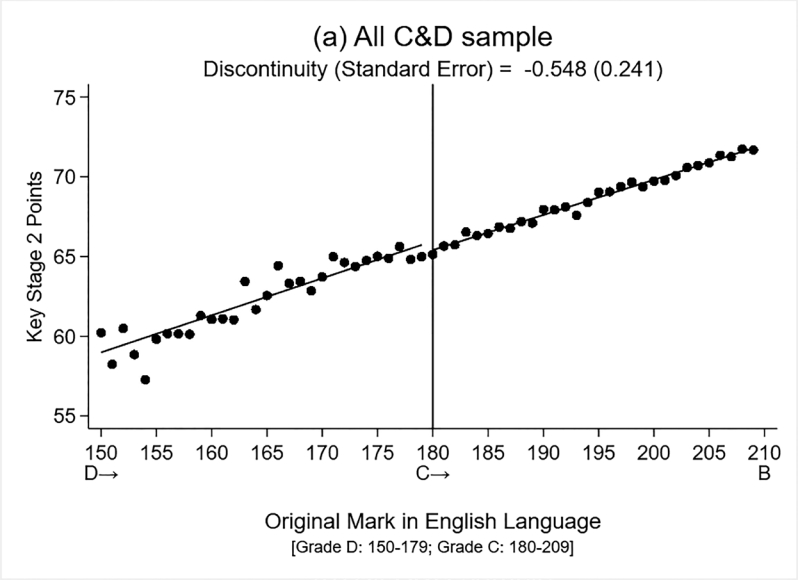


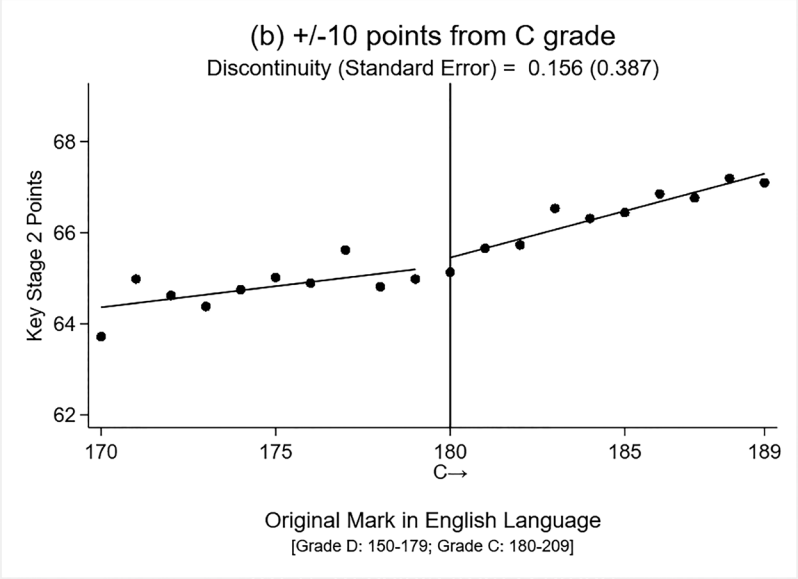



En cualquier caso, la principal complicación práctica vendrá porque normalmente no vamos a poder quedarnos en un intervalo muy pequeño alrededor del umbral por falta de suficientes observaciones. La comparación va a incluir un intervalo no tan pequeño y esto puede introducir ruido/sesgo en la estimación. Hay distintos métodos respecto a cómo hacer esto, incluyendo decidir la amplitud del intervalo y las propiedades de los mismos cuya descripción se escapa del objetivo de este capítulo. Más relevante para el propósito de este capítulo es comentar los supuestos necesarios para que este método nos de una estimación de un efecto causal y sus limitaciones.

Además de la condición de que exista esa discontinuidad, ese punto en una variable que haga que cambie de forma discontinua (que haya un salto en) la probabilidad de ser tratado, se necesita suponer que todo lo demás, en concreto los resultados potenciales, en ausencia del tratamiento cambian de forma continua alrededor de ese umbral. Esto se conoce como supuesto de continuidad local. Formas de buscar evidencia a favor o en contra de ese supuesto de continuidad son las siguientes:

1) Las características de los estudiantes o resultados de los mismos anteriores al tratamiento deben cambiar de forma continua alrededor del umbral. No sería sorprendente que cambiaran. Por ejemplo, no sería sorprendente que aquellos con un nivel inicial más alto, dispongan de más recursos en sus hogares. Lo que no tiene que suceder es que esto cambie de forma discontinua alrededor del umbral. El tratamiento no puede afectar a esta característica preexistente. En el siguiente gráfico, sacado de la figura 5 de Machin, McNally y Ruiz-Valenzuela (2020) puede verse como hay continuidad en la evolución de la relación entre la nota que obtienen a los 16 años (el "Original Mark In English Language" en el eje de abscisas del gráfico) y la nota que obtuvieron en el examen que realizaron cuando tenían 11 años de edad (el "key stage 2 Points" del eje de ordenadas del gráfico). Hay una correlación positiva, de forma que los que tienen más nota en uno tienden a tener más nota en otro. Pero en el paso de estar algo por debajo de 180 a estar algo por encima de 180 no hay ningún salto. Sigue una evolución continua. Así debe ser, puesto que aprobar o suspender a los 16 no puede tener ningún efecto causal en el resultado de un examen que se había hecho a los 11 años de edad.




2) Hacer pruebas de falsificación como las descritas en la sección 6.3.

3) Mirar la distribución de alumnos a cada lado del umbral. Si encontramos que se concentra una masa de ellos justo por encima y hay un vacío justo por debajo, esto sería evidencia de manipulación endógena, o posible sesgo de selección. Un ejemplo en el que frecuentemente se observa esto es en las notas

académicas cuando hay un umbral que marca una situación dicotómica como promocionar o repetir curso. Suele haber una proporción mucho más alta de aprobados justo por encima del umbral, que suspensos justo por debajo del umbral que separa el aprobado/suspenso. Y unos y otros no suelen ser comparables porque los maestros que han puesto esas calificaciones finales han tenido en cuenta más características e información para decidir qué alumnos de entre aquellos con una nota muy similar y cercana al aprobado finalmente quedaban por encima del umbral. Esto impedirá utilizar ese tipo de calificaciones en el método de regresión discontinua para evaluar el impacto de repetir curso sobre el rendimiento académico futuro. Repetidores y no repetidores, aunque tengan una nota ambos cercana al umbral, serán distintos en otras de las características tenidas en cuenta por la junta de evaluación para decidir quién se queda por encima y por debajo del umbral.

Otra amenaza a la validez sería la existencia de contaminación sobre el grupo de control, de forma que el tratamiento tiene efectos también sobre los no tratados, igual que se señaló en el caso de los experimentos en la sección 4.2. En el ejemplo en el que se quiere estudiar el impacto causal de una política que determina la repetición de curso, sobre el rendimiento académico de los alumnos, puede suceder que la amenaza de tener que repetir curso suponga un elemento de incentivo fuerte para los alumnos que no repiten, haciéndoles estudiar y aprender más. Este efecto contaría como un efecto negativo de la política al comparar repetidores con no repetidores, cuando en realidad es un efecto positivo que desaparece si se elimina la política y se pasa a la promoción automática de todos.

**Efecto causal que se estima:** el efecto causal que se estima es válido para aquellos que están alrededor del umbral, pero no tiene por qué ser válido para aquellos que están lejos del umbral. Este es también un efecto local. En la sección 11.5 se encuentran las expresiones matemáticas de los supuestos y el parámetro causal que identifica el RDD.

## 9. MODELOS ESTRUCTURALES: INVESTIGANDO LOS MECANISMOS QUE EXPLICAN LOS EFECTOS TOTALES DE UNA POLÍTICA

Un modelo estructural es un modelo económico que describe cómo se comportan los agentes objeto de estudio (estudiantes, familias, hogares, etc.), cómo toman sus decisiones, y qué restricciones de tiempo, presupuestarias y/o tecnoló-

gicas tienen. Por ejemplo, los modelos económicos que se presentan en el siguiente capítulo de este libro describen cómo deciden y se comportan los estudiantes y las familias respecto a la educación. Extendiéndolos para incorporar más variables, incluyendo elementos inobservables y dinámicos, tendríamos un modelo estructural que puede describir bien un conjunto de observaciones reales. Se usarán esos datos reales para estimar el valor de los parámetros que caracterizan ese modelo y que desconocemos.

Los métodos anteriores nos permiten conocer el efecto total o de forma reducida sobre $Y$ de cambiar $D$ de un valor $D_0$ a un valor $D_1$ sin necesidad de especificar cómo se determina la $Y$, ni qué otras variables afectan a $Y$, ni la forma de esos efectos. Pero, tal y como se ha explicado en la sección 2.2.1 y en la sección 4.2, eso no nos permite aprender nada sobre los mecanismos que causan ese efecto, ni separar el efecto directo de la $D$ del efecto que se produce a través de afectar otras variables que también determinan la $Y$, y puede presentar problemas de validez externa y de equilibrio general. Frente a eso, estimar un modelo estructural para conocer el efecto de $D$ presenta las siguientes ventajas:

1) Los modelos de comportamiento permiten estudiar mecanismos a través de los cuales la política tiene sus efectos sobre la variable de resultado, y, por lo tanto, separar los efectos directos de los efectos a través de otras variables. Tal y como se describe en el siguiente capítulo de este libro, los alumnos deciden sobre su esfuerzo y las familias deciden sobre el tiempo y recursos que dedican a la educación de sus hijos y ambas pueden responder a cambios en los insumos por parte del centro, siendo este un mecanismo adicional por el cual ese cambio en los insumos va a dar un efecto total sobre el progreso académico de los alumnos.

2) Validez externa. Al capturar los mecanismos y los efectos e interacciones con el resto de determinantes de la decisión, los efectos de una política se pueden conocer más allá de los individuos sobre los que se ha aplicado la política hasta ahora. En el modelo estructural se podrán tener en cuenta cómo las diferentes características y circunstancias del distinto grupo de individuos, sobre el que ahora se quiere conocer los efectos de una política, cambian el efecto de $D$ sobre la $Y$.

3) Permiten la evaluación de efectos de largo y corto plazo de una política educativa, puesto que estos modelos tienen en cuenta el carácter dinámico y basado en expectativas futuras de las decisiones de los individuos.

4) Los modelos estructurales permiten la simulación de experimentos de política contrafácticos. Para conocer el efecto causal de una política usando el método descrito en la sección 4 necesitamos ensayar la política. Y si tras el ensayo queremos saber qué pasaría si cambiamos alguno de los aspectos de esa política, por pequeño que sea el cambio, necesitamos volver a hacer otro ensayo. Lo mismo sucede si tras el primer ensayo los resultados no son los esperados y queremos saber qué pasa si hubiéramos ensayado otra política. Con ese método no se puede saber si no se vuelve a hacer otro ensayo. Sin embargo, una vez que tenemos un modelo estructural que caracteriza bien las decisiones, condicionantes y restricciones de los individuos, este nos permitirá simular y conocer, sin necesidad de implementarlo ni ensayarlo, el efecto que tendrá una política distinta a la implementada hasta ese momento.

Entre las desventajas, la más importante es que no resulta fácil validar el modelo escogido. La estimación requiere muchos datos y no es sencilla. Pero además, hay supuestos funcionales y paramétricos que elige el investigador y sería necesario validar. Para mostrar que el modelo es plausible se debe:

- Validar (ajusta bien las decisiones que observamos) dentro de muestra.
- Validar fuera de la muestra.
- Utilizar cambios de régimen y ver que es capaz de predecir bien sus efectos.
- Comparar con resultados de experimentos: la estimación del impacto total de una política ensayada debe de ser igual al impacto total que da el ensayo con el método descrito en la sección 4.

## 10. EFECTOS HETEROGÉNEOS

Implícito en todo lo escrito en este capítulo está el hecho de que hay efectos heterogéneos, esto es, el efecto es distinto para cada estudiante, y el objetivo ha sido estimar el efecto medio para toda la población de interés. Esto es lo que hace que en variables instrumentales estimemos el LATE en lugar del efecto causal para todos; o que en Regresión Discontinua tengamos el efecto sólo para aquellos que están cerca del punto de corte, pero no para los demás. Si el efecto fuera homogéneo, estimarlo bien para un subgrupo de la población significaría estimarlo bien para todos, puesto que todos tendrían el mismo efecto.

Normalmente no estamos interesados en conocer el efecto para cada estudiante (o en términos genéricos, para cada unidad de observación, sea esta estudian-

tes, trabajadores, hogares, etc.). Queremos conocer una medida informativa de ese efecto para nuestra población de interés. El efecto medio es la medida informativa más usada y para eso todos los métodos anteriores son interesantes y, en principio, suficientes. No obstante, si hay efectos heterogéneos, suficientemente distintos, entre individuos, tenemos dos nuevos problemas que enfrentar:

1) Esto es una de las principales razones para cuestionar la validez externa del efecto causal estimado mediante un experimento. Conocemos la media para ese grupo de individuos. Pero si los efectos son muy heterogéneos, otro grupo distinto de individuos puede tener un efecto medio muy distinto. Puede que incluso para algunos haya un efecto positivo y significativo y para otros un efecto cero.

2) También querremos poder usar varias medidas que caractericen esos efectos heterogéneos para conocerlo y poder, por ejemplo, saber sobre qué grupos la política educativa en cuestion tiene un efecto más positivo.

Una forma de estimar, al menos en parte, efectos heterogéneos es estimar efectos medios para varios subgrupos de individuos, definidos estos subgrupos por medio de variables observables en cuya dimensión pensamos que los efectos pueden ser distintos. Por ejemplo, permitir distintos efectos de reducir el tamaño de clase para alumnos con padres con distinto nivel educativo. Esto, en principio, es muy fácil de hacer con cualquiera de los métodos de todas las secciones anteriores. Por ejemplo, con datos experimentales, se trata de hacer la comparación entre tratados y controles subgrupo a subgrupo. El problema suele estar en la falta de un número suficiente de datos. Hacer esto requiere partir la muestra en varias submuestras. Un ensayo con 500 alumnos tratados y 500 controles puede ser suficientemente grande para estimar el efecto medio. Pero si tenemos que permitir efectos heterogéneos por nivel educativo de la madre, del padre y por el sexo de los alumnos, podemos encontrarnos que tenemos que hacer más de 15 subgrupos, usando cada estimación menos de 40 tratados y controles.

Otra forma de hacerlo es utilizar técnicas econométricas destinadas a estimar la distribución, o partes de la misma como varios cuartiles, de los efectos heterogéneos. Esta forma tiene mucho sentido en un contexto de estimación estructural, o cuando estemos usando más variables observables en los otros métodos de estimación.

## 11. EXPRESIONES FORMALES DE ALGUNOS DE LOS PARÁMETROS Y ESTIMADORES

### 11.1. Resultados potenciales, efecto causal, sesgo de selección y ensayos

$D_i$ es el valor de la variable cuyo efecto queremos conocer para el individuo *i*; *i* puede representar estudiantes, hogares, trabajadores, escuelas, empresas, etc. En cada caso representará aquellas unidades sobre las que se produce el efecto de *D* sobre *Y*. $Y_i$ representa el valor de la variable de resultado para el individuo *i*. Nosotros estamos interesados en conocer el efecto sobre *Y* de aumentar el valor de *D* de un valor $D_0$ a un valor $D_1$. Para todo individuo existen a priori dos resultados o valores de la $Y_i$ potenciales: 1) el que obtendrá si $D_i = D_0$ que llamamos $Y(D_0)_i$; y 2) el que obtendrá si $D_i = D_1$ que llamamos $Y(D_1)_i$. El efecto causal (total) para este individuo es $Y(D_1)_i - Y(D_0)_i$. Simplificando a la situación binaria, tratados $D_1 = 1$, controles $D_0 = 0$, tenemos,

Efecto medio para toda la población: $\mathbb{E}\,(Y(1)_i - Y(0)_i) \equiv \tau_{ATE}$

Efecto medio para los tratados: $\mathbb{E}\,(Y(1)_i - Y(0)_i|D_i = 1) = \mathbb{E}\,(Y(1)_i|D_i = 1) -$

$$-\mathbb{E}\,(Y(0)_i|D_i = 1) \equiv \tau_{ATT}$$

Observamos $\mathbb{E}\,(Y_i|D_i = 1) - \mathbb{E}(Y_i|D_i = 0) = \mathbb{E}\,(Y(1)_i|D_i = 1) - \mathbb{E}\,(Y(0)_i|D_i = 0)$.

Esto será igual a $\tau_{ATT}$ si

$$\mathbb{E}\,(Y(0)_i|D_i = 0) = \mathbb{E}\,(Y(0)_i|D_i = 1) \tag{3}$$

La diferencia entre $\mathbb{E}\,(Y(0)_i|D_i = 0)$ y $\mathbb{E}\,(Y(0)_i|D_i = 1)$ es lo que nos da el sesgo de selección. La condición (3) se va a dar cuando tengamos un experimento aleatorio bien implementado. Si se da esta condición, la estimación consistirá en obtener el equivalente muestral de $\mathbb{E}\,(Y_i|D_i = 1) - \mathbb{E}\,(Y_i|D_i = 0)$.

### 11.2. Selección sólo en observables

La regresión que añade controles observables

$$Y = \beta_0 + \beta_1 X + \tau D + \varepsilon \tag{4}$$

me va a dar el efecto causal si $\mathbb{E}\,(\varepsilon|X, D) = 0$. Esto es así porque

$$\mathbb{E}\,(Y|X, D = 1) - \mathbb{E}\,(Y|X, D = 0) = \mathbb{E}\,(\beta_0 + \beta_1 X + \tau + \varepsilon|X, D = 1) -$$
$$- \mathbb{E}\,(\beta_0 + \beta_1 X + \varepsilon|X, D = 0) \tag{5}$$

$$= \beta_0 + \beta_1 X + \tau + \mathbb{E}\,(\varepsilon|X, D = 1) - (\beta_0 + \beta_1 X) - \mathbb{E}\,(\varepsilon|X, D = 0) \tag{6}$$

$$= \tau + \mathbb{E}\,(\varepsilon|X, D = 1) - \mathbb{E}\,(\varepsilon|X, D = 0) \tag{7}$$

Siendo $\tau$ el efecto causal medio:

$$\mathbb{E}\,(Y(1)_i - Y(0)_i) = \mathbb{E}\,(\beta_0 + \beta_1 X + \tau + \varepsilon - (\beta_0 + \beta_1 X + \varepsilon)) = \tau$$

Efecto medio del tratamiento sobre los tratados, *ATT* estimado con emparejamiento en base a $P\,(X)$:

$$\tau_{ATT}^{PSM} = \mathbb{E}_{P\,(X)|D=1}\,\{\mathbb{E}\,(Y\,|P(X), D = 1) - \mathbb{E}\,(Y\,|P(X), D = 0)\} \tag{8}$$

El supuesto de independencia condicional, o de selección solo en observables, se puede expresar de forma general como

$$\{Y\,(0), Y\,(1)\} \perp D|\,X, \forall X \tag{9}$$

## 11.3. DiD básico y supuesto de identificación

Supuesto de identificación:

$$\mathbb{E}\,(Y(0)|D = 0, \textit{después}) - \mathbb{E}\,(Y(0)|D = 0, \textit{antes}) = \mathbb{E}\,(Y(0)|D = 1, \textit{después}) - \mathbb{E}\,(Y(0)|D = 1, \textit{antes}) \tag{10}$$

DiD:

$$\tau_{DiD} = [\mathbb{E}\,(Y\,|D = 1, \textit{después}) - \mathbb{E}\,(Y\,|D = 1, \textit{antes})] - [\mathbb{E}\,(Y\,|D = 0, \textit{después}) - \mathbb{E}\,(Y\,|D = 0, \textit{antes})] \tag{11}$$

$$= [\mathbb{E}\,(Y(1)|D = 1, \textit{después}) - \mathbb{E}\,(Y(0)|D = 1, \textit{antes})] - [\mathbb{E}\,(Y(0)|D = 0, \textit{después}) - \mathbb{E}\,(Y(0)|D = 0, \textit{antes})]$$

Bajo el supuesto de identificación $\tau_{DiD}$ es igual a

$$= [\mathbb{E}\,(Y\,(1)|D = 1, \textit{después}) - \mathbb{E}\,(Y(0)|D = 1, \textit{antes})] - [\mathbb{E}\,(Y\,(0)|D = 1, \textit{después}) - \mathbb{E}\,(Y(0)|D = 1, \textit{antes})] \tag{12}$$

$$= \mathbb{E}\,(Y\,(1)|D = 1, \textit{después}) - \mathbb{E}\,(Y(0)|D = 1, \textit{después}) = ATT$$

## 11.4. Variables instrumentales: estimador de Wald, Supuestos y LATE

En el caso más sencillo en el que $D$ es binario, y $Z$ también es binario, el estimador de VI se conoce como estimador de Wald y es el equivalente muestral de la siguiente ratio:

$$\tau_{VI} = \frac{\mathbb{E}\,(Y_i|Z_i = 1) - \mathbb{E}\,(Y_i|Z_i = 0)}{\mathbb{E}\,(D_i|Z_i = 1) - \mathbb{E}\,(D_i|Z_i = 0)} \tag{13}$$

Los 4 supuestos enunciados en la sección 7 son:

1) S.1. Independencia del instrumento

$$\{Y\,(D(1), Z = 1)i\,,\, Y\,(D(0)\,, Z = 0)i\,,\, D(1)i\,,\, D(0)i\} \perp Z_i$$

2) S.2. Restricción de exclusión

$$Y\,(d, Z = 1)i = Y\,(d, Z = 0)_i$$

para $d = 0,1$.

3) S.3. Condición de rango o de relevancia del instrumento

$$\mathbb{E}\,(D|Z = 1) \neq \mathbb{E}\,(D|Z = 0)$$

4) S.4. Supuesto de monotonicidad

$$D(1)_i \geq D(0)_i,\ \forall\ i$$

o bien

$$D(1)_i \leq D(0)_i,\ \forall\ i$$

Bajo esos 4 supuestos

$\tau_{VI}$ en (13) es igual a $\mathbb{E}\,(Y(1)i - Y(0)i\,|D\,(1)i > D\,(0)_i) \equiv \tau_{LATE}$

## 11.5. Regresión Discontinua

Sea $Z$ la variable en base a la cual se decide quien es tratado y $z$ el umbral a partir del cual cambia (pongamos, aumenta) la probabilidad de ser tratado. Condición de identificación o de discontinuidad en la probabilidad de ser tratado:

$$\lim_{\Delta \to 0} Pr\,(D = 1\,|\,Z = \bar{z} + \Delta) \neq \lim_{\Delta \to 0} Pr\,(D = 1\,|\,Z = \bar{z} - \Delta) \tag{14}$$

Supuesto de continuidad local necesario para identificar un efecto causal:

$$\lim_{\Delta \to 0} \mathbb{E}\,(Y(j) \mid Z = \bar{z} + \Delta) = \lim_{\Delta \to 0} \mathbb{E}\,(Y(j) \mid Z = \bar{z} - \Delta) \text{ para } j = 0,\ 1 \qquad (15)$$

Definición, en términos poblacionales, del parámetro que da lugar al estimador de Regresión Discontinua (caso borroso, genérico):

$$\tau_{RDB} = \frac{\lim_{\Delta \to 0} \mathbb{E}\,(Y \mid i\, Z_i = \bar{z} + \Delta) - \lim_{\Delta \to 0} \mathbb{E}\,(Y_i \mid Z_i = \bar{z} - \Delta)}{\lim_{\Delta \to 0} Pr(D_i = 1 \mid Z_i = \bar{z} - \Delta) - \lim_{\Delta \to 0} Pr(D_i = 1 \mid Z_i = \bar{z} - \Delta)} \qquad (16)$$

Si la discontinuidad fuera nítida, entonces $\lim_{\Delta \to 0} Pr\,(D_i = 1 \mid Z_i = z + \Delta) = 1$, y $\lim_{\Delta \to 0} Pr\,(D_i = 1 \mid Z_i = z - \Delta) = 0$, y el denominador valdría 1.

Se puede demostrar que

$$\tau_{RDB} = \mathbb{E}\,(Y(1)_i - Y(0)_i \mid Z_i = \bar{z}) \qquad (17)$$

lo que es el efecto del tratamiento para los individuos que están en el umbral $\bar{z}$.

## BIBLIOGRAFÍA

### Libros que cubren la mayoría de los métodos de evaluación de impacto descritos


Angrist, J. y J. Pischke (2009). *Mostly Harmless Econometrics*. Princeton University Press.

Bernal, R. y X. Peña (2011). *Guía Práctica para la Evaluación de Impacto.* Universidad de los Andes. ISBN: 978-958-695-599-7.

Cunningham, S. (2021). *Causal Inference. The Mixtape*. Yale University Press.


### Libros de texto de Econometría

Contienen materiales necesarios para entender bien las técnicas de estimación que se usan. Además, en algunos capítulos tratan de forma específica métodos descritos aquí:


Cameron A. y P. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Greene, W. (2003). *Econometric Analysis*, 5th edition. Prentice Hall.

Stock. J. H. y M. K. Watson (2012). *Introducción a la Econometría*, 3ª edición, versión en español. Prentice Hall.

Wooldridge, J. W. (2010). *Econometric Analysis of Cross Section and Panel Data,* 2nd edition. The MIT Press.

**Artículo sobre experimentos**

Artículo sobre el uso de experimentos y las limitaciones de los mismos:

Deaton, A. (2009). Instruments of Development: Randomisation in the Tropics, and the Search for the Elusive Keys to Economic Development, The Keynes Lecture in Economics, *NBER Working Paper 4690*.

**Artículos sobre emparejamiento**

Caliendo, M. y S. Kopeinig (2005). Some Practical Guidance for the Implementation of Propensity Score Matching, *IZA Discussion Paper No. 1588*.

Dehejia, R. (2005). Practical Porpensity Score Matching: A Reply to Smith and Todd, *Journal of Econometrics*, *125*:355-364.

Heckman, J., H. Ichimura y P. Todd (1997). Matching as an Econometric Evaluation Estimator, *Review of Economic Studies*, *64*:605-654.

Artículo en el que se demuestra que si puedo identificar el efecto causal emparejando en observables, también lo puedo hacer emparejando en función de la probabilidad de ser tratado:

Rosenbaum, P. y D. Rubin (1983). The central role of the Propensity Socore in Observational Studies for Causal Effects, *Biometrika*, *70*:41-50.

Smith, J. y P. Todd (2005). Does Matching Overcome LaLonde's Critique of Non-Experimental Estimators?, *Journal of Econometrics*, *125:*305-353.

**Uso de observables para intentar acotar el sesgo de selección**

Altonji, J., T. Elder y C. Taber (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools, *Journal of Political Economy*, *113*(1):151-184.

**Artículos sobre DiD**

Avances recientes sobre el supuesto de tendencias paralelas y las restricciones de forma funcional cuando se controla por observables:

Rambachan, A. y J. Roth (2023). A More Credible Approach to Parallel Trends, *Review of Economic Studies, 90*:2555-2591.

Roth, J. y P. Sant'Anna (2023). When is parallel trends sensitive to functional form?, *Econometrica, 91*(2):737-747.

Sun, L. y S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects, *Journal of Econometrics*, *225*(2):175-199.

Artículos recientes sobre contrastar tendencias paralelas:

Jaeger, D., T. Joyce y R. Kaestner (2020). A Cautionary Tale of Evaluating Identifying Assumptions: Did Reality TV Really Cause a Decline in Teenage Childbearing?, *Journal of Business & Economic Statistics, 38*:317-326.

Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends, *American Economic Review Insights*, *4*(3):305-322.

Artículos recientes sobre estimación cuando se dispone de varios periodos postratamiento:

Callaway, B. y P. Sant'Anna (2021). Difference-in-Differences with multiple time periods, *Journal of Econometrics, 225*:200-230.

Chaisemartin, C. y X. D'Haultfoeuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects, *American Economic Review*. *110*(9):2964-2996.

Mas allá del DiD:

Arkhangelsky, D., S. Athey, D. Hirshberg, G. Imbens y S. Wager (2021). Synthetic Difference-in-Differences, *American Economic Review, 111*(12):4088-4118.

Artículo resumen de los últimos y relevantes avances en DiD:

Roth, J., P. Sant'Anna, A. Bilinski y J. Poe (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature, *Journal of Econometrics*, *235*(2):2218-2244.

**Artículo sobre variables instrumentales**

Imbens, G. y J. Angrist (1994). Identification and Estimation of Local Average Treatment Effects, *Econometrica, 62*:467-476.

Artículo que va más allá del LATE y relaciona variables instrumentales con la literatura estructural:

Heckman, J. y E. Vytlacil (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation, *Econometrica, 73*(3):669-738.

**Artículos sobre Regresión Discontinua**

Hahn, J., P. Todd y W. van der Klaauw (2001). Identification and Estimation of Treatment Effects with a Regression-discontinuit Design, *Econometrica, 69*:201-209.

Imbens, G. y Lemieux (2008). Regression Discontinuity Designs: A Guide to Practice, *Journal of Econometrics*, *142*(2):615-635.

Lee, D. y T. Lemieux (2010). Regression Discontinuity Designs in Economics, *Journal of Economic Literature, 48*(2):281-355.

Ejemplo de aplicación que se ha usado en esta sección:

Machin, S., S. McNally y J. Ruiz-Valenzuela (2020). Entry through the narrow door: The costs of just failing high stakes exams, *Journal of Public Economics*, *90,* 104224. https://doi.org/10.1016/j.jpubeco.2020.104224

**Artículos sobre estimación estructural e inferencia causal**

Este artículo es un buen ejemplo de cómo usar estimación estructural para conocer los efectos causales de una política y similares políticas alternativas:

Todd, P. y K. Wolpin (2006). Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment toValidate a Dynamic Behavioral Model of Child Schooling and Fertility, *American Economic Review*, *96*(5), 1384-1417.

Artículos más generales sobre inferencia causal:

Heckman, J. (2000). Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective, *Quarterly Journal of Economics*, *115*(1):45-97.

Rosenzweig, M. y K. Wolpin (2000). Natural Experiments in Economics, *Journal of Economic Literature, 38*(4):827-874.

# Capítulo 3a
# EL IMPACTO DEL TAMAÑO DE LA CLASE EN EL RENDIMIENTO ACADÉMICO

Ismael Sanz
Departamento de Economía Aplicada I e Historia e Instituciones Económicas
Universidad Rey Juan Carlos
Department of Social Policy
London School of Economics
ORCID: https://orcid.org/0000-0003-1286-4124

**Resumen**

La relación entre el tamaño de las clases y el rendimiento académico representa un tema crucial en la discusión sobre estrategias educativas y asignación óptima de recursos. Este análisis revisa la literatura especializada para explorar las implicaciones y eficacia de las reducciones de las ratios alumnos por clase. Las clases más pequeñas son preferidas por docentes y padres por permitir una atención más personalizada y mejorar la gestión de las posibles interrupciones en clase. Sin embargo, las investigaciones demuestran que, si bien la reducción del tamaño de la clase tiene un impacto positivo en el aprendizaje de los alumnos, este efecto es limitado y conlleva un significativo coste presupuestario. Reducir la ratio alumnos por aula en los primeros niveles educativos de Infantil y Primaria y en áreas desfavorecidas, o estrategias como tutorías en pequeños grupos, se presentan como opciones con una relación coste-beneficio superior a la reducción general del tamaño de la clase. Futuras investigaciones podrían centrarse en la interacción entre el tamaño de la clase y las prácticas docentes, la incorporación de tecnología y las innovaciones pedagógicas para desarrollar soluciones más integrales y eficaces en la mejora del rendimiento académico.

## 1. INTRODUCCIÓN

La asociación entre el número de estudiantes por aula y el rendimiento académico es un aspecto central en el debate sobre la calidad y la asignación de

los recursos educativos. En economía de la educación se conoce a los artículos que abordan esta relación como la literatura del tamaño de clase. El número de alumnos por aula es una de las métricas más comunes para evaluar los recursos en el ámbito educativo. Su medición es sencilla y menos subjetiva que la mejora de la calidad. Entre las ventajas de disminuir el tamaño de la clase se encuentra su facilidad de implementación y, sobre todo, su popularidad con un gran apoyo entre los agentes implicados en el sistema educativo (padres y docentes). La variación del tamaño de las clases es observable para los padres como una medida de la inversión en educación. Los profesores prefieren también clases con menos alumnos. Las clases más pequeñas a menudo se ven como beneficiosas, ya que permiten a los docentes dedicar más tiempo, atención y apoyo a los alumnos y centrarse en las necesidades de cada uno de ellos. En cambio, en clases más grandes, los profesores necesitan más tiempo para lidiar con las interrupciones, al mismo tiempo que algunos estudiantes podrían desconectarse de la instrucción. Lazear (2001) elabora un modelo teórico en el que los tamaños de clase óptimos son más elevados para los estudiantes con buen comportamiento. Los alumnos con problemas de conducta interrumpen más la enseñanza afectando negativamente el aprendizaje (en el recuadro 4 de este capítulo se presenta con detalle este modelo de producción de educación). El modelo de Lazear (2001) es no lineal, por cuanto que la probabilidad de interrupción aumenta a un ritmo más rápido para las clases pequeñas que para las grandes, en las que las perturbaciones ya son más comunes.

La mayor desventaja de reducir el tamaño de las clases es que es una medida tremendamente cara. Crear nuevas clases implica contratar nuevos profesores, construcción de aulas adicionales, la compra de nuevos materiales para el aula, personal administrativo adicional y una serie de otros costes asociados (Handel y Hanushek, 2023). El tamaño de la clase tiene un impacto considerable en el nivel de gasto corriente en educación a través de los salarios de los profesores. Un aumento en los costes laborales de la educación que pone en jaque el ya limitado presupuesto de educación, donde el 72% en España y 58% en OCDE del gasto en centros públicos no universitarios se destina al pago de salarios de los docentes (OCDE, Education at a Glance, 2022, Table C6.3). De hecho, España es uno de los países desarrollados que más porcentaje del gasto total como del corriente dedica a remuneraciones de los docentes en centros públicos, una circunstancia que no ocurre en el caso de la enseñanza privada que está más cerca del promedio de la OCDE.

La ratio estudiante por clase es un input, no un output, de modo que no garantiza un mayor aprendizaje o mejor calidad de la educación. Para examinar si disminuir el tamaño de la clase es una política pública eficaz, hay que determinar su coste, evaluar como beneficiaría a los estudiantes (a corto y largo plazo), y comparar dicho coste-beneficio con el de otras medidas educativas que, por el mismo o menor presupuesto, podrían tener efectos en el aprendizaje igual o más positivos. Las ratios más bajas de estudiantes por clase tienen que ser sopesadas en relación con otras intervenciones como aumentar el salario de los docentes, invertir más en su desarrollo profesional, incrementar la tecnología en la enseñanza o un uso más generalizado de asistentes de profesores, por poner algunos ejemplos. Cualquier efecto de las reducciones del tamaño de las clases necesita ser comparado en términos de eficiencia con otras intervenciones educativas alternativas (Green e Iversen, 2022).

El tamaño de la clase se calcula dividiendo el número de estudiantes entre el número de clases y mide el promedio de alumnos que se agrupan en un aula. La ratio alumnos por profesor se computa, por el contrario, en términos de alumnos equivalentes a tiempo completo con respecto a docentes equivalentes a tiempo completo reflejando el nivel de recursos de enseñanza disponibles en un país en relación con su población estudiantil (véase recuadro 1 sobre las diferencias entre ratio alumnos por clase y alumnos por profesor). España tiene menos alumnos por profesor que la OCDE en Primaria y ESO y, sin embargo, tiene más alumnos por clase. Los últimos datos disponibles de 2019 en la tabla 1, muestran que el tamaño de la clase en España es de 22 alumnos por clase en Primaria (columna 5), por los 21 en la OCDE y los 19 de la UE-22. También en la ESO España tiene más alumnos por clase (columna 10): 25 por los 23 de la OCDE y los 21 de la UE-22. La ratio promedio de alumnos por profesor para ese mismo año de 2019 en España en Primaria (14) y ESO (12), se encuentra justo entre las medias de la UE (13 y 11) y de la OCDE (15 y 13). Con el mismo número o más de docentes en España que en la OCDE, el número de alumnos por clase es más elevado (véase recuadro1).

**Tabla 1. Tamaño promedio de la clase, por tipo de institución y nivel de educación (2013 y 2019).**

| | Primaria 2019 | | | | | Secundaria obligatoria 2019 | | | | | 2013 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pública | Centros privados | | | Total Primaria | Pública | Centros privados | | | Total Secundaria obligatoria | Total Primaria | Total Secundaria obligatoria |
| | | Total privada | Concertada | Privada | | | Total privada | Concertada | Privada | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (13) | (16) |
| Francia | 22 | 25 | 25 | – | 23 | 25 | 26 | 27 | 12 | 25 | 23 | 25 |
| Alemania | 21 | 21 | – | – | 21 | 24 | 23 | – | – | 24 | 21 | 24 |
| Italia | 19 | 19 | – | 19 | 19 | 21 | 21 | – | 21 | 21 | 19 | 22 |
| España | 21 | 24 | 25 | 20 | 22 | 25 | 27 | 27 | 21 | 25 | 22 | 25 |
| Reino Unido | 27 | 24 | 28 | 12 | 26 | 25 | 23 | 25 | 12 | 23 | 25 | 19 |
| Estados Unidos | 21 | 16 | – | 16 | 20 | 26 | 18 | – | 18 | 25 | 21 | 27 |
| Promedio OCDE | 21 | 20 | 21 | 18 | 21 | 23 | 21 | 22 | 19 | 23 | 21 | 23 |
| Promedio UE-22 | 20 | 18 | 19 | 15 | 19 | 21 | 20 | 20 | 17 | 21 | 19 | 21 |

Fuente: OCDE/UIS/Eurostat (2021, página 326) https://www.oecd-ilibrary.org/education/education-at-a-glance-2021_b35a14e5-en
Más información y el Anexo 3 para las notas https://www.oecd.org/education/education-at-a-glance/EAG2021_Annex3_ChapterD.pdf

**Recuadro 1. La relación entre el tamaño de la clase y la ratio alumnos por profesor. OCDE Education at a Glance (2012, 2020, 2021)**

El tamaño de clase mide el número de estudiantes que siguen el mismo plan de estudio, considerado como el mayor número de cursos en común y sin tener en cuenta los desdobles (OCDE, 2020). El cálculo se lleva a cabo como el cociente entre el número total de estudiantes y el número de clases existentes. En cambio, la ratio de alumnos por profesor se computa dividiendo el número de estudiantes equivalentes a tiempo completo entre el número de docentes equivalentes a tiempo completo en cada uno de los niveles educativos. Los dos indicadores están relacionados, pero también miden características diferentes del sistema educativo. La ratio alumnos por profesor proporciona información sobre los recursos humanos disponibles en un país en relación con su población de estudiantes, mientras que el tamaño de clase refleja cómo se usan esos recursos midiendo el promedio de estudiantes agrupados por cada aula. Es decir, que el tamaño de la clase se ve afectado no sólo por la ratio de alumnos por profesor, sino también por aspectos como el número de clases o estudiantes de los que un profesor es responsable, la cantidad de tiempo de instrucción que tienen que impartir los docentes, las horas de clase que recibe el alumnado o la posible existencia de asignaturas optativas o de desdobles... Así, por ejemplo, un mayor número de asignaturas optativas o desdobles conducirá a un mayor tamaño de las clases comunes. La relación entre el tamaño estimado de clase, que no es el mismo que el real por la existencia de asignaturas optativas y los posibles itinerarios, se aproxima a la ratio de alumnos por profesor multiplicado por el cociente entre horas de clase que reciben los estudiantes y las horas de instrucción de los docentes:

$$\textit{Tamaño estimado de clase} \approx \textit{Ratio alumnos por profesor} \times \frac{\textit{Horas de clase de estudiantes}}{\textit{Horas de instrucción de docentes}}$$

La ecuación anterior muestra que dos países con la misma ratio alumnos por profesor pueden tener distintos tamaños estimados de clase. Por ejemplo, si nos centramos en la ESO, España tiene más horas de instrucción de los estudiantes al año, 1.057 horas como muestra Education at a Glance 2023 de la OCDE, que las de la OCDE (916) o la UE-22 (876). Nuestro país tiene también menos horas de instrucción de los docentes al año en ESO (669) que en la OCDE (723) aunque muy similar a la UE-22 (660). El resultado es que una ratio de alumnos por profesor en España similar a la OCDE y la UE-22 (12, por 13 y 11; respectivamente), pero con un cociente más elevado de horas de clase de estudiantes dividido por horas de instrucción de docente, conduce a un mayor tamaño estimado de clase.

Ejercicio 1. *Calcula el tamaño estimado de clase de dos países, A y B, con la misma ratio de alumnos por profesor, 14, pero con diferentes horas de clase de estudiantes y de horas de instrucción de docentes:*

*País A (datos reales de España en Primaria, 2019):*

- *Ratio alumnos por profesor: 14*
- *Horas de clase de estudiantes al año: 792*
- *Horas de instrucción de docentes al año: 871*

*País B (datos aproximados de la UE-22, 2019):*

- *Ratio alumnos por profesor: 14*
- *Horas de clase de estudiantes al año: 766*
- *Horas de instrucción de docentes al año: 871*

*Suponga ahora que los datos de las horas de instrucción de los docentes al año cambian en el país B. ¿Cuál sería entonces el tamaño estimado de la clase?*

*País A (datos reales de España en Primaria, 2019):*

- *Ratio alumnos por profesor: 14*
- *Horas de clase de estudiantes al año: 792*
- *Horas de instrucción de docentes al año: 871*

*País B (datos reales de la UE-22, 2019):*

- *Ratio alumnos por profesor: 14*
- *Horas de clase de estudiantes al año: 766*
- *Horas de instrucción de docentes al año: 740*

Ejercicio 2. *Ir a la tabla V.B1.4.9 y la tabla V.B1.4.11 del Informe PISA 2018 OCDE en https://doi.org/10.1787/888934132241 obtener los datos de* Student-teacher ratio y Language-of-instruction class size. *Calcular el coeficiente de correlación entre la ratio alumnos por profesor y ratio alumnos por clase para los 37 países de la OCDE y comprobar la posición en la que se encuentra España entre los 37 en las dos variables.*

## 2. LA ASOCIACIÓN ENTRE EL TAMAÑO DE CLASE Y EL RENDIMIENTO EDUCATIVO: LA PRIMERA GENERACIÓN DE ARTÍCULOS NO CAUSALES

Durante décadas se ha argumentado que el gasto en educación, incluyendo la inversión que requiere reducir el tamaño de las clases, no estaba correlacionado con el rendimiento académico (véase recuadro 2). Hanushek y Woessmann (2011), por ejemplo, concluían que la evidencia general sobre el efecto del tamaño de la clase en el rendimiento de los estudiantes, medido a través de pruebas externas y estandarizadas, era mixta o al menos no identificaba beneficios notables derivados de la reducción del número de alumnos por aula.

**Recuadro 2. La asociación no causal entre el tamaño de la clase y resultados en PISA**

Los análisis de PISA revelan que algunos de los sistemas educativos con mejores resultados tienen, sin embargo, muchos estudiantes por clase. La tabla al final de este recuadro muestra los datos del volumen V del Informe PISA 2018 (OCDE, 2020b, tabla V.B1.4.11). La información sobre el tamaño promedio de la clase está basada en los datos que proporcionan los directores de los centros educativos de Secundaria que participan en la muestra de cada país en esta evaluación y se centran en las clases en las que se encuentran los estudiantes de 15 años. La tabla 1 anterior, por el contrario, estaba basada en datos administrativos proporcionada por los Ministerios de Educación de cada país y recogidos también por la OCDE. En cualquier caso, las conclusiones son similares pues se confirma que España tiene una elevada ratio de alumnos por clase (29,4), por encima del promedio de la OCDE y de los países de su entorno europeo, con la excepción de Francia. Nuestro país también tiene una gran heterogeneidad en el tamaño de las clases, como refleja la alta desviación típica. La tabla muestra que tanto en España como en el conjunto de la OCDE el tamaño de las clases en la asignatura de Lengua –la asignatura central de PISA 2018– es más reducido en los centros educativos con menor nivel económico, social y cultural.

Como se comprueba en la tabla de este recuadro, Alemania (con 25,0 estudiantes por clase) y Reino Unido (24,9) tienen mejores resultados en PISA 2018 en Lectura, Matemáticas y Ciencias que Italia, a pesar de que el tamaño promedio de la clase de este último es menor (22,9). También se pueden encontrar comparaciones en sentido contrario, pues Alemania y Reino Unido tienen también mejores resultados en PISA que países con más ratio alumnos por aula como Francia (29,8) o el conjunto de la OCDE (26,1). El análisis de la OCDE de la columna 3 muestra una asociación no causal positiva, contraria a la esperada, entre el tamaño de la clase y el rendimiento en Lectura incluso después de haber tenido en consideración el nivel socioeconómico promedio de los centros educativos y el individual de las familias de los alumnos. En concreto, en el conjunto de la OCDE un aumento en un estudiante de la ratio de alumnos por aula se asocia a un incremento en un punto en los resultados de PISA Lectura, y de hasta 4 en el caso de Francia. Esta asociación contraintuitiva entre tamaño de clase y rendimiento educativo del Informe PISA está afectada por posibles sesgos de variables omitidas y endogeneidad de los recursos. En

su reciente survey Handel y Hanushek (2023) señalan que este tipo de análisis correlacionales y no causales sólo muestran relaciones limitadas entre los recursos escolares –en este caso la ratio alumnos por aula– y los resultados de los estudiantes. La evidencia reciente ofrece una nueva perspectiva, prestando mucha más atención a la identificación de los parámetros clave de las políticas. Durante el resto de este capítulo se resumirá evidencia de lo que Handel y Hanushek (2023) han llamado la "revolución de credibilidad" en el análisis empírico moderno, que llamaremos artículos de "segunda generación" en este capítulo, con estudios que aprovechan la variación exógena de las ratios de alumnos por clase y que muestran una relación inversa o nula entre ambas, pero nunca positiva.

Por lo demás, la tabla no muestra los resultados de PISA 2018 de España porque la OCDE concluyó que *"no se puede garantizar completamente la comparabilidad de los datos de PISA 2018 para España con los de evaluaciones anteriores de PISA"* (OCDE, 2020c).

Tamaño de clase en Lengua, características del centro y rendimiento en Lectura. Informe PISA a alumnos de 15 años. Resultados basados en informes de directores de escuela.

| | **Promedio** | **Desviación estándar** | **Más bajo** | **Más alto** | **Cambio en Lectura por aumento de la clase en un estudiante** | | **Lectura** | **Matemáticas** | **Ciencias** |
|---|---|---|---|---|---|---|---|---|---|
| Francia | 29,8 | 4,9 | 26,0 | 33,2 | **4** | (0,5) | 493 | 495 | 493 |
| Alemania | 25,0 | 4,5 | 22,0 | 26,8 | **2** | (0,9) | 498 | 500 | 503 |
| Italia | 22,9 | 5,1 | 21,5 | 23,9 | 1 | (0,7) | 476 | 487 | 468 |
| España | 29,4 | 9,7 | 28,9 | 30,6 | – | – | – | – | – |
| Reino Unido | 24,9 | 4,3 | 24,8 | 23,1 | 0 | (0,5) | 504 | 502 | 505 |
| Estados Unidos | 26,2 | 5,4 | 26,9 | 26,4 | **1** | (0,4) | 505 | 478 | 502 |
| OCDE | 26,1 | 5,6 | 23,8 | 27,4 | **1** | (0,1) | 487 | 489 | 489 |

Nota: El perfil socioeconómico se mide mediante el índice PISA de estatus económico, social y cultural. Los valores en negrita son estadísticamente significativos.
Fuente: OCDE. PISA 2018 Results (Volume V). Tabla V.B1.4.11 https://doi.org/10.1787/888934132241
https://doi.org/10.1787/888934028140

Ejercicio 3. *Ir a la tabla V.B1.4.11 de la OCDE en https://doi.org/10.1787/888934132241 y descargar en un Excel los datos de la ratio alumnos por clase en los países de la OCDE. Bajar la información sobre los resultados de cada país en PISA 2018 en Lectura, Matemáticas y Ciencias (https://doi.org/10.1787/888934028140, PISA. Volume I). Calcular, usando la función en Excel de correlación, la asociación entre la ratio alumnos por clase, con los resultados en Lectura, Matemáticas y Ciencias.*

Estos análisis de "primera generación", que no suelen encontrar una relación entre el tamaño de clase y el rendimiento educativo, generalmente no han tenido en cuenta la naturaleza no aleatoria del tamaño de la clase. No consideran la posibilidad de que los agentes educativos asignen el número de alumnos por aula en función de factores como la capacidad de los estudiantes o que las familias seleccionen el centro educativo en por el tamaño esperado de la clase. Pero la realidad nos muestra que algunos de los cambios en los inputs educativos que experimentan los estudiantes se producen como resultado de decisiones llevadas a cabo por sus padres, docentes, directores de centros educativos o gestores de la educación. Alguno de estos estudios de "primera generación" usa datos observacionales que a menudo son transversales o no son datos individuales sino agregados. Incluso en aquellos trabajos que emplean bases de datos longitudinales, no se puede estar seguro de que se estén incluyendo todas las variables relevantes y donde la obtención de relaciones causales firmes puede ser muy compleja (Cabrales, 2010). Como concluyen Angrist y Pischke (2015), la correlación observada entre el tamaño de la clase y los resultados de los estudiantes no debe tomarse como evidencia rigurosa. Si, por ejemplo, algunos de los estudiantes con dificultades de aprendizaje se asignan a clases más pequeñas para ofrecer una atención más individualizada, la evidencia empírica utilizando Mínimos Cuadrados Ordinarios (MCO) detectaría un impacto positivo del tamaño de la clase en el rendimiento del estudiante que no sería real. Otro factor endógeno en la determinación del tamaño de las clases es el caso de que los centros con menores ratios de alumnos por aula tengan un mejor rendimiento académico debido a que sean muy selectivos o porque disponen de más recursos económicos, que incidirá de forma directa en el aprendizaje del alumnado (Graves y Zuchelli, 2022).

## 3. EL IMPACTO CAUSAL DEL TAMAÑO DE LAS CLASES EN EL RENDIMIENTO EDUCATIVO

Establecer una relación causa-efecto del tamaño de las clases y el aprendizaje de los estudiantes es una tarea difícil. Los trabajos de "segunda generación" abordan los problemas de endogeneidad que sesgan las estimaciones de MCO, utilizando fuentes exógenas de variación que son fundamentales para una identificación creíble del efecto de la disminución del número de alumnos por aula. Los estudios que veremos a partir de esta sección 3 emplean métodos de identificación convincentes. Se tienen en cuenta variaciones en el tamaño de las clases que no guardan relación con factores que puedan repercutir en el rendimiento académico. Son investigaciones que pueden atribuir claramente sus efectos estimados al número de alumnos por aula. Para determinar el efecto causal del tamaño de la clase se han utilizado los siguientes métodos analíticos que revisamos a continuación.

### 3.1. Experimentos aleatorios controlados

Como señalábamos anteriormente la asignación no aleatoria de estudiantes a clases, como la asistencia de alumnos con dificultades de aprendizaje a aulas más pequeñas de apoyo, genera relaciones espurias entre el tamaño de clase y el rendimiento educativo. Los ensayos controlados aleatorios (RCTs por sus siglas en inglés) resuelven este problema. Este procedimiento asigna aleatoriamente distintos tamaños de clases para estimar los efectos causales evitando el sesgo de selección. Se elige un grupo de estudiantes al azar para beneficiarse de un tamaño de clase reducido, y se compara su desempeño con el de un grupo de control que ha estado en aulas con el número habitual de alumnos. El proyecto STAR (*Student/Teacher Achievement Ratio)* desarrollado en Tennessee en los años 80, es precisamente el experimento más célebre en este ámbito. Entre 1985 y 1989, este programa STAR asignó aleatoriamente entre 1985 y 1989 a 11.571 alumnos de Educación Infantil de Tennessee a clases pequeñas (15 alumnos) o a clases grandes (22 alumnos). Los estudiantes de las 79 escuelas que se ofrecieron para participar permanecieron desde Infantil hasta tercer Grado (tercero de Primaria en España) en los grupos originales de tratamiento y control y fueron evaluados al final de cada curso. El análisis de Krueger (1999) sobre el ensayo controlado aleatorio del tamaño de clase de Tennessee STAR muestra efectos causales sólidos y robustos de la reducción del tamaño de clase en el aprendizaje. En concreto, Krueger (1999) concluyó que una disminución de 7 alumnos

(de 22 a 15) en la ratio alumnos por aula mejora los resultados en un 22% de la desviación estándar, el equivalente a lo que un estudiante aprende en seis meses (Hanushek y Woessmann, 2020). Los efectos son más positivos para estudiantes de origen minoritario o aquellos que eran elegibles para las becas de comedor, un indicador de pobreza en EE. UU. La estimación de Krueger (1999) basada en los datos del Proyecto STAR tiene uno de los impactos estimados más elevados de todos los que veremos en este capítulo. Aunque es una evidencia de por sí muy importante, Krueger y Whitmore (2001) muestran que el efecto de la intervención desaparece con el paso del tiempo. Para el año siguiente al experimento, la diferencia entre los estudiantes del grupo de tratamiento y de control se reduce entre la mitad y tres cuartas partes de la diferencia inicial. La mayoría de las mejoras encontradas en las clases pequeñas ocurrieron en Educación Infantil, con beneficios mucho menores en primero de Primaria y casi inexistentes en segundo y tercero de Primaria.

Chetty et al. (2011) analizan también el mismo proyecto STAR y concluyen que estar asignado a clases pequeñas incrementa en 1,8 puntos porcentuales la probabilidad de llegar a estudios superiores a los 20 años, una mejora significativa dado que la media era del 26,4%. Además, los autores muestran un impacto considerable de asistir a clases pequeñas en Educación Infantil en la calidad de la Universidad a la que se asiste y en los salarios a los 27 años. En concreto, cada punto porcentual de mejora en la distribución de las notas se asocia a un aumento en los salarios con 27 años de 94 dólares anuales. Los estudiantes de clases más reducidas también tienen otros resultados positivos más a largo plazo como una mayor probabilidad de tener una casa en propiedad o ahorros en planes de pensiones.

Después del STAR, ha habido pocos ensayos aleatorizados más con una escala suficientemente grande como para hacer inferencias significativas, debido en muy buena medida al elevado coste que implican estos experimentos. Entre los ensayos llevados a cabo más recientemente se encuentra el de Duflo et al. (2015) en el oeste de Kenia, donde se asignó aleatoriamente un docente adicional a 70 centros, reduciendo efectivamente el tamaño promedio de las clases de 82 a 44 estudiantes. Estos autores no encuentran efectos positivos en el aprendizaje en los estudiantes cuyos profesores ya estaban previamente contratados a pesar de la reducción de la ratio de alumnos por aula. Sin embargo, los resultados mejoraron en 0,24 desviaciones estándar para aquellos estudiantes que fueron asignados a los nuevos docentes contratados. Una posible explicación es que estos docentes recién incorporados presentaban bajas tasas de absentismo, mientras

que los antiguos profesores de los centros educativos del grupo de tratamiento redujeron su esfuerzo de manera endógena.

Ejercicio 4. *Esther Duflo y Michael Kremer, dos de los autores del este artículo Duflo, E., Dupas, P., y Kremer, M. (2015). School governance, teacher incentives, and teacher-pupil-ratios: experimental evidence from Kenyan primary schools.* Journal of Public Economics*, 123, 92-110 https://doi.org/10.1016/j.jpubeco.2014.11.008 obtuvieron el Premio Nobel de Economía en 2019. En este ejercicio te pedimos que leas el artículo, de libre acceso en https://web.stanford.edu/~pdupas/DDK_ETP.pdf, y elabores otras posibles interpretaciones que expliquen la razón por la que estos investigadores no encuentran un efecto de la reducción del tamaño de la clase en los estudiantes de los profesores que ya están contratados mientras que sí hallan un impacto significativo en los alumnos de los docentes recién contratados.*

Andersen et al. (2020) evalúan los efectos de dos tratamientos en Dinamarca en centros educativos de Primaria. En el primer grupo de tratamiento, los estudiantes tienen un asistente adicional con certificado de docente mientras que, en el segundo grupo, el asistente no tiene esta cualificación. Los dos tratamientos tienen un efecto positivo y significativo en la comprensión lectora de los alumnos, pero no en Matemáticas. En general, en economía de la educación la conclusión suele ser más bien la contraria, que las intervenciones tienden a ser más efectivas para aumentar el rendimiento en Matemáticas en comparación con el rendimiento en Lectura, como muestra el análisis de 196 experimentos controlados y aleatorios de Fryer (2017). Los estudiantes asignados a la intervención con un co-docente con título de enseñanza mejoraron sus calificaciones de Lectura en 0,086 desviaciones estándar mientras que los estudiantes que se beneficiaron de un asistente adicional sin el certificado aumentaron su competencia lectora en 0,131 desviaciones estándar, debido a que con el mismo presupuesto se podían contratar a más ayudantes (Andersen et al., 2020). El impacto perdura en el tiempo para los estudiantes rezagados y en desventaja socioeconómica. Un análisis más profundo de los mecanismos indica que incorporar un asistente no certificado de profesor no es simplemente una reducción en el tamaño de la clase, sino que es particularmente eficaz cuando comparte la responsabilidad de la instrucción en el aula y en grupos con estudiantes con problemas de comportamiento. En el caso de la co-docencia, la buena coordinación, con las funciones y roles bien determinados, son clave para el buen funcionamiento de esta intervención, con uno de los dos trabajando en pequeños grupos.

MacKinnon et al. (2023) estiman de nuevo el efecto de la disminución del tamaño de la clase en el rendimiento académico con los datos del proyecto STAR como aplicación práctica de los test que proponen para determinar el nivel correcto de agrupamiento en modelos de regresión. Estos autores encuentran que las clases pequeñas conducen a mejores resultados en Lectura en primero de Primaria y muestran que si se incluyen efectos fijos por colegio, el cluster a nivel de centro es el más apropiado con respecto al de aula. Incluyendo variables de control adicionales, como género, raza del alumno y del docente, condición de becario de comedor escolar, experiencia y cualificación del docente, resultados previos del estudiantes, y trimestre de su nacimiento, además de efectos fijos del centro educativo, la dummy de si el estudiante asistió a una clase pequeña (15 alumnos) en lugar de 22, tiene un efecto positivo y significativo en el rendimiento del alumno en primero de Primaria con los mismos datos del proyecto STAR analizados por Krueger (1999).

### 3.2. Regresiones de discontinuidad

Los ensayos controlados aleatorios que veíamos en la sección anterior no siempre son fáciles de llevar a cabo por el elevado coste económico que implican y por razones también éticas, pues el alumnado asignado a clases con muchos estudiantes podría tener menos oportunidades educativas y de inserción laboral en el futuro. De hecho, Chetty et al. (2011) muestran que estar en clases con menos alumnos de pequeño tiene efectos positivos a largo plazo en el desarrollo académico en estudios superiores y en inserción laboral. Hanushek et al. (2015) hallan además que las habilidades medidas por las puntuaciones de las pruebas externas están estrechamente relacionadas con los resultados económicos posteriores, incluyendo los ingresos individuales en el mercado laboral.

La dificultad de realizar ensayos aleatorios ha provocado que se haya desarrollado un cuerpo de investigación que busca aprovechar las variaciones naturales en el tamaño de la clase, cambios que no sean influenciados por las elecciones de los agentes educativos. Un enfoque ha sido emplear reglas administrativas con respecto a los tamaños máximos de clase como fuente de variación exógena cuasi-aleatoria (véase Comisión Europea, 2023: National Education Systems sobre regulación en los Estados miembros de la UE del tamaño máximo de las clases https://eurydice.eacea.ec.europa.eu/national-education-systems). Este método utiliza las reglas sobre el tamaño máximo de la clase para comparar los cursos en los que se abren más clases por pasarse del límite (disminuyendo el tamaño me-

dio de la clase) con los que se quedan en la frontera, pero que tienen un número de estudiantes total por curso muy similar. Por ejemplo, si la regla fija en 30 el número máximo de alumnos por clase, y un curso pasa justo a 31 estudiantes, tendrán que abrir una segunda clase con un tamaño promedio de las dos aulas de 15,5, frente al 30 anterior. Lo mismo ocurriría si la matrícula de un curso aumenta de 60 a 61 estudiantes, con un cambio en ratio de alumnos por aula de 30 a 20,3. Los trabajos que repasamos en esta sección utilizan como instrumento estas discontinuidades en el tamaño de clase que no reflejan las elecciones de los centros educativos o los padres.

Angrist y Lavy (1999) se centran en los estudiantes de 4º y 5º curso de centros de Primaria en Israel en los años 1991 y 1992, donde se utiliza la regla de Maimónides que especifica que las clases no deben tener una ratio superior a 40. Esta regla lleva a discontinuidades en el tamaño de la clase cuando la matriculación en un curso supera el número de 40, 80, 120 y en general múltiplos de 40. En los centros con 40 estudiantes en el mismo curso, los alumnos estarán en una misma clase. Si llega uno más, haciendo un total de 41, la cohorte se dividiría en dos clases más pequeñas. En la práctica no siempre se cumple de forma rigurosa la división de clases, lo que lleva a un enfoque de variable instrumental, en el que la clase real se instrumenta con el tamaño de clase predicho derivado de la regla de división de clases. La identificación solo proviene de contextos donde estas reglas son vinculantes. Utilizando esta metodología, Angrist y Lavy (1999) muestran que un aumento de 8 estudiantes en el tamaño de la clase provoca una reducción de 0,29 de una desviación estándar en las puntuaciones. Bajo el supuesto de linealidad, sería un efecto aproximado de 0,035 de una desviación estándar por un incremento de un estudiante en el tamaño de la clase.

Fredriksson et al. (2013) utilizan la variación exógena que genera el límite máximo de 30 alumnos por aula en Primaria en Suecia, empleando como instrumento los diferentes umbrales como puntos de discontinuidad en los que el tamaño de la clase da un salto. El cumplimiento de la norma es elevado, pero no perfecto, lo que conduce a un enfoque de variable instrumental. Estos autores también llevan a cabo estimaciones de forma reducida del diseño de "fuzzy RD" (regresión discontinua difusa, véase el capítulo 2 de métodos empíricos y la introducción a la inferencia causal), empleando como un instrumento para el tamaño de la clase las discontinuidades de los diferentes puntos de corte. La regla del tamaño de la clase es un predictor muy fuerte del tamaño real de la clase, por lo que las estimaciones de variables instrumentales y las regresiones de discontinuidad no difieren cualitativamente. Fredriksson et al. (2013) encuentran que

un aumento de un estudiante en el tamaño de la clase causa una reducción de entre 0,032 y 0,047 de una desviación estándar en el rendimiento del estudiante en habilidades cognitivas (habilidades verbales y lógicas). Argaw y Puhani (2018), por su parte, aprovechan la regla de la ratio máxima de 25 alumnos por aula en el estado de Alemania de Hesse para examinar el efecto del tamaño de la clase en los resultados al final de la enseñanza Primaria. Estos autores encuentran que, aunque sólo en el caso de los chicos, un aumento en el tamaño de la clase de 10 estudiantes reduce la oportunidad de asistir al itinerario más académico en Secundaria en 3 puntos porcentuales, una elección que tiene importantes efectos a lo largo de la vida.

Angrist et al. (2019) volvieron a analizar el impacto del tamaño de la clase en el rendimiento académico en Israel, una investigación que ya habían llevado a cabo Angrist y Lavy (1999), pero con datos más recientes del periodo 2002-2011. Utilizaron de nuevo la regla de 40 estudiantes por clase de Maimónides para llevar a cabo estimaciones de variables instrumentales del impacto del tamaño de la clase en los resultados de los exámenes. Además, Angrist et al. (2019) observaron que hay cierta evidencia de manipulación de la matrícula en algunos centros de Israel, en el sentido de que hay demasiadas clases con el número exacto de estudiantes que provoca su desdoblamiento. Utilizando una estrategia que emplea la fecha de nacimiento de los estudiantes para calcular la matrícula y evitar el sesgo de manipulación, llegaron a la conclusión de que no hay un efecto significativo del tamaño de las aulas en los resultados de los exámenes estandarizados y externos. Es decir, Angrist et al. (2019) encuentran que con datos más actuales desaparece el efecto positivo que los mismos Angrist y Lavy (1999) hallaron de la reducción del número de estudiantes por aula en los resultados académicos.

Ejercicio 5: *Explica las razones por las que el artículo de Angrist et al. (2019, 10.1257/aeri.20180120) no encuentra el mismo resultado que Angrist y Lavy (1999, https://www.jstor.org/stable/2587016) a pesar de analizar datos del tamaño de las clases de Israel, aunque sea en periodos diferentes.*

Connolly y Haeck (2022) estiman el impacto de la reducción del tamaño de la clase en el rendimiento del estudiante aprovechando las regulaciones que limitan la ratio máxima de estudiantes por aula a 20 en Educación Infantil en Quebec. Estos autores se diferencian de Angrist y Lavy (1999) en el sentido de que no emplean directamente la regla de tamaño de clase como un instrumento por los problemas que representan el posible agrupamiento de aulas alrededor

de los umbrales y composición de la clase. Su enfoque es más cercano al diseño de "fuzzy RD" (regresión discontinua difusa) de Fredriksson et al. (2013). Al analizar la Educación Infantil, Connolly y Haeck (2022) se centran en un nivel de enseñanza que reduce la probabilidad de que el centro pueda clasificar de alguna manera a los alumnos, dado que se sabe muy poco del niño antes de que comience la escuela. Estos autores encuentran que la disminución del número de alumnos por clase aumenta el desarrollo cognitivo en 0,014 desviaciones estándar. aproximadamente la mitad de impacto que haya Krueger (1999). Estos autores también hallan que los niños que viven en áreas geográficas desfavorecidas se benefician más de la reducción de la ratio alumnos por aula, mientras que el efecto para el resto de los alumnos no es significativo. Connolly y Haeck (2022) también encuentran un impacto positivo, aunque de menor medida, de la reducción del tamaño de la clase en las habilidades no-cognitivas (como la competencia social, autoconfianza, sentido de la responsabilidad, autonomía, hábitos de trabajo y el respeto por los compañeros) de los niños de familias desfavorecidas. Estos autores concluyen que el impacto del tamaño de la clase en el rendimiento del estudiante es no lineal pues el efecto marginal de la disminución de la ratio de alumnos por aula es más elevado en las clases más pequeñas. La conclusión de este trabajo es que la disminución de la ratio alumnos por clase centrada a colegios ubicados en zonas vulnerables lograría una mejora educativa más elevada que una disminución generalizada de las ratios.

En resumen, los análisis de regresiones discontinúas basados en las normas de máximos de ratio de aula por clase, encuentran generalmente un efecto positivo de la reducción del tamaño de la clase en los resultados educativos. En su survey, Green e Iversen (2022) también señalan algunas excepciones a esta conclusión de efectos positivos de la disminución del tamaño de la clase. Leuven et al. (2008) utilizan la regla del número máximo de 30 alumnos por aula en Secundaria inferior (nuestra ESO) en Noruega, en donde esta norma es un fuerte predictor del tamaño de clase real, y encuentran efectos nulos de la reducción de la ratio alumnos por clase en el rendimiento educativo en Lenguaje o Matemáticas para los estudiantes en 2002 y 2003.

### 3.3. Variables instrumentales

Uno de los problemas del uso de reglas de máximo número de estudiantes por aula es que para algunos centros educativos estos umbrales no son vinculantes, porque se encuentran lejos de la discontinuidad (Hoxby, 2000). Un en-

foque alternativo a las reglas del tamaño es utilizar cambios en la matrícula por curso dentro de un mismo centro educativo debido a la variación demográfica, que es también poco influenciable por las elecciones de los agentes educativos. Los padres no pueden prever la variación de año en año en el tamaño de la matriculación en un curso en un centro determinado. Se emplean datos longitudinales para explotar la variación aleatoria anual en el tamaño de las cohortes escolares, que inducen cambios en el tamaño de las clases. Una ventaja adicional es que esta metodología conduce a estimaciones de cambios en un estudiante en el tamaño de la clase sin la necesidad de hacer suposiciones de linealidad a partir de cambios mayores en la ratio alumnos por aula. El diseño dentro de la escuela es ampliamente aplicable y permite estudiar los efectos del tamaño de la clase incluso si no existen reglas de tamaño de clase o cuando no se puede identificar fácilmente el umbral correcto de tamaño de clase, porque están en vigor diferentes umbrales que dependen de características no observables para el investigador (Bach y Sievert, 2020). Los diseños dentro de la escuela permiten la estimación de efectos heterogéneos del tamaño de la clase a lo largo de toda la distribución del tamaño de la clase, pudiendo contrastar la existencia de efectos no lineales.

Sin embargo, el empleo de la variación año tras año en las matrículas por curso y centro educativo implica requisitos de datos longitudinales que no siempre están disponibles, y una pérdida de precisión estadística debido tanto a la falta de grandes cambios en la matrícula como a un mayor potencial de error en la medición de los números de matrícula (Green e Iversen, 2022). El artículo seminal que utiliza este método es el de Hoxby (2000). Esta autora utiliza como instrumento las desviaciones de los datos de matrícula con respecto a la tendencia en colegios de Primaria en Connecticut. Hoxby (2000) no encuentra efectos significativos de la variación del tamaño de la clase en el rendimiento de los estudiantes y descarta efectos tan pequeños como 0,04 desviaciones estándar para una reducción del 10 por ciento en el tamaño de la clase. Heinesen (2010) examina el efecto de la ratio de alumnos por aula en los resultados académicos de los estudiantes de Dinamarca en la asignatura de francés. La variación en este caso proviene tanto de la matrícula por curso y centro como de la elección de idioma de los estudiantes entre francés o alemán a lo largo del tiempo. Este autor halla que la reducción en un alumno en el tamaño de la clase aumenta en 0,029 desviaciones estándar el rendimiento educativo. Cho et al. (2012) utilizan la variación en matriculación total dentro del colegio para examinar los resultados de los alumnos de 3º y 5º de Primaria en Minnesota. Encuentran que una dis-

minución de diez estudiantes en el tamaño de la clase llevaría a un aumento en las puntuaciones de los exámenes entre 0,04 y 0,05 de una desviación estándar.

Más recientemente, Kedagni et al. (2021) también utilizan el enfoque de Hoxby (2000) de variaciones en la matrícula general como shocks exógenos para analizar el impacto de la ratio alumnos por clase en el aprendizaje de los estudiantes de Secundaria en Grecia. En este país, la gran mayoría de los estudiantes asisten al centro educativo más cercano y los institutos tienen que admitir a todos los estudiantes elegibles. Por tanto, las variaciones entre promociones de un mismo instituto son exógenas a los padres y centros. Kedagni et al. (2021) encuentran que la disminución en un estudiante del tamaño de la clase en el décimo curso (nuestro 4º de la ESO) en Grecia incrementa el rendimiento académico en 0,017 desviaciones estándar. Kedagni et al. (2021) también proporcionan evidencia de una relación no lineal entre la proporción de alumnos por clase y los resultados académicos. El tamaño de la clase inicialmente aumenta el rendimiento académico, dado que los estudiantes aprenden tanto del docente como entre ellos. Los estudiantes aprenden de compañeros con un nivel de competencias similar al suyo. Cuanto más grande es la clase, más probable es que tengan alumnos con resultados académicos parecidos. Pero alcanzado ya un límite en la ratio alumnos por aula, el tamaño comienza a tener un impacto negativo en el aprendizaje de los estudiantes porque los efectos de aglomeración y las fuerzas de congestión empiezan a prevalecer. Además, Kedagni et al. (2021) elaboran un modelo estructural dinámico con costes de despido y contratación de docentes. Reducir los costes de despido a cero tiene un efecto directo al facilitar la desvinculación de los docentes, lo que aumenta el tamaño de la clase en comparación con el status quo. Pero abaratar el despido de los docentes también tiene un efecto indirecto, por cuanto que hace más probable que sean contratados, lo que reduce el tamaño de la clase. Los resultados del contrafactual muestran que el efecto neto es que reducir el coste de despido de los docentes a cero aumenta el tamaño promedio de la clase entre 5 y 6 estudiantes y reduce su resultado académico en aproximadamente un punto en una escala de 1 a 20. La disminución de los costes de contratación reduce el tamaño de la clase entre 4 y 5 estudiantes, dado que se contratan más profesores, y aumenta los resultados académicos en 0,3 puntos. Bach y Sievert (2020) resumen la evidencia de los artículos que emplean la metodología de la variación en el tamaño de las cohortes dentro de cada centro educativo como instrumento en la estimación de los efectos del tamaño de la clase, y concluyen que, de los 11 artículos resumidos, cinco no encuentran que menores ratio alumnos por aula mejoren el aprendizaje (tabla E.1, Bach y

Sievert, 2020). En el recuadro 3 analizamos con más profundidad el artículo de Bach y Sievert (2020) y los problemas que los análisis de variables instrumentales pueden tener para estimar el efecto causal del tamaño de las clases en el resultado de los estudiantes.

Una diferencia importante entre los estudios que emplean regresiones discontinúas o variables instrumentales y el estudio de experimento aleatorio controlado del proyecto STAR (*Student/Teacher Achievement Ratio)* es que los centros educativos que participaron en este experimento STAR recibieron recursos adicionales para reducir el tamaño de las clases. Por el contrario, en los casos que se analizan mediante regresiones discontinuas o variables instrumentales probablemente tuvieron que desviar recursos de otros ámbitos educativos cuando la variación natural de la población llevó a clases más pequeñas. Por lo tanto, una posible interpretación de los resultados divergentes es que los efectos positivos encontrados en el experimento STAR se lograron, al menos en parte, gracias a los recursos adicionales.

**Recuadro 3. El artículo de Bach y Sievert (2020) con más detalle**

Una gran parte de la literatura cuasi-experimental aprovecha la variación dentro de la escuela en el tamaño de la cohorte a lo largo del tiempo para analizar el impacto del número de alumnos por clase en el aprendizaje de los estudiantes (véase, por ejemplo, Hoxby, 2000; Leuven et al., 2008; Cho et al., 2012). Estos estudios generalmente encuentran efectos pequeños o nulos del tamaño de la clase, lo que contrasta con la evidencia experimental disponible que muestra efectos sustanciales del tamaño de la clase (véase, por ejemplo, Krueger, 1999; Krueger y Whitmore, 2001).

Bach y Sievert (2020) encuentran una explicación a esta diferencia en los resultados sobre el efecto del tamaño de la clase en el aprendizaje de los estudiantes. Estos dos autores muestran que en sistemas educativos con repetición de curso o "redshirting" (retrasar la entrada de un alumno a Primaria por un año), el tamaño de la cohorte está negativamente relacionado con la proporción de estudiantes repetidores. Intuitivamente, en cohortes más grandes, los estudiantes retenidos de la cohorte anterior (más pequeña) constituyen mecánicamente una proporción menor de estudiantes en la cohorte dada (más grande). Esta asociación negativa provoca una relación positiva entre el tamaño de la cohorte de nacimiento y las puntuaciones en las pruebas a nivel de curso, ya que los estudiantes que han repetido en el pasado son seleccionados

negativamente. La conclusión es que hay un efecto positivo espurio del tamaño de la clase y el aprendizaje, lo que lleva a un sesgo al alza (haciendo menos negativo) las estimaciones del tamaño de la clase basadas en la variación dentro de la escuela en el tamaño de la cohorte. En términos de variables instrumentales, esto es una violación de la restricción de exclusión del tamaño de la cohorte de nacimiento como instrumento.

Bach y Sievert (2020) ponen como ejemplo una situación en la que generalmente hay un número constante de nacimientos cada año, pero en un año específico (*t*) hay un aumento de alumnado que entra a primero de Primaria. Debido a ese aumento en las matriculaciones en el año *t*, más estudiantes pasan a segundo de Primaria al siguiente año (*t+1*). Por lo tanto, los estudiantes que repitieron del año anterior representan un porcentaje menor en segundo de Primaria en *t+1*. Además, debido al aumento en primero de Primaria en el año *t*, hay más estudiantes que repiten. Estos estudiantes pasan a segundo de Primaria un año después (*t+2*), aumentando el porcentaje de estudiantes repetidores en ese grupo. Es decir, un aumento inesperado en las matriculaciones en primero de Primaria en el año *t* se relaciona con una disminución en el porcentaje de estudiantes repetidores en segundo de Primaria al siguiente año (*t+1*). Si hay menos estudiantes repetidores, que normalmente tienen competencias más bajas, el nivel de habilidades promedio del grupo será más alto. La conclusión es que, incluso si el tamaño inicial de la cohorte no está relacionado con las habilidades académicas y los umbrales de repetición, el efecto neto del sesgo reducirá el tamaño estimado del impacto negativo del tamaño de la clase.

Utilizando datos de todos los colegios de Primaria en el estado federal de Alemania de Saarland, Bach y Sievert (2020) demuestran que el tamaño de la cohorte de nacimiento está sistemáticamente relacionado con la composición de los estudiantes a nivel de curso. Al corregir el sesgo resultante, controlando si un estudiante ha repetido previamente en un curso, hallan que no solo las clases más pequeñas son beneficiosas para los resultados de los exámenes, sino también para reducir las repeticiones.

La conclusión de Bach y Sievert (2020) es que una disminución de un estudiante en el tamaño de la clase desde 1º a 3º de Primaria mejora los resultados en las pruebas en Lengua y Matemáticas al concluir 3º en 0,019 y 0,014 de una desviación estándar, respectivamente. Las tasas de repetición disminuyen también en 0,15 puntos porcentuales (7%) cuando se reduce en

un estudiante el tamaño de la clase. Además, los efectos del tamaño de la clase no son lineales, con efectos grandes en clases más numerosas y sin efectos en clases más pequeñas. Se estima que una reducción de un estudiante en clases con más de 20,5 estudiantes mejora los resultados de las pruebas de Lengua y Matemáticas en un 0,048 y 0,038 de una desviación estándar, respectivamente. Al mismo tiempo, Bach y Sievert (2020) no hallan evidencia que sugiera que reducir el tamaño de la clase mejore los resultados de los estudiantes en clases con menos de 20,5 estudiantes. Los estudiantes en desventaja, por una menor competencia en alemán o trastorno de aprendizaje, se benefician más de asistir a clases más pequeñas con un efecto positivo de cerca del 0,050 de una desviación estándar en Lengua y Matemáticas por la reducción en un estudiante del tamaño de clase. Este impacto es de una magnitud similar a los del experimento del Proyecto STAR.

### 3.4. Leyes de reducción del tamaño de la clase

Los estudios de los tres apartados anteriores aíslan el efecto causal del número de alumnos por aula analizando variaciones en el tamaño de las clases que no guarden relación con factores que puedan repercutir en el rendimiento académico. Estas investigaciones pueden atribuir los efectos estimados exclusivamente al número de alumnos por aula. Sin embargo, en la aplicación real de tales políticas es muy improbable que el resto de los factores educativos permanezcan invariables. Una literatura creciente de investigación utiliza variaciones reales en la ratio alumnos por aula generadas por cambios de políticas educativas, específicamente mandatos de reducción del tamaño de la clase. Estos análisis tienen la ventaja de analizar los efectos de equilibrio general de disminuciones generalizadas de ratio de alumnos por clase en un sistema educativo, incluyendo sus impactos indirectos asociados (Green e Iversen, 2022). Por ejemplo, la disminución del tamaño de la clase requiere de la contratación de muchos docentes, una circunstancia que podría acarrear menor calidad de los profesores recién incorporados. En los apartados previos, los análisis se concentraban en las reducciones de la ratio de alumnos por clase de forma aislada sin considerar los efectos generales de estas medidas.

Ejercicio 6. *A partir del survey de Green e Iversen (2022) y del artículo de Gilraine et al. (2018) señala los posibles efectos indirectos derivados de la reducción del tamaño de clase, distintos de los cambios en el aprendizaje, en los estudiantes, docentes y padres.*

La iniciativa para Reducir el Tamaño de la Clase (Class Size Reduction, CSR) en California, que se implementó en 1996, bajó el número de estudiantes en los grados desde Educación Infantil hasta tercero de Primaria de 30 a 20 alumnos por aula. Según Jepsen y Rivkin (2009), esta intervención en la educación implicó un proceso de contratación masiva de docentes, lo que condujo a la incorporación de profesores sin la formación o experiencia necesaria. Como resultado de esta medida, solo se observó un pequeño aumento en los rendimientos de Matemáticas y Lectura. Adicionalmente, Jepsen y Rivkin (2009) indican que el programa CSR de California generó una alta concentración de docentes sin experiencia en los centros educativos con estudiantes de menor estatus socioeconómico, donde la contratación de profesores resultaba más difícil. Chingos (2012) examina también el mandato de reducción del tamaño (CSR) de las clases de Florida, aprobado en 2002, y que disminuía la ratio alumnos por clase hasta 18 desde Educación Infantil a 3º de Primaria, 22 desde 4º de Primaria a 8º curso (nuestro 2º ESO), y 25 desde 9º (3º ESO) a 12º curso (nuestro 2º de Bachillerato o 2º de FP Media). Chingos (2012) aprovecha que el grado de cumplimiento previo a los nuevos máximos de tamaño de clase era diferente entre centros educativos, con algunos colegios e institutos cerca ya de esos límites mientras que otros se encontraban muy lejos. En concreto, este autor estima el impacto de la disminución de la ratio alumnos por aula comparando las desviaciones con respecto a las tendencias del rendimiento previo en los distritos que tuvieron que reducir el tamaño de la clase con las desviaciones en comparación a los distritos que recibieron recursos equivalentes, pero no se vieron obligados a disminuir las ratios porque ya las tenían bajas. Este autor encuentra que hay efectos pequeños a nulos de las reducciones del tamaño de la clase en el aprendizaje de los estudiantes. La conclusión de Jepsen y Rivkin (2009) y Chingos (2012) es que la contratación masiva de docentes en un mismo ciclo escolar en lugar de hacerlo de manera gradual mitiga los beneficios potenciales de disminuir la ratio de estudiantes por aula, porque conlleva la incorporación de docentes con menos experiencia y con un nivel de preparación inferior, especialmente si los salarios no aumentan. Adicionalmente, una reducción del tamaño de las clases a gran escala puede afectar la distribución de los profesores entre centros educativos. Por ejemplo, la creación de nuevos puestos en colegios e institutos en zonas acomodadas puede atraer a docentes experimentados que previamente trabajasen en poblaciones desfavorecidas.

Han y Ryu (2017) analizan la disminución promedio de 44 a 35 alumnos por clase implementada en centros educativos de Secundaria en Seúl en 2002. Su identificación explota tanto la variación generada por esta medida, así como

las diferencias de cuándo y cómo se implementó en cada instituto. Estos autores encuentran también efectos pequeños en los resultados académicos de los alumnos. Gilraine et al. (2018) analizan de nuevo los efectos del mandato de reducción del tamaño de las clases en California sobre el rendimiento académico de los alumnos. Estos autores utilizan una estrategia de diferenciación que aprovecha la implementación gradual de la reforma por curso centrándose, a diferencia de Jepsen y Rivkin (2009), en el distinto impacto de la medida en los centros públicos y privados.

Jackson et al. (2016) estudian el efecto de los cambios inducidos por las reformas financieras en la enseñanza no universitaria que se implementaron en EE. UU. desde principios de la década de 1970 y que se aceleraron en la década de 1980. Estos autores utilizan el momento de la implementación de las reformas mandatadas por los tribunales y el cambio en la fórmula de financiamiento como un shock exógeno del gasto escolar. Aprovechando que los nacidos entre 1955 y 1985 tuvieron una exposición diferente a las reformas financieras escolares, dependiendo del lugar y año de su nacimiento, muestran que aumento del 10% en el gasto por alumno cada año durante los doce años de enseñanza resulta en 0,27 años adicionales completados de educación, un 7,25% de aumento en salarios, y una reducción del 3,67% en la incidencia anual de pobreza en adultos. Además, los efectos son mucho más notables para los niños de familias con bajos ingresos. Los incrementos en el gasto escolar inducidos por la reforma estuvieron asociados con una reducción en la ratio alumnos por profesor, más días de clase al año y aumento del salario de los docentes. Jackson et al. (2016) concluyen que los efectos positivos del aumento de la financiación de la educación están impulsados, al menos en parte, por la reducción en el tamaño de la clase.

Gilraine et al. (2018) encuentran efectos positivos directos de la disminución de la ratio alumnos por clase en los resultados de los alumnos. La reducción del tamaño de las clases en los centros públicos también redujo la ventaja competitiva de los centros privados de California. La entrada en centros públicos de estudiantes de mayor nivel socioeconómico provenientes de centros privados incrementa el potencial del "efecto compañero" que se analiza en el capítulo 3c de este libro, además de proporcionar la posibilidad de que los centros públicos incorporen a docentes de la privada, que ahora tienen menos estudiantes, para alcanzar las nuevas ratios de alumnos por clase. Gilraine et al. (2018) analizan el beneficio a largo plazo de la ley de reducción del tamaño de la clase (CSR) en California utilizando estimaciones en Chetty et al. (2011). El coste del CSR analizado es de alrededor de 1.100 dólares por estudiante (a precios de 2009). El aumento en los ingresos

del valor presente individual creado por la ley de reducción del tamaño de clase de California –basándose en la investigación de Chetty et al. (2011) sobre la relación entre tamaño de clase, aprendizaje y salarios futuros– es de 18.300 dólares (a precios de 2009). Por lo tanto, el beneficio neto del CSR analizado por Gilraine et al. (2018) implica una relación coste-beneficio sustancialmente positiva.

En conjunto, los papers que analizan la introducción de leyes que reducen los máximos de ratios de alumnos por aulas y examinan los efectos directos e indirectos de equilibrio general de estas medidas encuentran efectos menos positivos de las disminuciones del tamaño de las clases que las metodologías previas, con la excepción del último artículo de Gilraine et al. (2018).

Ejercicio 7. *El artículo de Banerjee et al. (2017) (con dos Premios Nobel de Economía entre sus autores, el propio Abhijit Banerjee y Esther Duflo) señala que "La promesa de los ensayos controlados aleatorios es que la evidencia recopilada a través de la evaluación de un programa específico nos ayuda, posiblemente después de varias rondas de ajustes y múltiples replicaciones en diferentes contextos, a informar la política. Sin embargo, los críticos han señalado que una posible limitación en esta agenda es que los resultados de pequeños estudios realizados por "prueba de concepto" de organizaciones no gubernamentales pueden no aplicarse a políticas a gran escala que puedan ser implementadas por gobiernos. Después de discutir los posibles problemas, este trabajo escribe la transición desde el concepto original hasta el diseño y evaluación de una política escalable. Banerjee et al. (2017) concretan su análisis con el ejemplo de un experimento aleatorio controlado de la implementación de la metodología "Enseñando al Nivel Correcto" de la organización no gubernamental Pratham en los colegios de Primaria de India. Resume los principales retos con los que se encuentra una política educativa para que se implemente a gran escala, basándose en este importante artículo.*

### 3.5. Comparación de ratios de los alumnos por clase y rendimiento por país a partir de PISA, TIMSS o PIRLS

Algunos estudios han analizado el efecto del tamaño de la clase utilizando las pruebas internacionales estandarizadas que se comentaron en el capítulo 1 de este libro, como PISA, TIMSS o PIRLS. Son análisis que se aproximan algo más a la causalidad que los del recuadro 1 aunque pueden conllevar supuestos de identificación que resultan más restrictivos que los de los de las cuatro secciones anteriores. Emplear datos de evaluaciones internacionales a gran escala como PISA, TIMSS o PIRLS implica potencialmente errores de medición que podrían provo-

car un sesgo a la baja en las estimaciones de los efectos de la reducción del tamaño de la clase. Estos artículos instrumentan la ratio alumnos por aula con el tamaño promedio de clase en todo el curso o etapa educativa, pero este instrumento no es suficientemente fuerte para algunos países o no es aplicable si el centro educativo no tiene más de un grupo por curso (Green e Iversen, 2022). Woessmann y West (2006) utilizan los datos de TIMSS de 18 países y la variación por curso dentro de un mismo centro de manera similar a los artículos iniciados por Hoxby (2000) que usan la variación de la población en modelos de variables instrumentales. Además, incluyen efectos fijos por centro educativo que controlan la selección entre institutos. Estos autores encuentran efectos heterogéneos del tamaño de las clases en diferentes países: en particular, los tamaños de clase más pequeños se asocian con mejores resultados en Islandia y Grecia, mientras que el efecto es básicamente nulo en la mayoría de los demás países.

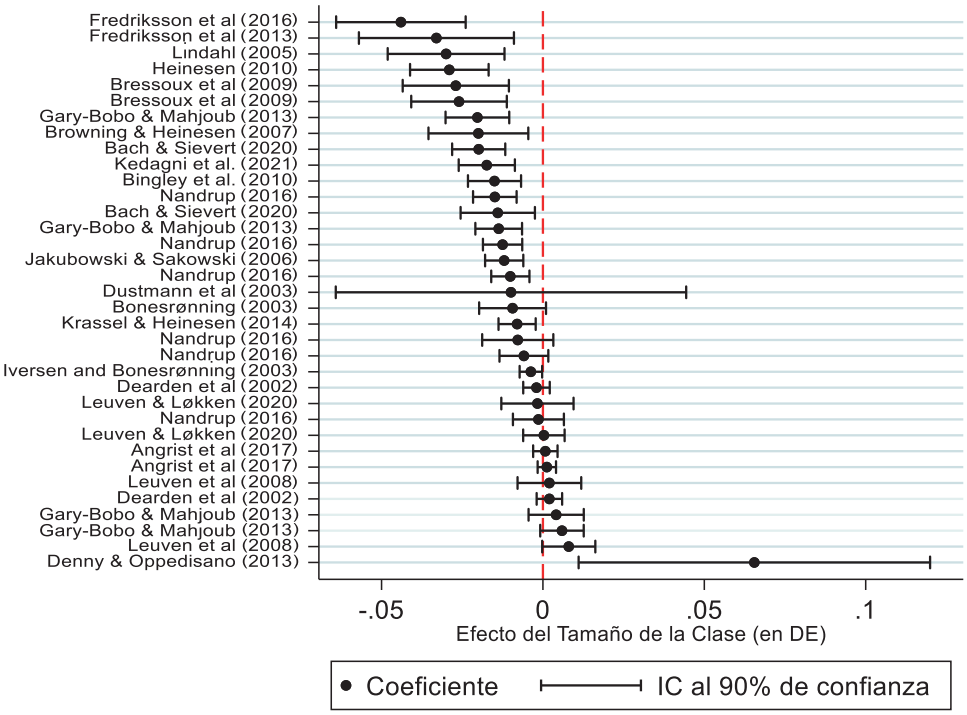
Ejercicio 8. *Analice el efecto del tamaño de la clase en los resultados de Matemáticas y Ciencias que Woessmann y West (2006) encuentran para el caso de España, uno de los 18 países analizados, y compárelo con el impacto que estiman para el resto de los países examinados.*

Altinok y Kingdon (2012) utilizan la variación de los resultados entre Matemáticas y Ciencias en TIMSS de 47 países de un mismo estudiante. En concreto, estos autores relacionan la diferencia en la ratio alumno por aula de un mismo estudiante en su clase de Matemáticas y Ciencias con la diferencia en el rendimiento de ese mismo estudiante entre las dos mismas asignaturas, Matemáticas y Ciencias. Este enfoque de diferencias-en-diferencias evita el problema del emparejamiento no aleatorio de niños con centros educativos y con clases dentro de los institutos. Los resultados muestran un efecto negativo significativo de la ratio alumnos por aula en el rendimiento de los estudiantes de 14 países, pero de magnitud limitada en la mayoría de los casos. Además, Altinok y Kingdon (2012) hallan que los efectos del tamaño de la clase son menores en países con mayor calidad docente. Altinok y Kingdon (2012) Denny y Oppedisano (2013) emplean los datos de PISA de EE. UU. y Reino Unido instrumentando con la variación del tamaño de cursos adyacentes. Estos autores no encuentran un efecto negativo de la ratio alumnos por aula en los resultados de Matemáticas de los estudiantes, e incluso apuntan a un impacto positivo en el caso del Reino Unido. En conjunto, esta literatura demuestra evidencia más mixta y con efectos menos negativos del tamaño de la clase en el aprendizaje de los estudiantes que los de experimentos aleatorios, e incluso que los de regresiones discontinuas o variables instrumentales.

## 4. ESTIMACIONES EMPÍRICAS DEL IMPACTO DE LA REDUCCIÓN DEL TAMAÑO DE LA CLASE. EFECTOS EN EL LARGO PLAZO Y HETEROGENEIDAD

El gráfico 1 muestra el resumen de los resultados encontrados en la literatura sobre la relación entre el tamaño de la clase y el rendimiento educativo. El cuadro está realizado a partir del survey de Leuven y Oosterbeek (2018) y actualizado por Montalbán (2021). Se ha incorporado un artículo nuevo con respecto a Montalbán (2021), el de Kedagni et al. (2021) en European Economic Review. Los artículos incluidos en la revisión cumplen las siguientes características: a) están publicados o en proceso de publicación en una revista internacional que utilice revisión por pares; b) utilizan datos europeos; y c) su diseño de investigación debe ser válido para estimar efectos causales.

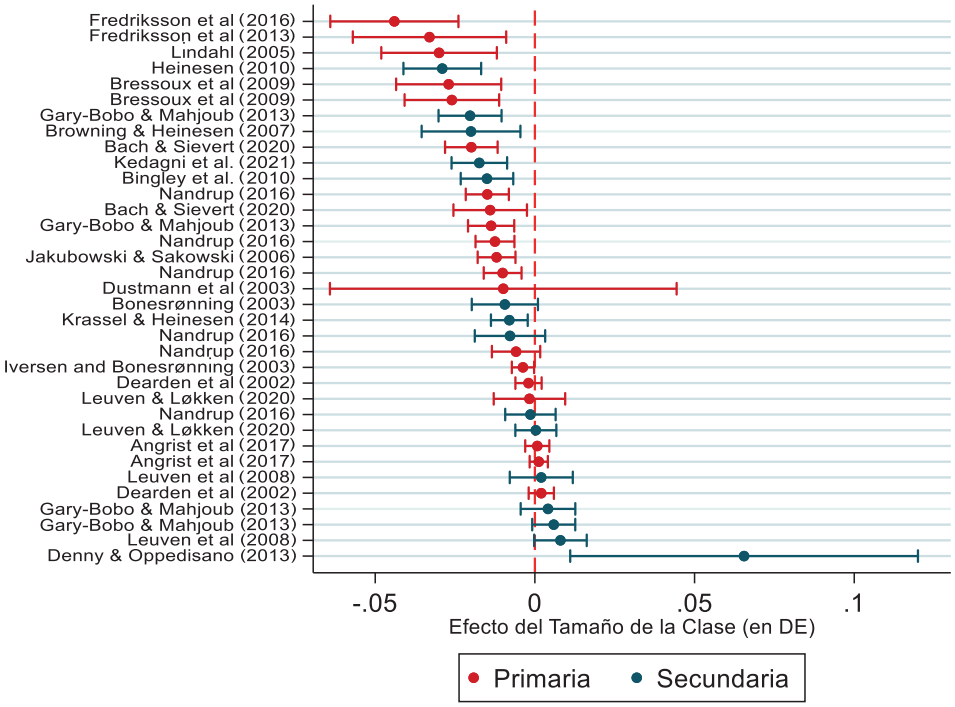
**Gráfico 1. Estimaciones sobre el Efecto del Tamaño de la Clase.**




Nota: Elaboración basada en Leuven y Oosterbeek (2018) y Montalban (2021), con la incorporación de un artículo adicional (Kedagni et al., 2021). DE se refiere a desviaciones estándar. Los coeficientes indican el cambio en el rendimiento académico, medido en desviaciones estándar, causado por el incremento de un alumno por clase. Los coeficientes negativos indican efectos positivos de disminuir el tamaño de la clase. Los puntos indican los coeficientes. Las líneas negras horizontales muestran los intervalos de confianza al 90%.

En conjunto, estas estimaciones del gráfico 1 indican que disminuir el tamaño de la clase aumenta el aprendizaje de los estudiantes. El promedio muestra que, por un aumento en un estudiante en el tamaño de la clase, el aprendizaje se reduce en –0,009 desviaciones estándar (el artículo que hace la mediana tiene un impacto de –0,0099). Sin embargo, el tamaño de los efectos es relativamente pequeño, el equivalente a lo que un aprende un alumno en una semana, y únicamente la mitad de ellos son estadísticamente significativos. El rango de los coeficientes estimados va desde –0,044 de Fredriksson et al. (2016) a 0,065 de Denny y Oppedisano (2013). De 35 coeficientes estimados: 18 son negativos estadísticamente significativos, 8 son negativos no significativos, 8 positivos no significativos y 1 positivo significativo.

**Gráfico 2. Estimaciones sobre el Efecto del Tamaño de la Clase por Nivel Educativo.**




Nota: Elaboración basada en Leuven y Oosterbeek (2018) y Montalban (2021), con la incorporación de un artículo adicional (Kedagni et al., 2021). DE se refiere a desviaciones estándar. Los coeficientes indican el cambio en el rendimiento académico, medido en desviaciones estándar, causado por el incremento de un alumno por clase. Los coeficientes negativos indican efectos positivos de disminuir el tamaño de la clase. Los puntos indican los coeficientes. Las líneas negras horizontales muestran los intervalos de confianza al 90%.

El gráfico 2 muestra que las reducciones del tamaño de la clase en Educación Primaria son ligeramente más efectivas, con un 65% de los artículos encontrando efectos estadísticamente significativos, frente al 53% en Secundaria. El coeficiente que hace la mediana en Primaria es –0,0123 y en Secundaria –0,0078. Estos resultados están en línea con la mayor rentabilidad de intervenciones tempranas (Heckman, 2006) y con los estudios que también concluyen que los efectos son más fuertes y se estiman con mayor precisión en los estudiantes más jóvenes (Iversen y Bonesrønning, 2013; Nandrup, 2016 y Green e Iversen, 2022).

A continuación, comparamos la eficiencia del coste de reducir el tamaño de la clase en comparación con destinar esa inversión a otras medidas educativas. Para ello, la tabla 2 recoge la tabla 11 y la tabla 16 de Handel y Hanushek (2023). En la parte de arriba muestra las estimaciones revisadas por estos dos autores sobre el impacto de la reducción del tamaño de la clase en un estudiante en los exámenes del alumnado. En este caso, y a diferencia del gráfico 1 y 2, el efecto es positivo dado que Handel y Hanushek (2023) analizan una disminución (y no un aumento) en la ratio de alumnos por aula. Las estimaciones presentadas, con un impacto promedio de 0,003 desviaciones estándar (un tercio en términos absolutos del que presentamos nosotros en el gráfico 1) han sido estandarizadas por los dos autores con el objetivo de que las estimaciones sean comparables. En la misma tabla 2 se muestra las estimaciones recogidas y estandarizadas por Handel y Hanushek (2023) sobre el efecto de un aumento del 10% en el gasto en educación en los resultados del alumnado en los exámenes. El impacto promedio de un incremento en el gasto en las puntuaciones de los exámenes es de 0,069 desviaciones estándar. Reducir el tamaño de la clase en un 10%, supondría disminuir como máximo en 2,5 la ratio estudiante por aula (véase tabla 1 de este mismo capítulo), lo que conduciría a un impacto en el aprendizaje (asumiendo linealidad) de cerca de 0,01 desviaciones estándar. Handel y Hanusek (2023) concluyen que el impacto positivo estimado en el aprendizaje de reducir el tamaño de la clase es significativamente menor que el impacto que produce dedicar esos fondos al gasto general en educación. Si en lugar de suponer un efecto promedio de la reducción de la ratio estudiante por aula de 0,003 que alcanzan Handel y Hanushek (2023), utilizásemos el mostrado en el gráfico 1 de 0,009, la comparación seguiría siendo favorable a emplear los fondos públicos en otras medidas educativas diferentes a la disminución del tamaño de clase, pero la diferencia sería mucho menor (0,069 frente a 0,027 desviaciones estándar). Esta conclusión está en la misma línea de Caneiro y Heckman (2003) que señalaban que, aunque las clases más pequeñas aumentan los ingresos de los estudiantes

cuando son adultos, las ganancias salariales que reciben los estudiantes no compensan los costes de contratar profesores adicionales.

**Tabla 2. El efecto de una reducción en el tamaño de la clase en comparación con un aumento del 10% del gasto en educación (extraído de Handel y Hanushek, 2023).**

| | Mediana | Mínimo | Máximo | Número estudios |
|---|---|---|---|---|
| ***Distribución de estimaciones del efecto de la reducción en un estudiante del tamaño de clase*** | | | | |
| Efecto en resultados de pruebas externas | 0,003 | –0,066 | 0,036 | 29 |
| ***Distribución de estimaciones del impacto de un aumento del 10% del gasto en educación*** | | | | |
| Efecto en resultados de pruebas externas | 0,069 | –0,244 | 0,543 | 23 |

Fuente: Tabla 11 y tabla 16 de Handel y Hanushek (2023).

Bettinger et al. (2017) analizan también los efectos de aumentar el tamaño de la clase sobre los resultados y la permanencia de estudiantes universitarios en cursos en línea, una modalidad habitual en muchas instituciones de educación superior. Un experimento cuasi-aleatorio, por el que alumnos en DeVry University fueron asignados a clases de tamaño regular con 31 compañeros o ligeramente más grandes (un 10% más). La conclusión es que pequeños cambios en el número de estudiantes de las clases en línea no conducen a diferencias en el rendimiento de los estudiantes, ni en la permanencia en la Universidad. Bettinger et al. (2017) tampoco encuentran heterogeneidad en los efectos en función de la disciplina académica que los alumnos estudian en la Universidad. El gran tamaño de la muestra analizada por estos autores permite estimaciones precisas que descartan efectos significativos.

La formación que reciben los jóvenes está destinada a tener efectos a lo largo de toda su vida. Por eso es importante contrastar el impacto de las medidas educativas a largo plazo, más allá de su repercusión en los resultados de las pruebas. Podría ser que la reducción del tamaño de clase tenga unos beneficios que son valorados en el mercado laboral. Por el contrario, los efectos del tamaño de la clase podrían tener fecha de caducidad. Como señalamos al principio de este capítulo, Krueger y Whitmore (2001), encontraron que el impacto de la reducción del tamaño de clase se reduce entre la mitad y tres cuartas partes un año después de la intervención, incluso después de mantener la nueva y más limitada ratio de alumnos por clase. Chetty et al. (2011) también encuentran que el efecto del tamaño de la clase en los resultados académicos de los alumnos

desaparece en el corto plazo. Pero este importante trabajo también halla que la menor ratio de alumnos por aula sí afecta significativamente a la probabilidad de llegar a estudios superiores. La evidencia empírica de Chetty et al. (2011) supone una paradoja, pues indica que la disminución de la ratio alumnos por aula no tiene efectos sustanciales en el corto plazo, pero sí en el medio y largo plazo. Una hipótesis que podría explicar este puzle sería la de que estudiar en clases reducidas podría no afectar de forma significativa a las habilidades cognitivas, pero sí ser beneficioso para las habilidades no-cognitivas como la perseverancia, la motivación, el esfuerzo o la determinación. La escasa incidencia de las clases reducidas en la mejora de las habilidades cognitivas explicaría que el impacto del tamaño de las clases en los resultados académicos sea en todo caso modesto y desaparezca con el tiempo. Por el contrario, estudiar en clases con pocos alumnos mejora las habilidades no-cognitivas, un efecto que no se manifiesta en las pruebas externas y estandarizadas pero que tiene un impacto a largo plazo incrementando la probabilidad de llegar a estudios superiores y de gozar de una situación financiera más estable en la vida adulta (Chetty et al., 2011). Chingos (2012) tampoco encontraba un impacto significativo del mandato de reducción del tamaño (CSR) de las clases de Florida, pero sí hallaba evidencia de efectos en resultados no cognitivos, incluyendo una mayor asistencia a clase en Primaria y menos problemas de disciplina en Secundaria. También Fredriksson et al. (2013) encuentran un efecto significativo de las clases reducidas en la adquisición de habilidades no-cognitivas por parte de los alumnos. Una menor ratio de alumnos por aula en los cursos de entre 4º a 6º de Primaria mejora a los 13 años la perseverancia, las aspiraciones en los estudios y la confianza y la seguridad en uno mismo. Los efectos de las clases reducidas en Primaria en las habilidades no-cognitivas permanecen a los 18 años medido en esta ocasión en términos de perseverancia, carácter social, estabilidad emocional, disposición a asumir responsabilidades y llevar a cabo iniciativas. De hecho, Fredriksson et al. (2013) encuentran que el menor tamaño de las clases en los últimos cursos de Primaria incrementa significativamente la probabilidad de finalizar estudios superiores y aumenta los ingresos laborales entre los 27 y los 42 años, particularmente a los individuos que se encuentran en la media de la distribución salarial. En definitiva, los efectos en las habilidades no-cognitivas de los estudiantes parecen ser más consistentes (Sevilla et al., 2020).

Como ya se comentó en el capítulo 4a el Proyecto Preescolar Perry y el Proyecto Abecedarian son dos programas de intervención en Educación Infantil que han seguido a los participantes durante más de medio siglo, mostrando re-

sultados prometedores a lo largo del camino. Se trata de dos intervenciones tempranas en Educación Infantil evaluadas mediante el método de aleatorización con seguimiento a largo plazo. El Proyecto Preescolar Perry (PPP) se llevó a cabo en Michigan, a partir de 1962, proporcionando un programa de educación preescolar de los 3 a los 5 años con hasta 3 horas al día durante todas las semanas del curso, además de orientación a los padres con visitas semanales de los maestros a los hogares. Para ello se contrató docentes con alta cualificación, manteniendo una ratio de menos de 6 niños por maestro. El Proyecto Abecedarian (ABC) se desarrolló en Carolina del Norte, a partir de 1972, desde el nacimiento hasta los 5 años, con un programa intensivo de más de 6,5 horas al día, aplicando un desarrollo curricular especifico, que incluyó tanto actividades educativas como un componente de atención médica y nutricional. La elegibilidad de los dos Programas (PPP y ABC) se basó en indicadores de desventaja socioeconómica de los hogares de los niños, aunque los elementos específicos y los pesos diferían entre las dos intervenciones. Décadas de seguimiento han mostrado resultados alentadores en las habilidades no-cognitivas como el autocontrol o la motivación con efectos positivos a largo plazo en educación, ingresos, salud o menor participación en delitos (Conti et al., 2016).

El efecto directo del tamaño de la clase varía entre distintos tipos de estudiantes dando lugar a efectos heterogéneos (Lazear, 2001; y Albornoz et al., 2018). Green e Iversen (2022) resumen las evidencias de que los efectos del tamaño de la clase son más positivos en estudiantes provenientes de familias de ingresos o niveles educativos bajos, que provienen de hogares disfuncionales, de grupos minoritarios o en alumnos previamente rezagados, aunque en algunos casos las diferencias no son significativas (Leuven y Løkken, 2020). Estos patrones heterogéneos tienen importantes implicaciones políticas. Sería eficiente que los tamaños máximos de clase fueran más reducidos a medida que aumenta la proporción de niños desfavorecidos en un curso (Bach y Sievert, 2020).

Las variaciones en la ratio alumnos por clase, podrían tener complementariedades o ser substitutivos con otros inputs de la educación como el apoyo familiar y la labor del profesorado. Por ejemplo, los padres cuyos hijos ha sido asignados a clases con más estudiantes podrían reaccionar dedicando más apoyo educativo, ayudándoles más con las tareas o facilitarles tutorías externas. El capítulo 3b sobre familia analiza el modelo teórico de Albornoz, Berlinski y Cabrales (2018), en el que el esfuerzo y los resultados de aprendizaje de los estudiantes, el comportamiento de los padres y la escuela, y los recursos públicos dedicados a la educación se determinan conjuntamente en un marco teórico integrado.

Las conclusiones del modelo de Albornoz et al. (2018) predicen que cuando las clases son más grandes, los padres pasan más tiempo ayudando a los estudiantes con las tareas, pero los profesores tienden a poner más responsabilidad en sus estudiantes, una hipótesis confirmada por Fredriksson et al. (2016). Los docentes pueden aprovechar la reducción del tamaño de las clases cambiando sus metodologías para llevar a cabo prácticas docentes más interactivas entre los profesores y los alumnos, y de estos entre sí. La revisión de Green e Iversen (2022) no encuentra que exista una evidencia clara y concluyente de la posible respuesta de los padres ante cambios en el tamaño de la clase. Bonesrønning et al., (2020) si muestran que la disminución de la ratio alumnos por aula conducen a aumentos sustanciales en el esfuerzo del profesor, motivación o satisfacción laboral.

**Recuadro 4. El modelo teórico de Lazear (2001)**

Lazear (2001) presenta un modelo de producción educativa en el que cuando un estudiante interrumpe la clase, el aprendizaje se reduce para los demás estudiantes. El tamaño óptimo de la clase es mayor para los estudiantes con un mejor comportamiento en clase, lo que ayuda a explicar por qué es difícil encontrar efectos relacionados con el tamaño de la clase en los datos. La estructura básica comienza con el reconocimiento de que la educación en un entorno de aula es un bien público. Como con la mayoría de los bienes públicos, el aprendizaje en el aula tiene efectos de congestión, que son externalidades negativas creadas cuando un estudiante impide el aprendizaje de sus compañeros de clase.

El modelo teórico de Lazear (2001) comienza por definir $p$ como la probabilidad de que un estudiante en particular no esté impidiendo su propio aprendizaje o el de otros en cualquier momento. Entonces, la probabilidad de que todos los estudiantes en una clase de tamaño $n$ esté teniendo un buen comportamiento es $p^n$, de modo que la interrupción ocurre $1 - p^n$, del tiempo. Ni el estudiante ni sus compañeros de clase pueden aprender mucho cuando un estudiante tiene un mal comportamiento, lo que hace que el docente destine su tiempo a él. Lazear (2001) señala que $p$ será relativamente alto porque incluso teniendo $p = 0{,}98$ en una clase de 25 estudiantes resultará en interrupciones el 40% del tiempo $1 - (0{,}98)^{25} = 0{,}40$. Para determinar las acciones de los colegios bajo diferentes condiciones, Lazear (2001) comienza preguntándose cuánto pagaría una familia para que un estudiante esté en una clase de tamaño $n$. Se supone que el valor de una unidad de aprendizaje está dado por $V$, que es determinado por el valor de mercado del capital humano

y la probabilidad de que un estudiante esté centrado en el aprendizaje en ese instante dado. Para determinar el tamaño óptimo de la clase, Lazear (2001) considera un colegio con $Z$ estudiantes con $m$ profesores y $m$ clases. Se asume que el coste de un docente y el valor de alquiler del capital asociado para su aula está dado por $W$, que se supone independiente de $p$ y otras condiciones de trabajo. Entonces, un colegio privado que quiere maximizar las ganancias puede vender cada momento en la escuela por $ZVp^n$, con un coste total de $Wm$. La maximización de las ganancias significaría elegir $m$ de manera que se maximice (Lazear, 2001):

$$Beneficio = ZV\,p^{\frac{z}{m}} - Wm \tag{1}$$

$$Beneficio\ por\ estudiante = ZVp^{n-\frac{w}{n}} \tag{2}$$

$$\text{Dado que } n = \frac{z}{m} \tag{3}$$

la condición de primer orden es entonces

$$-V\frac{z^2}{m^2}\,p^{\frac{z}{m}}\ln(p) - W = 0 \tag{4}$$

Y usando (2)

$$Vp^n \ln(p) - \frac{W}{n^2} = 0 \tag{5}$$

Las variaciones en $V$ pueden resultar ya sea porque el mercado para individuos educados en relación con individuos de menos formación cambia, o porque se modifica la cantidad de aprendizaje que tiene lugar durante un momento ininterrumpido de escolaridad (como resultado, por ejemplo, de variaciones en la calidad del docente o las competencias del estudiante).

El equilibrio del mercado se logra de dos maneras, dada la existencia de beneficios positivos. La entrada competitiva de empresas en la industria educativa hace subir $W$ a través de la presión de demanda sobre los salarios en el mercado de docentes. Al mismo tiempo, la oferta de graduados educados al mercado laboral reduce $V$. El equilibrio se logra cuando la maximización de beneficios resulta en cero beneficios totales para el proveedor competitivo de educación.

El tamaño de la clase es una variable de elección y el tamaño óptimo de la clase varía inversamente con la capacidad de atención de los estudiantes. Es eficiente usar menos profesores y una mayor proporción de estudiantes por profesor cuando los estudiantes tienen un mejor comportamiento. Los efec-

tos negativos del tamaño de la clase son mayores para los estudiantes que se comportan peor en clase.

Ejercicio 8. *A partir del modelo básico de la ecuación (2) demuestra la proposición 1 del paper de Lazear (2001, pp. 782 y 799, https://www.jstor.org/stable/2696418). El tamaño óptimo de la clase aumenta con el salario del profesor, disminuye con el valor de una unidad de educación y, lo más importante, aumenta con la probabilidad de que los estudiantes se comporten bien. Es óptimo reducir el tamaño de la clase cuando los estudiantes se comportan peor.*

Ejercicio 9. *Supongamos que la relación de W a* $Vp^n$ *es 5. El tiempo del profesor es cinco veces más valioso que lo que cualquier estudiante obtiene de la clase.* $Z = 100$ *y* $p = 0{,}99$*, lo que significa que el 99 por ciento del tiempo, cualquier estudiante no está causando una interrupción suficiente para que no se produzca aprendizaje en el aula. Utilizando la condición de primer orden (3), calcula el óptimo m y el óptimo n de tamaño de clase.*

*a) Calcula el óptimo m de número de clases y el tamaño de clase n, si* $p = 0{,}98$*, es decir si hay más interrupciones de clase que cuando* $p = 0{,}99$*.*

*b) Calcula el óptimo m de número de clases y el tamaño de clase n, si* $p = 0{,}995$*, es decir si hay menos interrupciones de clase que cuando* $p = 0{,}99$*.*

*c) Calcula la producción educativa por estudiante* ($p^n$)*, con el n óptimo cuando* $p = 0{,}98$*;* $p = 0{,}99$*; y* $p = 0{,}995$*. Comprueba como una clase con menos estudiantes n puede tener una menor producción educativa si el aprovechamiento de la clase p baja lo suficiente.*

*d) Explica por qué la relación de n con p produce que las clases óptimas de Educación Infantil son más pequeñas que las de la Universidad.*

## 5. TUTORÍAS EN PEQUEÑOS GRUPOS

En la sección 3.3 y 3.4 se ha comprobado que la reducción del tamaño de la clase mejora las habilidades cognitivas en el corto plazo y tiene un efecto más a largo plazo en las habilidades no-cognitivas. Pero incluso con las estimaciones más elevados del impacto de la reducción de la ratio alumnos por aula como la de Krueger (1999), la disminución generalizada del tamaño de la clase no es una política muy eficiente dado la enorme inversión que requiere (Caneiro y Heckman, 2003). Bonesrønning et al. (2022) muestran evidencia de que se pueden adoptar enfoques más flexibles y menos costosos como las tutorías en

pequeños grupos para centrarse en algunas asignaturas y brindar refuerzo y apoyo a los alumnos que están rezagados al permitir una instrucción más individualizada por parte de los tutores. Se ha encontrado evidencia empírica causal que demuestra la efectividad de las tutorías en pequeños grupos. En su metaanálisis, Nickow et al. (2020) revisan 96 artículos que examinan los efectos de esta medida a partir de experimentos aleatorios controlados y concluyen que el impacto promedio es significativo y de 0,37 desviaciones estándar, el equivalente a lo que un estudiante aprende en un curso escolar. Los efectos para las intervenciones de Lectura y Matemáticas son similares, aunque la tutoría de Lectura tiende a producir tamaños de efectos más altos en los primeros cursos, mientras que la tutoría de Matemáticas tiene impactos más elevados en los cursos posteriores. Aparte del apoyo académico, estas intervenciones pueden mejorar también el desarrollo socioemocional de los estudiantes y aumentar la asistencia a la escuela. La irrupción de la pandemia de Covid-19 afectó de forma desigual a los niños en situaciones de desventaja (Betthäuser et al., 2023), situando a los programas de tutoría en pequeños grupos en un lugar prominente como una política eficiente para reducir las brechas educativas que se han ampliado durante la pandemia.

La efectividad de las tutorías en pequeños grupos disminuye cuanto más grande es el grupo, con una caída más pronunciada cuando el grupo es de más de 6 alumnos (Burgess, 2020). El impacto será también mayor cuanto más intenso sea el programa. Por ejemplo, el mismo Burgess (2020) indica que una tutoría de media hora al día durante 12 semanas produce un progreso adicional para los estudiantes que es equivalente al aprendizaje de cuatro meses en el centro educativo. Kraft y Falken (2020) señalan los factores que contribuyen al éxito en la configuración de los programas de tutorías en pequeños grupos, que incluyen un proceso riguroso de selección de los tutores, contratación a tiempo completo a lo largo del año académico con el mismo grupo de alumnos, formación continua en mejores prácticas docentes y coordinación con los profesores del colegio o instituto en el que se encuentra el estudiante. Los programas de tutorías son más eficientes si son impartidos por docentes, o en general profesionales de la educación, que en el caso de que estén dirigidos por voluntarios o padres. Las tutorías realizadas por voluntarios no suelen proporcionar buenos resultados educativos porque no funcionan durante todo el curso académico y porque no tienen una formación específica para la enseñanza de alumnos rezagados. El desafío de brindar tutoría de alta calidad en una amplia gama de asignaturas y cursos se vuelve aún más complejo cuando se trata de satisfacer las necesidades de los estudiantes con discapacidades (Kraft, 2015). El resultado final de las tutorías realizadas por

voluntarios es, generalmente, el de bajas tasas de participación de los alumnos y escasa mejora en su rendimiento académico. La falta de coordinación entre tutores y docentes regulares y una alta rotación de los primeros ha llevado también al fracaso de algunas intervenciones. Kraft y Falken (2020) también apuntan a que las tutorías realizadas durante horario escolar tienden a tener mayores impactos que las llevadas a cabo después del colegio. En conclusión, la implementación de programas de tutoría en pequeños grupos bien diseñados y coordinados podría ofrecer oportunidades significativas para mejorar la educación. Sin embargo, es vital para garantizar su éxito que se considere una planificación cuidadosa, selección meticulosa de tutores, con un compromiso continuo por parte de estos y su adaptación a las diversas necesidades de los estudiantes. También Loeb et al. (2023) muestran que el apoyo escolar proporcionadas por las tutorías en pequeños grupos es más significativo cuando los tutores fomentan relaciones positivas con los estudiantes, dirigidas a la instrucción de sus fortalezas y necesidades. Revisando 33 artículos sobre la implementación del apoyo escolar, estos autores llegan a la conclusión de que el entorno y horario del apoyo, el reclutamiento y formación, y el currículo afectan de forma significativa en que las tutorías individuales o de pequeños grupos sean una instrucción de calidad para los estudiantes.

Entre los estudios causales más recientes sobre el efecto de las tutorías en pequeños grupos se encuentra Bonesrønning et al. (2022), que realizan un experimento aleatorio controlado a gran escala en Noruega. Los estudiantes de Primaria beneficiados por asignación aleatoria de tutorías de Matemáticas en pequeños grupos salen de sus clases habituales, de modo que esta iniciativa tiene el efecto adicional de reducir el tamaño del grupo principal en Matemáticas. Bonesrønning et al. (2022) muestran que las clases más pequeñas conducen a aumentos sustanciales en una variedad de medidas de esfuerzo del profesor, motivación y satisfacción laboral. Las estimaciones de *intention-to-treat*, muestran que los estudiantes en los centros educativos tratados aumentaron su rendimiento en 0,06 desviaciones estándar en las pruebas nacionales (equivalente a dos meses de aprendizaje), sin un efecto heterogéneo según el nivel previo, educación de los padres o género.

En España, Gortázar et al. (2023) proporcionan evidencia de un experimento controlado aleatorizado sobre la efectividad de un programa de tutoría online en Matemáticas, dirigido a estudiantes de Secundaria de 12-15 años de zonas desfavorecidas. Se trata de un programa intensivo de ocho semanas durante las horas extraescolares en grupos de dos estudiantes de la misma clase. Las tutorías fueron impartidas por docentes de Secundaria cualificados y remunerados en

colaboración con Empieza por Educar (ExE), la rama española de Teach for All, una ONG especializada en la formación de jóvenes profesores que trabajan en centros a las que asisten estudiantes vulnerables y de bajos ingresos. La intervención aumentó en 0,26 desviaciones estándar las puntuaciones en las pruebas estandarizadas de Matemáticas, y redujo la probabilidad de repetir el año escolar en 9,4 puntos porcentuales, equivalente a una disminución del 78% con respecto al grupo de control, que tenía una tasa de repetición del 12%. Gortázar et al. (2023) proporcionan evidencia de que los efectos positivos del programa persisten un año después. La intervención también mejoró algunos resultados no cognitivos, como las aspiraciones: los estudiantes del grupo de tratamiento tienen 13,6 puntos porcentuales más de probabilidades de declarar que les gustaría continuar los estudios académicos de Bachillerato, equivalente a un aumento del 33,2% en comparación con la media del grupo de control. Los autores no encuentran un impacto en las competencias matemáticas percibidas por los estudiantes, en el gusto por las Matemáticas, locus de control, perseverancia, o el bienestar general.

Gortázar et al. (2023) hallan evidencia a favor de las hipótesis de Kraft y Falken (2020) sobre la importancia de que los tutores estén cualificados y remunerados, puesto que el programa fue completado por un 96,6% de los alumnos del grupo de tratamiento y hubo una rotación muy baja durante el programa. Estos autores llevaron a cabo este experimento aleatorio y controlado varios meses después de que los centros educativos españoles reabrieran completamente después de la pandemia. La evidencia positiva del que encuentran muestra que los programas de tutorías en pequeños grupos pueden ser considerada un complemento en lugar de un sustituto de la enseñanza regular en clase. La intervención tiene un coste aproximado de 300 euros por estudiante, lo que se indica que la relación beneficio-coste es de un aumento de 0,087 desviaciones estándar (0,26/3) por cada 100 euros.

## 6. CONCLUSIONES

La identificación de la relación causal entre el tamaño de la clase y los resultados educativos de los estudiantes a través de correlaciones o estimaciones de mínimos cuadrados ordinarios (OLS) conducen a asociaciones espurias. Los estudios causales muestran que la reducción del tamaño de la clase en un estudiante mejora los resultados de las pruebas externas de los alumnos de un aula en cerca de 0,01 desviaciones estándar, equivalente a lo que se aprende en una semana

de colegio. Las estimaciones del impacto de la disminución de la ratio alumnos por aula registran una gran variedad, con algunos estudios encontrando un impacto nulo y otros un efecto significativo de hasta 0,044 desviaciones estándar. Reducir el tamaño de la clase tiene un impacto positivo en el aprendizaje, pero es un efecto limitado para una medida que tiene un coste presupuestario muy elevado. Como concluía ya Chingos (2011) las reducciones del tamaño de la clase "a gran escala también son extremadamente costosas y suponen una pérdida de oportunidades para realizar inversiones educativas más inteligentes" (p. 2). Otras medidas educativas podrían tener una mejor relación coste-beneficio que la reducción de la ratio alumnos por clase. La mediana de los estudios revisados en Handel y Hanusheck (2023) muestra que un aumento del 10% del gasto en otras políticas educativas conduce a una mejora del aprendizaje 7 veces más elevada que el efecto de una reducción también en el 10% del tamaño de la clase.

Una alternativa puede ser llevar a cabo reducciones del tamaño de la clase de forma más dirigida y concentrada en los primeros niveles educativos, Infantil y Primaria. Como muestra el capítulo 4a de Educación Infantil de este mismo libro, la evidencia empírica coincide en mostrar la conveniencia de llevar a cabo intervenciones tempranas de alta calidad. Muchos buenos hábitos se crean muy temprano. El impacto positivo de la reducción del tamaño de clase parece concentrarse, dentro de la Educación Primaria, en la asignatura de Matemáticas (Handel y Hanushek, 2023).

La reducción en el tamaño de la clase beneficia en mayor medida a los estudiantes de familias desfavorecidas o de bajo nivel educativo, que vienen de entornos familiares desestructurados, que forman parte de grupos minoritarios o rezagados en su rendimiento académico (Green e Iversen, 2022). En consecuencia, es más eficiente concentrar las disminuciones de la ratio alumnos por clase en centros educativos ubicados en zonas desfavorecidas o con mayor proporción de alumnos con dificultades de aprendizaje, pues estos estudiantes se pueden beneficiar más de una mayor atención de los profesores. Esta conclusión refuerza la idea de que la capacidad del docente para controlar y gestionar el aula se dificulta cuando los estudiantes son disruptivos o necesitan apoyo adicional (Lazear, 2001), de modo que los efectos de la reducción del tamaño de la clase son más perceptibles en entornos con un mayor número de estudiantes con problemas de seguimiento de las clases o de comportamiento. Políticas públicas como tutorías en pequeños grupos y recuperación escolar registran un coste-beneficio muy superior a la reducción generalizada del tamaño de la clase.

En cuanto a futuras líneas de investigación, se podría explorar la interacción entre el tamaño de la clase y las prácticas docentes, el ambiente en el aula y las expectativas, incentivos, restricciones y comportamientos de estudiantes, padres y profesores. El estudio del impacto de la ratio alumnos por aula se ha concentrado en los enfoques convencionales de instrucción. Sin embargo, en la era actual, la introducción de tecnologías modernas e innovaciones en la pedagogía en los centros educativos ha llevado a cambios sustanciales. Finalmente, el estudio del efecto de la reducción del tamaño de la clase podría enfocarse en áreas particulares o procesos cognitivos donde se identifiquen más obstáculos en el aprendizaje (Connolly y Haeck, 2022).

## BIBLIOGRAFÍA


Albornoz, F., Berlinski, S. y Cabrales, A. (2018). Motivation, resources and the organization of the school system. *Journal of the European Economic Association*, 16 (1), 199-231. https://doi.org/10.1093/jeea/jvx001

Altinok, N. y Kingdon, G. (2012). New evidence on class size effects: a pupil fixed effects approach. *Oxford Bulletin of Economics and Statistics, 74*(2), 203-234. https://doi.org/10.1111/j.1468-0084.2011.00648.x

Andersen, S.C., Beuchert, L., Skyt Nielsen, H. y Thomsen, M.K. (2020). The effect of teacher's aides in the classroom: evidence from a randomized trial. *Journal of the European Economic Association, 18*(1), 469-505 https://doi.org/10.1093/jeea/jvy048

Angrist, J.D., Battistin, E. y Vuri, D. (2017). In a small moment: class size and moral hazard in the Italian Mezzogiorno. *American Economics Journal: Applied Economics, 9*(4), 216-249. DOI: 10.1257/app.20160267.

Angrist, J.D. y Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics, 114*(2), 533-575. https://doi.org/10.1162/003355399556061

Angrist, J.D., Lavy, V., Leder-Luis, J. y Shany, A. (2019). Maimonides' rule redux. *American Economic Review: Insights, 1*(3), 309-324. DOI: 10.1257/aeri.20180120

Angrist, J. y Pischke, J. (2015). *Dominar la Econometría*. Antoni Bosch, Barcelona.

Argaw, B.A. y Puhani, P.A. (2018). Does class size matter for school tracking outcomes after elementary school? Quasi-experimental evidence using ad-

ministrative panel data from Germany. *Economics of Education Review, 65*, 48-57. https://doi.org/10.1016/j.econedurev.2018.06.005

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M. y Walton, M. (2017). From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. *Journal of Economic Perspectives, 31*(4), 73-102. DOI: 10.1257/jep.31.4.73

Bach, M. y Sievert, S. (2020). Birth Cohort Size Variation and the Estimation of Class Size Effects. *ZEW - Centre for European Economic Research Discussion Paper*. 20-053. http://dx.doi.org/10.2139/ssrn.3719762

Betthäuser B.A., Bach-Mortensen A.M. y Engzell P. (2023*).* A systematic review and meta-analysis of the evidence on learning during the Covid-19 pandemic. *Nature Human Behaviour, 7*(3), 375-385. 10.1038/s41562-022-01506-4.

Bettinger, E., Doss, C., Loeb, S., Rogers, A. y Taylor, E. (2017). The Effects of Class Size in Online College Courses: Experimental Evidence. *Economics of Education Review, 58*, 68-85. https://doi.org/10.1016/j.econedurev.2017.03.006

Bingley, P., Jensen, V.M. y Walker, I. (2010). *The effect of class size on education and earnings: Evidence from Denmark.* Unpublished Working Paper. https://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/bingley.pdf

Bonesrønning, H. (2003). Class size effects on student achievement in Norway: Patterns and explanations. *Southern Economic Journal, 69*(4), 952-965. https://doi.org/10.1002/j.2325-8012.2003.tb00542.x

Bonesrønning, H., Green, C., Iversen, J.M.V. y Nyhus, O.H. (2020). *How do teachers respond to changes in class size? Evidence from a randomised trial.* Norwegian University of Science and Technology. (mimeo).

Bonesrønning, H., Finseraas, H., Hardoy, I., Iversen, J.M.V., Nyhus, O.H., Opheim, V., Salvanes, K.V., Sandsør, A.M.J. y Schøne, P. (2022). Small-group instruction to improve student performance in mathematics in early grades: Results from a randomized field experiment. *Journal of Public Economics, 216*, 104765 https://doi.org/10.1016/j.jpubeco.2022.104765.

Bressoux, P., Kramarz, F. y Prost, C. (2009). Teachers training, class size and students' outcomes: Learning from administrative forecasting mistakes. *The*

*Economic Journal, 119*(536), C540-C561. https://doi.org/10.1111/j.1468-0297.2008.02247.x

Browning, M. y Heinesen, E. (2007). Class size, teacher hours and educational attainment. *Scandinavian Journal of Economics, 109*(2), 415-438. https://doi.org/10.1111/j.1467-9442.2007.00492.x

Burgess, S. (2020). *How we should deal with the lockdown learning loss in England's schools.* Voxeu. https://cepr.org/voxeu/columns/how-we-should-deal-lockdown-learning-loss-englands-schools (accedido 16 Octubre 2023).

Cabrales, A. (2010). *El tamaño de la clase y el experimento STAR a largo plazo. Nada es Gratis*. https://nadaesgratis.es/cabrales/el-tamano-de-la-clase-y-el-experimento-star-a-largo-plazo (accedido 16 octubre 2023).

Carneiro P. y Heckman J.J. (2003). Human Capital Policy, en Heckman J., Krueger A. (eds), *Inequality in America: What Role for Human Capital Policy?* MIT Press, Cambridge, pp. 77-240.

https://www.nber.org/system/files/working_papers/w9495/w9495.pdf

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W. y Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics, 126*(4), 1593–1660. https://doi.org/10.1093/qje/qjr041

Chingos, M.M. (2011). *The false promise of class-size reduction*. Center for American Progress. (accedido 16 Octubre 2023) https://files.eric.ed.gov/fulltext/ED536071.pdf

Chingos, M.M. (2012). The impact of a universal class-size reduction policy: evidence from Florida's statewide mandate. *Economics of Education Review, 31*(5), 543-562. https://doi.org/10.1016/j.econedurev.2012.03.002

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, 31*(3), 77-95. https://doi.org/10.1016/j.econedurev.2012.01.004

Cohen-Zada, D., Gradstein, M. y Reuven, E. (2013). Allocation of students in public schools: Theory and new evidence. *Economics of Education Review, 34*(C), 96-106.

Connolly, M. y Haeck, C. (2022). Nonlinear Class Size Effects on Cognitive and Noncognitive Development of Young Children. *Journal of Labor Economics, 40*(S1), 341-382. https://doi.org/10.1086/718328

Conti, G., Heckman, J.H. y Pinto, R. (2016). The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *Economic Journal, 126*(596), 28-65 https://doi.org/10.1111/ecoj.12420

Dearden, L., Ferri, J. y Meghir, C. (2002). The effect of school quality on educational attainment and wages. *Review of Economics and Statistics, 84*(1), 1-20. https://doi.org/10.1162/003465302317331883

Denny, K. y Oppedisano, V. (2013). The surprising effect of larger class sizes: evidence using two identification strategies. *Labour Economics, 23*, 57-65. https://doi.org/10.1016/j.labeco.2013.04.004

Duflo, E., Dupas, P. y Kremer, M. (2015). School governance, teacher incentives, and teacher-pupil-ratios: experimental evidence from Kenyan primary schools. *Journal of Public Economics, 123*, 92-110. https://doi.org/10.1016/j.jpubeco.2014.11.008

Dustmann, C., Rajah, N. y van Soest, A. (2003). Class size, education and wages. *Economic Journa,. 113*, F99-F120. https://doi.org/10.1111/1468-0297.00101

European Commission, 2023. *National Education Systems.* Eurydice, Luxemburgo. https://eurydice.eacea.ec.europa.eu/national-education-systems

Fredriksson, P., Öckert, B. y Oosterbeek, H. (2013*).* Long-term effects of class size. *Quarterly Journal of Economics, 128*(1), 249-285. https://www.jstor.org/stable/26372498

Fredriksson, P., Öckert, B. y Oosterbeek, H. (2016). Parental responses to public investments in children: Evidence from a maximum class size rule. *Journal of Human Resources, 51*(4), 832-868. DOI: 10.3368/jhr.51.4.1114-6779R1

Fryer, R.G. (2017). The Production of Human Capital in Developed Countries: Evidence From 196 Randomized Field Experiments, en Banerjee, A.V. y Duflo, E. (eds), *Handbook of Economic Field Experiments*, Volume 2, North-Holland, Amsterdam, pp. 95-322.

https://scholar.harvard.edu/sites/scholar.harvard.edu/files/fryer/files/handbook_fryer_03.25.2016.pdf

Gary-Bobo, R.J. y Mahjoub, M.B. (2013). Estimation of class-size effects, using Maimonides' rule and other instruments: The case of French junior high

schools. *Annals of Economics and Statistics,* 111-112, 193-225. https://doi.org/10.2307/23646331

Gilraine, M., Macartney, H. y McMillan, R. (2018). *Education reform in general equilibrium: evidence from California's class size reduction*. National Bureau of Economic Research. Report No. w24191 https://www.nber.org/system/files/working_papers/w24191/revisions/w24191.rev0.pdf

Gilraine, M., (2020). A Method for Disentangling Multiple Treatments from a Regression Discontinuity Design. *Journal of Labor Economics, 38*(4), 1267-1311. DOI: 10.1086/706740

Gortázar, L., Hupkau, C. y Roldán, T. (2023). Online tutoring works: Experimental evidence from a program with vulnerable children. ESADE Working Paper. https://www.esade.edu/ecpol/es/publicaciones/online-tutoring-works-experimental-evidence-from-a-program-with-vulnerable-children/

Graves, J. y Zucchelli, E. (2022). *Consecuencias no intencionadas de las medidas de política pública: ejemplos de sanidad y educación*. Papeles de Economía Española. 172, 30-42. https://www.funcas.es/articulos/consecuencias-no-intencionadas-de-las-medidas-de-politica-publica-ejemplos-de-sanidad-y-educacion/

Green, C. y Iversen, J.M.V. (2022). The effect of class size in schools on student outcomes, en Mendolia, S., O'Brien, M., Paloyo, A.R. y Yerokhin, O. (eds.). *Critical Perspectives on Economics of Education.* Routledge, Abingdon, pp. 166-181 https://doi.org/10.1080/02188791.2023.2224634

Han, J. y Ryu, K. (2017*). Effects of class size reduction in upper grades: evidence from Seoul, Korea. Economics of Education Review, 60*, 68-85. https://doi.org/10.1016/j.econedurev.2017.07.004

Handel, D.V. y Hanushek, E.A. (2023). U.S. School Finance: Resources and Outcomes, en Machin, S.J., Woessmann, L. Hanushek, E.A. (eds.), *Handbook of the Economics of Education.* Volume 7. North Holland, Amsterdam (en prensa).

Hanushek, E.A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis. 19*(2), 141-164. https://doi.org/10.3102/01623737019002141

Hanushek, E.A., Schwerdt, G., Wiederhold, S. y Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review, 73*, 103-130 https://doi.org/10.1016/j.euroecorev.2014.10.006

Hanushek, E. y Woessmann L. (2011). The Economics of International Differences in Educational Achievement, en Machin, S. y Woessmann, L. (eds.), *Handbook of the Economics of Education,* Volume 3. North Holland, Amsterdam, pp. 89-200. https://www.sciencedirect.com/handbook/handbook-of-the-economics-of-education/vol/3/suppl/C

Hanushek, E. y Woessmann, L. (2020). The economic impacts of learning losses. *OECD Education Working Papers, 225*. OECD Publishing, Paris, https://doi.org/10.1787/21908d74-en.

Heckman, J.J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science, 312*(5782), 1900-1902. https://doi.org/10.1126/science.1128898

Heinesen, E. (2010). Estimating class-size effects using within-school variation in subject-specific classes. *Economic Journal, 120*(545), 737-760. https://doi.org/10.1111/j.1468-0297.2009.02301.x

Hojo, M. (2013). Class-size effects in Japanese schools: A spline regression approach. *Economics Letters, 120*(3), 583-587. https://doi.org/10.1016/j.econlet.2013.06.027

Hoxby, C.M. (2000). *The effects of class size on student achievement: new evidence from population variation. Quarterly Journal of Economics, 115*(4), 1239-1285. https://doi.org/10.1162/003355300555060

Iversen, J.M. y Bonesrønning, H. (2013). Disadvantaged students in the early grades: Will smaller classes help them? *Education Economics, 21,* 305-324 https://doi.org/10.1080/09645292.2011.623380

Jackson, C.K., Johnson R.C. y Persico, C. (2016). The Efects of School Spending on Educational and Economic Outcomes: Evidence From School Finance Reforms. *Quarterly Journal of Economics, 131*, 157-218. https://doi.org/10.1093/qje/qjv036

Jakubowski, M. y Sakowski, P. (2006). Quasi-experimental estimates of class size effect in primary schools in Poland. *International Journal of Educational Research, 45,* 202-215. https://doi.org/10.1016/j.ijer.2006.11.003

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, 44*(1), 223-250. DOI: https://doi.org/10.3368/jhr.44.1.223

Kedagni, D., Krishna, K., Megalokonomou, R. y Zhao, Y. (2021). Does class size matter? How, and at what cost? *European Economic Review, 133*(C), 103664 https://doi.org/10.1016/j.euroecorev.2021.103664

Krassel, K.F. y Heinesen, E. (2014). Class-size effects in secondary school. *Education Economics, 22*(4), 412-426 https://doi.org/10.1080/09645292.2014.902428

Krueger, A.B. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics, 114*(2), 497-532. https://doi.org/10.1162/003355399556052

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. *The Economic Journal, 111*(468), 1-28. https://doi.org/10.1111/1468-0297.00586

Lazear, E.E. (2001). *Educational Production. Quarterly Journal of Economics*. 116(3), 777-803. https://doi.org/10.1162/00335530152466232

Leuven, E., Oosterbeek, H. y Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *Scandinavian Journal of Economics, 110*(4), 663-693. https://doi.org/10.1111/j.1467-9442.2008.00556.x

Leuven, E. y Oosterbeek, H. (2018). Class size and student outcomes in Europe. *European Expert Network on Economics of Education Analytical Report, 33*. Publications Office of the European Union, Luxemburgo https://doi.org/10.2766/59709

Leuven, E. y Løkken, S.A. (2020). Long-term impacts of class size in compulsory school. *Journal of Human Resources, 55*(1), 309-348. https://doi.org/10.3368/jhr.55.2.0217.8574R2

Lindahl, M. (2005). Home versus school learning: A new approach to estimating the effect of class size on achievement. *Scandinavian Journal of Economics, 107*(2), 375- 394. https://doi.org/10.1111/j.1467-9442.2005.00413.x

MacKinnon, J.G., Nielsen, M.O. y Webb, M.D. (2023). Testing for the appropriate level of clustering in linear regression models. *Journal of Econometrics, 235*(2), 2027-2056. https://doi.org/10.1016/j.jeconom.2023.03.005.

Montalbán, J. (2021). *¿Es realmente efectiva la política educativa más popular?. Nada es Gratis.* https://nadaesgratis.es/jose-montalban/es-realmente-efectiva-la-politica-educativa-mas-popular (accedido 16 octubre 2023)

Nandrup, A.B. (2016). Do class size effects differ across grades? *Education Economics, 24*(1), 83-95. https://doi.org/10.1080/09645292.2015.1099616

OCDE (2012). *Education at a Glance 2012: OECD Indicators*, OCDE, Paris. https://doi.org/10.1787/eag-2012-en

OCDE (2014). *PISA 2012 Results: Students and Money (Volume VI): Financial Literacy Skills for the 21st Century*. OCDE, Paris. https://dx.doi.org/10.1787/9789264208094-en .

OCDE (2020a) *Education at a Glance 2020: OECD Indicators*. OCDE, Paris. https://doi.org/10.1787/69096873-en

OCDE (2020b). *PISA 2018 Results (Volume I): What Students Know and Can Do, Annex A9. A note about Spain in PISA 2018: Further analysis of Spain's data by testing date* (updated on 23 July 2020). PISA, OCDE, Paris. https://www.oecd.org/pisa/PISA2018-AnnexA9-Spain.pdf

OCDE (2020c). *PISA 2018 Results (Volume V): Effective Policies, Successful Schools.* OCDE, Paris, https://doi.org/10.1787/ca768d40-en.

OCDE (2021). *Education at a Glance 2021: OECD Indicators*. OCDE, Paris. https://doi.org/10.1787/b35a14e5-en

OCDE (2022). *Education at a Glance 2022: OECD Indicators*. OCDE, Paris. https://doi.org/10.1787/3197152b-en

Sevilla, A., Sainz, J. y Sanz, I. (2020). *La reducción en un 7,7% de la ratio alumnos por clase podría compensar parte del impacto negativo de la Covid-19 en la educación: Maimónides nueve siglos después (II). Nada es Gratis*. https://nadaesgratis.es/admin/la-reduccion-en-un-77-de-la-ratio-alumnos-por-clase-podria-compensar-parte-del-impacto-negativo-de-la-covid-19-en-la-educacion-maimonides-nueve-siglos-despues-ii (accedido 16 octubre 2023)

Urquiola, M. y Verhoogen, E. (2009). Class-Size Caps, Sorting and the Regression Discontinuity Design. *American Economic Review, 99*(1),179–215. DOI: 10.1257/aer.99.1.179

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, 50*(3), 695-736. https://doi.org/10.1016/j.euroecorev.2004.11.005

# Capítulo 3b
# EL IMPACTO DE LA FAMILIA EN LA EDUCACIÓN

Antonio Cabrales
Departamento de Economía
Universidad Carlos III de Madrid
ORCID: https://orcid.org/0000-0003-0209-0399

**Resumen**

Este capítulo estudia otro de los factores que impacta en el desempeño estudiantil: las familias. Pero intentaremos a ir más allá de la mera observación de que el estatus socioeconómico es importante para los resultados. Su participación activa es decisiva. De paso, se introduce la metodología de teoría de juegos como herramienta para entender interacciones estratégicas en entornos educativos, algo que puede ser de utilidad en otras secciones del libro.

## 1. INTRODUCCIÓN

El objetivo de este capítulo es doble. Por un lado, se pretende continuar con el estudio de los diferentes factores que afectan al rendimiento de los estudiantes y del sistema educativo. En particular, vamos a examinar el efecto de las familias en este sistema. Por otro lado, queremos introducir a los estudiantes a los métodos de modelización de teoría de juegos en problemas educativos. Estos métodos pueden resultar útiles y aplicables más allá de este capítulo.

Las familias son claramente cruciales en los resultados del sistema educativo. Por un lado, los padres proveen de la herencia genética y una parte importante del entorno en el que los estudiantes crecen. Y esto tiene una influencia directa, aunque "pasiva" en el desempeño de los estudiantes. En este capítulo queremos concentrarnos sobre todo en otro tipo de influencia más "activa", en los padres como proveedores de incentivos y otros elementos para mejorar los resultados de sus hijos.

En cuanto a la modelización, en la educación operan de manera simultánea muchos tipos de agentes. Los estudiantes, sus familias, los profesores, los gestores de los centros educativos, los ciudadanos que disfrutan, y a menudo finan-

cian, el sistema, y sus representantes políticos que son elegidos en parte para hacerlo funcionar. Esta complejidad de decisores que interactúan entre sí, y cuyos intereses no siempre están perfectamente alineados es la que hace apropiado el uso de la teoría de juegos para comprender cómo se toman esas decisiones. El capítulo no pretende ser exhaustivo, y preferimos dar unos pocos ejemplos de contribuciones de gran relevancia con cierto detalle y profundidad, antes que un simple listado de contribuciones académicas. En primer lugar, describiremos alguna evidencia sobre el impacto del esfuerzo parental en los resultados educativos de sus hijos en la sección 2 y en la sección 3 un artículo que describe el impacto de una intervención concreta para mejorar la involucración parental en la educación.

Luego pasamos a describir en detalle tres modelos que describen y explican la implicación parental en la educación. En la sección 4 explicamos el modelo de Doepke y Zilibotti (2017) sobre por qué existen y cómo se deciden los distintos estilos parentales de educación. Este modelo es importante porque explica los motivos económicos que inciden en la elección por parte de los padres de diferentes maneras de enfocar su esfuerzo educativo. Después, en la sección 5 estudiamos el modelo de Bisin y Verdier (2000, 2001) que, en contraste con el anterior, se enfoca más bien en la transmisión de valores "culturales" en educación. Esto sirve para señalar que los economistas somos bien conscientes de que la educación es algo más que "formación de capital humano" y que los temas culturales son bien importantes y generan después retroalimentaciones cruciales para la economía. Finalmente, en la sección 6 utilizamos el modelo de Albornoz et al. (2018) para aclarar cómo las interacciones de múltiples decisores –familias, estudiantes, escuelas, y autoridades políticas– son cruciales para entender el desempeño de los estudiantes y la asignación de recursos y esfuerzos en la educación.

Finalmente, la sección 7 recoge otros estudios, que no describimos en detalle, pero que pueden ser útiles para que el estudiante de este volumen continúe su exploración de este tema.

## 2. EVIDENCIA SOBRE LOS PADRES Y SU INFLUENCIA EN LA EDUCACIÓN

En esta sección discutimos en detalle algunos artículos representativos sobre la influencia de los padres en la educación. En la sección 7 proporcionamos evidencia adicional sobre este impacto de manera más resumida.

Una primera referencia relevante es un artículo de De Philippis y Rossi (2021) que utiliza datos de PISA. La pregunta que se hacen los autores es qué parte de la diferencia de nivel educativo entre países puede ser atribuida a las características o acciones de los padres. Una primera aproximación se puede obtener de una regresión estándar que intenta explicar el rendimiento individual del estudiante a partir de características del alumno, la familia, y el país. Cuando se hace esto, las variables que se observan de los padres (educación, profesión, si están empleados, número de libros en casa, y si la lengua que se habla en el hogar no es la oficial), explican una proporción modesta de las diferencias entre países (entre un 13% y un 25% de la varianza una vez que se controla por la calidad de la escuela, con variables observables o efectos fijos). El problema de esta aproximación es que centrarlo todo en el estatus socioeconómico (que es básicamente lo que miden esas variables) es reduccionista. Sabemos que hay otras cosas que hacen los padres. Por ejemplo, aquí vamos a hablar de los estilos parentales (Doepke y Zilibotti, 2017) y también de que la transmisión de valores culturales es importante (Bisin y Verdier, 2000, 2001).

Y por esto De Philippis y Rossi (2021) utilizan otra estrategia diferente. Lo que hacen los autores es comparar dentro de un mismo país estudiantes hijos de nativos y de inmigrantes con distintas nacionalidades, para un mismo nivel socioeconómico. Obviamente esto no es perfecto, porque los emigrantes pueden estar seleccionados sistemáticamente, y porque puede haber dificultades de asimilación para los inmigrantes que varían entre países de origen. En el artículo se relacionan las notas medias de PISA por países de un estudiante que es hijo de inmigrantes del país A con las medias de los estudiantes del país A. Se puede observar una correlación positiva y cercana a 1. En otras palabras, si las notas medias de PISA del país de origen son un porcentaje superior a la media, en el país de destino los hijos de inmigrantes de ese país mejoran la media de ese país en el mismo porcentaje. El hecho de que estas diferencias entre países se preserven a pesar de que los estudiantes en el país de acogida disfrutan de condiciones materiales, tanto de las escuelas como institucionales, similares a las de los nativos sugiere que la influencia parental es importante.

El análisis de regresión muestra que esta observación es robusta a la inclusión de muchos controles. Y el efecto no es debido fundamentalmente a las madres del este de Asia, porque persiste cuando se excluyen los estudiantes de estas nacionalidades. El estudio tiene muchos controles de robustez de los resultados, incluyendo la utilización de datos del censo americano, que tienen más observa-

ciones sobre características familiares. Una preocupación seria es que los padres de los inmigrantes estén seleccionados (positiva o negativamente, aunque la investigación sobre inmigración reciente, como Albornoz y et al. (2018) sugiere una selección positiva), pero dado que lo que importa son las comparaciones relativas entre grupos de inmigrantes, esto solamente es importante si hay una selección diferente entre países. Una forma de ver que esto no es un problema se puede comprobar en la figura 3 del artículo. En el se asocia el nivel de educación relativo de los emigrantes respecto a los nativos de su país de origen con la nota media del país de acogida y se puede observar que no hay ninguna correlación.

Aunque estos resultados de De Philippis y Rossi (2021) son interesantes, aún nos queda por entender cuáles son los mecanismos por los cuales los padres afectan a la instrucción de los hijos. Y para esto vamos a otro artículo, de Chen, Chung y Wang (2023). La pregunta en este caso es cómo afecta la presencia de hijos de padres con poder (en este caso altos funcionarios del partido comunista chino) al desempeño escolar de los compañeros de estos chicos.

Los autores utilizan una encuesta, el China Educational Panel Survey. Y aprovechan una característica importante del sistema escolar chino. La asignación de estudiantes a clases, cuando hay más de una clase por curso, es aleatoria. Por tanto, el número de hijos de altos funcionarios por clase es, dentro de la escuela, aleatoria. Y como utilizan efectos fijos de escuela y curso, están aislando el efecto de recibir algún hijo de cuadros de partido más en el entorno para el que esta asignación es aleatoria.

El primer resultado relevante es que tener a un hijo de alto funcionario más en la clase aumenta el rendimiento de sus compañeros en un 2%. En cambio, el tener a un padre alto funcionario no mejora el rendimiento de su propio hijo. Esto no quiere decir que los padres no influyan sobre sus hijos. En escuelas donde se aprecia interferencia de los padres en la asignación de clase (vía una comprobación de los informes de asignación computerizados y la asignación efectiva) y que se extraen de la muestra general, sí que hay un efecto importante de los padres funcionarios en sus propios hijos. Esta última observación sugiere que la influencia de los padres funcionarios en sus hijos surge por la influencia en el entorno social (mejores recursos a la clase, o mejores redes sociales), y no por aumentar su capital humano directamente.

Pero lo más interesante es de dónde surge el efecto del hijo de funcionarios en sus compañeros. El efecto de estos chicos es prácticamente el mismo si se

controlan por sus características personales, en particular sus propios resultados escolares. Recuerden que los funcionarios no tienen efecto directo en sus hijos. Esto, unido a que el impacto de los hijos de funcionarios es independiente de los resultados escolares de estos, indica que hay que explicar ese impacto de otra manera.

Una explicación razonable es que la presencia de estos chicos induce a los padres de los compañeros a aumentar su esfuerzo con estos. La razón más probable es que los padres de los compañeros quieren beneficiarse del contacto con el funcionario a través de sus hijos. Y en China un buen desempeño escolar del hijo es una buena forma de que los compañeros, y sobre todo sus padres, lo tomen en cuenta. Una evidencia a favor de esta hipótesis (que se puede obtener por la riqueza de datos de la encuesta) es que cuando aumentan los hijos de funcionarios en la clase, los padres de los compañeros se involucran más en las amistades que eligen sus hijos, aumentan su comunicación con ellos sobre lo que pasa en la escuela, y ayudan más en las tareas de los niños.

No es fácil saber si estos resultados serán aplicables fuera de China, en el sentido de la influencia de los padres de los compañeros en que los estudiantes se esfuercen. Lo que es claro es que conseguir la implicación de los padres en la escuela de sus hijos es un factor crucial en que estos mejoren sus resultados, y medidas de política que hagan más probable esta implicación podrían conseguir este objetivo. Esto es lo que pasamos a estudiar en la siguiente sección 3.

## 3. INTERVENCIONES DE POLÍTICA PARA AUMENTAR LA IMPLICACIÓN PARENTAL

La sección anterior 2 contiene alguna evidencia acerca del efecto de los padres en la educación de sus hijos con su intervención activa, más allá de proporcionar un entorno afectivo, y material más o menos potente. En esta sección explicamos una intervención representativa que se ha propuesto recientemente en la literatura para mejorar la implicación de los padres en la educación. En la sección 7 discutimos brevemente más evidencia sobre intervenciones.

Un artículo de Barrera Osorio y coautores (2020) describe dos ensayos aleatorios (ver capítulo 1) realizados por el gobierno de México relacionados con el programa "Apoyo a la Gestión Escolar (AGE)". Los ensayos aleatorios se realizaron en cuatro estados (Chiapas, Guerrero, Puebla y Yucatán) con una población indígena importante, que ha sufrido una discriminación histórica.

La primera intervención proporciona financiación a las asociaciones de madres y padres de alumnos (AMPAs). La beca estándar es de entre 500 y 700 dólares anuales en función del tamaño de la escuela y se puede utilizar para infraestructuras, material o actividades. No se puede usar para salarios a los directores o profesores. El tratamiento dobla esta beca hasta una cantidad entre 1.000 y 1.400 dólares. Esto cubre aproximadamente el 83% de los costes "extra" que tienen que aportar los padres a la educación más allá de lo que aporta el estado. El experimento se realizó en aproximadamente 250 escuelas. El experimento comienza en 2007, cuando se recogen los datos para la línea de base y continúa en 2008, 2009 y 2010 cuando se recogen los datos sobre los resultados

La otra intervención proporciona información a los padres sobre cómo implicarse más en la escuela y cómo ayudar a los hijos en la misma. La parte fundamental de la intervención consiste en cinco sesiones conducidas por un asesor comunitario que facilita la discusión. La primera sesión introduce el programa AGE, y explica la importancia de la implicación parental. La segunda explica a los padres cómo implicarse en la escuela y ayudar en la toma de decisiones. La tercera informa sobre recursos de educación y salud en su comunidad. La cuarta explica los hitos en el desarrollo de los niños y adolescentes e introduce actividades adaptadas a la edad para apoyar el aprendizaje. La quinta anima a los padres a crear un plan de acción sobre cómo mejorar la educación de sus hijos. Esta intervención se realizó en aproximadamente 180 escuelas. La intervención comienza en 2009 cuando se recogen los datos para la línea de base y continúa hasta 2010 cuando se recogen los resultados.

Los investigadores contaron con datos de exámenes estandarizados de los participantes en el experimento, así como de cuestionarios de estudiantes (de tercero a quinto de Primaria), padres y profesores. En el primer experimento, la partida de gastos más grande fue la relacionada con material escolar (libros, pizarras, material para escribir, 38% en el tercer año de intervención) y material sanitario (de primeros auxilios, 16% en el tercer año). Las reparaciones y mejoras de infraestructuras fueron ganando importancia en el tiempo, pasando del 35% en el primer año al 44% en el tercer año. El tratamiento informativo se realizó en el 95% del grupo de tratamiento.

En términos de resultados, comenzamos por la implicación parental. Para el primer tratamiento, no parece que hubiera un impacto significativo en la implicación en actividades escolares, o en reuniones con los profesores. En cambio, hubo un aumento en la participación en los órganos de decisión escolares (votaciones, reu-

niones) de alrededor de un 18%. Pero ese efecto desaparece en los años siguientes. En cambio, la intervención informativa aumentó la participación en actividades en un 16,9%. Este aumento se produjo fundamentalmente entre padres que ya eran miembros de la asociación porque no cambió el número de miembros de estas.

Seguimos viendo resultados con el comportamiento de los padres en casa. Para el primer experimento, no se observan cambios. Ni en el conocimiento de las tareas escolares, ni en la ayuda con sus tareas. Tampoco los maestros cambian, ni en términos de sus ausencias de clase, ni en hacer la educación más personalizada. Para el experimento de información, sí que hay cambios significativos. Los padres tienen 5,7% más de probabilidades de estar al tanto de las tareas y 8,8% más de probabilidad de ayudar con las tareas de sus hijos. Es decir, esta información no solo aumentó la participación dentro de las escuelas, sino que también mejoró el comportamiento de los padres fuera de las escuelas. Una observación importante es que estos efectos fueron mucho más potentes en los padres indígenas. Por ejemplo, el porcentaje de estos que ayuda con las tareas sube del 20,5% al 31,9% lo que representa un aumento del 50%.

Finalmente llegamos a los resultados escolares propiamente dichos. Para el primer experimento, de manera quizá no sorprendente vistos los anteriores resultados, no se encuentra ninguna diferencia en resultados escolares de ningún tipo. Para la intervención informativa, se observa una disminución del 1,2% en el abandono escolar y una disminución del 6% en las acciones disciplinarias (algo que tiene más impacto en familias de menor nivel socioeconómico). Sin embargo, no parece haber diferencias en las notas escolares. Aunque los autores son ligeramente negativos sobre esta falta de impacto en las notas, es posible argumentar que los resultados son de hecho positivos. Sabemos por la literatura sobre intervenciones escolares tempranas descrita en el capítulo 4a, que a veces una intervención que afecta a las habilidades no cognitivas no tiene efectos inmediatos en las notas, pero tiene un impacto importante en la probabilidad de graduarse, tener un buen trabajo y una vida plena. Los resultados aquí apuntan en esa dirección también. Fíjense que aquellos resultados del Perry Preschool y Abecedarian a los que aludimos en el capítulo 4a justamente reducen el abandono, y que un mecanismo era el mejor comportamiento de los escolares, como se observa aquí también.

Para acabar, los autores se fijan en el impacto de las intervenciones en la confianza de los padres en los profesores. Aquí observan una mejora importante de

la segunda intervención en esa confianza, lo que probablemente es el mecanismo para la mejora de la implicación en casa.

Aunque los resultados en sí no son espectaculares, tampoco son intervenciones muy caras. Así que para mí el mensaje es optimista. Se pueden hacer intervenciones de bajo coste que convencen a algunos padres de mejorar su atención a la educación.

## 4. ESTILOS PARENTALES Y EDUCACIÓN: DOEPKE Y ZILIBOTTI, 2017

Esta sección presenta una versión simplificada del modelo de Doepke y Zilibotti (2017), al que en adelante llamaremos DZ17, para comprender cómo interactúan los padres y sus hijos en la toma de decisiones sobre el esfuerzo que los estudiantes dedican a la educación[1]. La perspectiva original de DZ17 es que las familias quieren tener influencia en la dedicación de sus hijos, porque estos no tienen una perspectiva del futuro suficientemente formada. Por ese motivo, los estudiantes pueden tomar decisiones que no contemplen de manera suficientemente profunda el impacto en el largo plazo de la adquisición de capital humano.

El modelo es muy útil para entender las "fricciones" que hacen que los de deseos de padres e hijos no coincidan, y las estrategias que usan los padres para guiar a sus hijos en la dirección que prefieren. Al mismo tiempo el modelo ilumina sobre cómo las condiciones del entorno económico hacen más o menos probable que se utilice una estrategia u otra. DZ17 describen cuatro estilos parentales, tomados de la literatura psicológica, a los que vamos a llamar: autoritario[2], persuasivo, permisivo, y abandonados. El estilo autoritario influencia de manera directa las acciones que puede tomar el estudiante, el persuasivo intenta afectar a las preferencias del estudiante, el permisivo y el abandonado simplemente no intervienen en esas decisiones, aunque se diferencian en que a los permisivos les importan los estudiantes y sus decisiones, y los abandonados simplemente no tienen interés en lo que les pase. En DZ17 no se explora el estilo abandonado

---

[1] La simplicación se adapta de una versión del modelo derivada en Cabrales y Hauk (2023).

[2] En el artículo original, los autores utilizan la palabra inglesa "authoritative" que puede traducirse en español por "autoritativo" que está reconocida en el diccionario de la RAE, pero nos parece que es un vocablo poco conocido y que no transmite bien el concepto.

porque para su forma de modelizar se confunde con el permisivo, pero proponen extensiones donde la diferencia entre permisivo y abandonado es importante.

Vamos a ver primero cómo toman decisiones los estudiantes, y después veremos lo que influye en las decisiones de los padres. Finalmente describiremos alguna evidencia consistente con los resultados del modelo.

### 4.1. La decisión del estudiante

El tiempo total del estudiante se normaliza a una unidad. El estudiante tiene que decidir la proporción de ese tiempo total dedicado a esfuerzo educativo. Vamos a denotar esa proporción como la variable $x$. Cada unidad de tiempo empleado en educación tiene un retorno futuro acumulado que denotamos $R$. Por tanto, el esfuerzo educativo tiene un rendimiento total de $Rx$. El resto del tiempo $1 - x$, se dedica a otras actividades, que denominamos de manera genérica como ocio. Cada unidad de tiempo dedicado a ocio tiene un rendimiento $aK$, que depende de un parámetro de preferencia $a$ y de un retorno marginal $K$, y por tanto el tiempo total dedicado a ocio rinde $aK(1 - x)$. La utilidad total del estudiante es simplemente la suma de los dos retornos, el del esfuerzo educativo y el del ocio, y se puede expresar como:

$$U_E(x) = aK(1 - x) + Rx \tag{1}$$

Esto se puede rescribir como:

$$U_E(x) = aK + (R - aK)x$$

Nótese que esta función de utilidad es lineal, y por tanto el valor de la variable $x$ que maximiza (1) es:

$$x = \begin{cases} 1 \text{ si } R > aK \\ 0 \text{ si } R \leq aK \end{cases} \tag{2}$$

Esta solución de esquina es típica de una función lineal, y es muy poco realista. Como es natural, los estudiantes no se pasan todo el tiempo estudiando o dedicados al ocio. Pero esta especificación hace que las derivaciones que surgen a continuación sean muy sencillas, y revelan algunos factores importantes en las decisiones, que es para lo que sirven los modelos teóricos. En concreto, como se puede observar, si el retorno de la educación $R$ es muy elevado comparado con el "disfrute" derivado del ocio $aK$, los estudiantes pasarán mucho tiempo estu-

diando, y lo contrario sucede si $R$ es bajo. El siguiente ejercicio ayudará al lector a deducir lo que pasa para funciones de utilidad más "realistas".

Ejercicio 1. *¿Cuál sería la decisión del estudiante si la función de utilidad fuera?:*

$$U_E(x) = aK(1 - x) + Rx^{\alpha} \tag{3}$$

*donde* $0 \leq \alpha \leq 1$.

### 4.2. La elección entre persuasión y permisividad

Los padres persuasivos (los que deciden cambiar la $a$ del estudiante) son más frecuentes, como veremos en la sección 4.4, en lugares donde la renta se reparte de manera más desigual. Estos padres son conscientes de que el esfuerzo educativo tiene unos rendimientos muy elevados en la vida adulta comparado con no hacer ese esfuerzo. Una forma de modelar esta forma de ver el mundo es suponer que estos padres tienen una función de utilidad como la que ahora presentamos, donde $\mu \in (0, 1)$ representa el peso que la familia da al disfrute del ocio del hijo. Es decir, estas familias no valoran tanto el ocio del chico como él mismo, lo cual es equivalente a decir que se dan cuenta del gran valor futuro de la educación.

$$U_P(a) = \mu aK(1 - x) + Rx \tag{4}$$

Estos padres tienen un poder/estrategia importante. Pueden influenciar el valor de $a$ que tienen los estudiantes. Concretamente, pueden escoger cualquier valor de $a \in [\underline{a}, \overline{a}]$. Esto es lo que quiere decir que puedan ser persuasivos. Cuanto más amplio sea este intervalo, mayor es el poder de persuasión potencial de los padres. Y a la inversa si el intervalo es pequeño. Nosotros vamos a interpretar la elección de $a$ como una elección entre persuasividad y permisividad. Si los padres deciden influenciar la $a$, son persuasivos. En caso contrario son permisivos.

Utilizando la respuesta óptima del estudiante (2) la utilidad de la familia se puede escribir como:

$$U_P(a) = \begin{cases} R \text{ si } \dfrac{R}{K} \geq a \\ \mu aK \text{ si } \dfrac{R}{K} < a \end{cases} \tag{5}$$

Vamos a examinar un poco más en detalle esta expresión, teniendo en cuenta los límites a la elección parental (teniendo en cuenta que los padres "persuasivos" son los que deciden influenciar la $a$, y los "permisivos" los que no lo hacen).

1. Si $\frac{R}{K} \geq \bar{a}$ los estudiantes utilizarán $x = 1$ para todos los valores de $a$. En este caso los padres no cambian la elección de los estudiantes para ningún valor de su variable de elección.

2. Si $\frac{R}{K} < \underline{a}$ Los chicos siempre utilizan $x = 0$, así que los padres no pueden cambiar tampoco la elección de los hijos.

3. Si, por el contrario, $\underline{a} < \frac{R}{K} < \bar{a}$ la elección de a de los padres sí que puede determinar si el estudiante escoge $x = 1$ o $x = 0$.

Ejercicio 2. *Demostrar que las siguientes afirmaciones son ciertas:*

*1. La familia prefiere $x = 1$ si $\bar{a} \leq \frac{R}{\mu K}$. Por tanto, la condición para $x = 1$ es $\frac{R}{K} \leq \bar{a} \leq \frac{R}{\mu K}$.*

2. *La familia prefiere que los chicos usen $x = 0$ si $\bar{a} > \frac{R}{\mu K}$.*

Dado lo que se muestra en el ejercicio es fácil ver que la respuesta óptima de la familia en términos de su elección de $a$ es la siguiente:

$$(1)\ a = \text{mín}\left[\bar{a} < \frac{R}{K}\right] \text{cuando máx}\{\underline{a}, \bar{a}\mu\} < \frac{R}{K} \text{ lo que implica } x = 1 \qquad (6)$$

$$(2)\ a = \bar{a} \text{ para } \frac{R}{K} < \text{máx}\,(\underline{a}, \bar{a}\mu\} \text{ lo que implica } x = 0 \qquad (7)$$

Vamos ahora a interpretar estos resultados. Cuando $R/K$, el rendimiento relativo de la inversión en capital humano respecto al valor intrínseco del ocio es alto, antes de ser afectado por la influencia parental, la familia desea ejercer más influencia. Esto hace que la familia quiera utilizar su persuasión. Este es el tipo de resultado que permite explicar como veremos en la sección 4.4 por qué en países desiguales, en los que el premio al Grado universitario es mayor, hay más padres persuasivos y menos permisivos.

Algo similar sucede para diferentes valores relativos otorgados por los padres al ocio estudiantil. Si $\mu$ es elevado (el ocio es considerado importante), los padres deciden no intervenir en $a$, y se convierten en permisivos. Si $\mu$ es bajo, decidirán intervenir, y se convierten en persuasivos. Esto predice que en un entorno donde la creatividad es importante, la permisividad aumentará.

### 4.3. La elección entre autoritarismo y permisividad

Los padres autoritarios desean escoger directamente el valor del esfuerzo que el estudiante utiliza. Es decir, no usan la *a* como variable de influencia, sino directamente la *x*. Esto es posible en un entorno donde los estudiantes pueden ser vigilados de manera directa por los padres y pueden supervisar el proceso de aprendizaje. Por tanto, esto es más razonable que ocurra en las etapas educativas iniciales, pero ciertamente no en las más avanzadas, donde justamente al no ser obligatorias, las decisiones son más cruciales. Por otro lado, el autoritarismo, como veremos, es más razonable cuando la capacidad de persuasión parental es baja. Esto es más frecuente en entornos donde los padres tienen un nivel de educación menor. Como consecuencia de esto entienden de manera más confusa el mundo laboral al que se enfrentarán los estudiantes, y por tanto no son capaces de transmitir la importancia de la educación.

Veamos ahora cómo modelizar estas elecciones de los padres. Suponemos que la familia puede escoger un límite inferior a *x*, $(\underline{x})$. Pero esta supervisión tiene un coste, que suponemos cuadrático. Para que este tipo de influencia sea óptimo, tiene que ser cierto que la persuasión no funcione, y controlar las preferencias es difícil, lo que como hemos visto sucede cuando:

$$\frac{R}{K} < \text{máx}\,\{\underline{a}, \bar{a}\mu\}$$

En ese caso podemos escribir:

$$U_p(\bar{a}, \underline{x}) = \mu\bar{a}K\,(1 - \underline{x}) + R\underline{x} - \frac{b}{2}\underline{x}^2 \qquad (8)$$

Ejercicio 3. *Si la utilidad de los padres es como se representa en 8, la elección óptima de los padres es:*

$$\underline{x} = \frac{R - \mu\bar{a}K}{b}$$

Observen que $\underline{x}$ es más alto, y por tanto el grado de control parental es mayor, cuando $R - \mu\bar{a}K$ es más alto. Es decir, cuando la diferencia entre el rendimiento profesional y el valor del ocio estudiantil percibido por la familia es grande, esta quiere intervenir. Esto explicaría también la relación positiva que encuentran DZ17 entre autoritarismo y desigualdad y que revisamos en 4.4. La razón por la que esa relación no es tan fuerte como con la persuasividad tiene que ver con otros factores que se adivinan aquí. Nótese por ejemplo que si la tecnología de

control parental es buena ($\underline{a}$ bajo) o el coste de control directo parental es alto ($b$ alto), no se quiere intervenir tanto en la $x$. Es decir, b es alto en países más avanzados donde los estudiantes tienen un alto grado de libertad. En esos mismos países, los padres tienen un alto poder de persuasividad, o sea que $\underline{a}$ es bajo, los padres pueden convencer a sus hijos de que vale la pena esforzarse por los retornos futuros, y consiguen que estos den un menor valor al ocio.

### 4.4. Evidencia sobre estilos culturales y relación con los entornos económicos

Una primera evidencia importante tiene que ver con el uso del tiempo de los padres en diversos países. Doepke, Zilibotti y Sorrenti (2019) documentan usando la encuesta del Estudio Multinacional de Uso del Tiempo para todos los países para los que hay suficientes datos, como Canadá, Italia, Países Bajos, España, Reino Unido y Estados Unidos, que los padres han aumentado el tiempo que pasan con sus hijos en las últimas décadas, en las que al mismo tiempo ha crecido la desigualdad dentro de los países. En un par de países, Estados Unidos y los Países Bajos, para los que la serie de datos es más larga, se observa que los padres con educación universitaria han aumentado mucho el tiempo dedicado a la educación de los hijos respecto a los que tienen menores niveles de educación. De no haber prácticamente diferencias en 1970, a una diferencia de más de tres horas por semana actualmente.

Para identificar los tipos de estilos parentales en los datos, DZ17 utilizan el *World Values Survey*, que tiene preguntas sobre los valores que los padres piensan que son importantes en el comportamiento de los hijos. Estos autores piensan que una pregunta que identifica bien el estilo autoritario es que los padres piensen que un valor importante es la obediencia. El estilo persuasivo lo identifican con padres que piensan que lo importante es el trabajo duro. El estilo permisivo se identifica con padres que piensan que la imaginación es importante. En la siguiente tabla observamos la comparación para varios países de algunos de estos datos.

Lo relevante de estos datos es que varían mucho entre países, y DZ17 muestra que hay correlaciones muy interesantes, por ejemplo, con la desigualdad. El estilo permisivo es mucho más frecuente en países con poca desigualdad, y mucho menos en países más desiguales. El estilo persuasivo tiene el patrón contrario, es muy frecuente en países desiguales y mucho menos frecuente en los más igualitarios. El estilo autoritario tiene un patrón parecido, aunque menos fuerte, al del estilo persuasivo.

**Tabla 1: World Values Survey Wave (ola) 7, 2017-2022. Preguntas: Q17, Q9, Q11.**

| Valor/País | Francia | Alemania | Italia | España | Suecia | UK | US |
|---|---|---|---|---|---|---|---|
| Obediencia (autoritario) | 25,9 | 11,8 | 23,3 | 40,4 | 7,3 | 22,2 | 20,5 |
| Trabajo duro (persuasivo) | 41,3 | 40,4 | 29 | 23,7 | 8,2 | 47,3 | 67,9 |
| Imaginación (permisivo) | 21 | 23,8 | 17,9 | 36 | 43,8 | 35,5 | 29,8 |

De hecho, en la tabla 2 de Doepke, Zilibotti y Sorrenti (2019), reproducida a continuación se muestran asociaciones entre variables a nivel individual, no ya sólo de país. La tabla presenta regresiones logísticas donde la variable dependiente es un indicador que recibe valor 1 si el estilo de crianza es intensivo (persuasivo o autoritario). Los coeficientes mostrados son ratios de probabilidades ("odds-ratios") y por tanto un número mayor que uno quiere decir que la variable aumenta la probabilidad del estilo parental intensivo, y un número menor que uno indica que la variable disminuye la probabilidad de ese estilo. Todos los modelos incluyen efectos fijos de ola de la encuesta. Los modelos en las columnas 2, 3, 5 y 6 incluyen controles por género, edad, edad al cuadrado y logaritmo del PIB per cápita (basado en PIB real en paridad de poder adquisitivo encadenado, de las Penn World Table 9.0). Los modelos en las columnas 3 y 6 también incluyen efectos fijos de país. La desigualdad se define como la relación entre el percentil 90 y el 10 de los ingresos brutos de las personas con empleo por cuenta

**Tabla 2. Desigualdad, redistribución y estilos parentales.**

| | (1) Logit Intensivo | (2) Logit Intensivo | (3) Logit Intensivo | (4) Logit Intensivo | (5) Logit Intensivo | (6) Logit Intensivo |
|---|---|---|---|---|---|---|
| **Desigualdad** | 2,38*** | 2,50*** | 2,12** | 1,74*** | 1,74*** | 27,22** |
| | (0,44) | (0,29) | (0,72) | (0,37) | (0,28) | (35,21) |
| **Progresividad fiscal** | | | | 0,20** | 0,24** | 5,35 |
| | | | | (0,13) | (0,17) | (5,88) |
| **Gastos sociales** | | | | 0,70 | 0,58 | 0,21** |
| | | | | (0,29) | (0,25) | (0,14) |
| **Controles** | No | Sí | No | Sí | No | Sí |
| **Efecto fijo país** | No | No | Sí | No | No | Sí |
| **Observaciones** | 45.482 | 45.482 | 45.482 | 32.196 | 32.196 | 32.196 |

Los errores estándar (entre paréntesis) tienen cluster a nivel de país.
*, **, *** indican significación estadística al 10%, 5% y 1%, respectivamente.

ajena a tiempo completo. La progresividad fiscal está tomada de Andrew Young School Policy Studies (2010). Los gastos sociales se expresan como porcentaje del PIB. Como se puede observar, la desigualdad es un factor muy robusto a favor de los estilos parentales más intensivos, el persuasivo y autoritario. Y, por contra, la progresividad fiscal o los gastos sociales disminuyen la probabilidad de estos estilos parentales.

## 5. TRANSMISIÓN CULTURAL: BISIN Y VERDIER, 2000 Y 2001

Los economistas tendemos a hablar de la educación por su importancia en la formación de un factor de producción, el trabajo. Es decir, pensamos en la educación como input en la creación de "capital humano". Pero también somos conscientes de que la educación es mucho más, porque tiene influencia en la creación de valores, preferencias y normas sociales. Se puede decir, estirando la metáfora, que además de "capital humano" influye en la creación de "capital cultural".

Y esto nos lleva a un potencial conflicto entre las preferencias de los padres y las preferencias sociales. Por poner un ejemplo concreto, alrededor de un 20% de los adultos americanos rechazan la teoría de la evolución, en el sentido de que piensan que los humanos hemos existido siempre en nuestra forma actual. Este porcentaje llega a un 38% entre los protestantes evangélicos (Masei, 2023). Estos padres no están muy felices de que se enseñe la teoría de la evolución por selección natural en una escuela. Y sus representantes políticos han intentado en muchos estados prohibir esta enseñanza en las escuelas, aunque por suerte los tribunales lo han impedido.

Los artículos de Bisin y Verdier (2000, 2001) inauguran una literatura muy interesante en la que se intenta modelizar esta tensión entre los deseos parentales y los de la sociedad en su conjunto. El modelo está relacionado con una literatura de antropología que estudia la transmisión cultural en modelos dinámicos de evolución de preferencias que comenzaron Cavalli-Sforza y Feldman (1981) y Boyd y Richerson (1988).

Vamos a presentar primero una versión muy simple del modelo, adaptada de Cabrales y Hauk (2023) donde además conectamos con el modelo de DZ17 de la sección anterior y después en la sección 5.4, discutimos la evidencia consistente con el modelo.

### 5.1. Culturas de capital humano y evolución cultural

Supongamos que existen dos "culturas del capital humano" como las de la sección 4.2 en la que un grupo da más peso que el otro al retorno futuro de la educación y por tanto $\mu_1 < \mu_2$. Las proporciones de la población que tienen cada una de estas culturas son $q_1$ y $q_2$. La socialización de los estudiantes combina una dimensión "vertical," por parte de la familia y "horizontal," por parte del resto de la sociedad. Esto quiere decir que la probabilidad de que el estudiante "herede" el valor cultural de su familia es $d_i$, y con probabilidad $1 - d_i$ será como el resto de la sociedad. Por tanto, las probabilidades de cada tipo de estudiante son:

$$P_{ii} = d_i + (1 - d_i)\, q_i$$
$$P_{ij} = (1 - d_i)\,(1 - q_i)$$

### 5.2. La decisión del estudiante

La utilidad del estudiante con estilo cultural $i$ es:

$$U_i^E\ (x) = \mu_i K\,(1 - x) + Rx$$

Por tanto, la decisión óptima sobre cuánto estudiar es:

$$x_i = \begin{cases} 1 \text{ si } R \geq \mu_i K \\ 0 \text{ si } R < \mu_i K \end{cases}$$

Para que el problema sea interesante, vamos a suponer que los dos tipos de estudiantes tienen elecciones diferentes. Esto quiere decir que suponemos que:

$$\mu_2 K > R > \mu_1 K$$

y por tanto,

$$U_1^E\,(x_1^*) = R;\ U_2^E\ (x_2^*) = \mu_2 K$$

### 5.3. Decisiones de la familia

Las familias tienen "empatía imperfecta," es decir, que no internalizan completamente la utilidad del estudiante. Por tanto, una familia de estilo cultural $i$ con un estudiante de estilo $j$ tiene una utilidad:

$$U_i^P\,(x_j^*) = \mu_i K\,(1 - x_j^*) + Rx_j^*$$

Si $i \neq j$ la utilidad de los padres es menor que si los hijos tuvieran el mismo estilo cultural.

Es decir, $U_i^P\ (x_j^*) < U_i^P\ \ (x_i^*)$. Para ver que esto es cierto, observen que:

$$U_1^P\ (x_2^*) = \mu_1 K < R = U_1^P\ \ (x_1^*)$$

y

$$U_2^P\ (x_1^*) = R < \mu_2 K = U_2^P\ (x_2^*)$$

En estas condiciones, parece claro que la familia querría invertir en cambiar el tipo de su hijo. Por ejemplo, si aumentar el $d_i$ (la probabilidad de que el estudiante sea como sus padres) tuviera un coste cuadrático, la utilidad total de la familia de tipo $i$ sería:

$$\begin{aligned} U_i^P &= P_{ii} U_i^E\ (x_i^*) + P_{ij} U_i^E\ (x_j^*) - \frac{b}{2} d_i^2 \\ &= \left(di + (1 - d_i)\, q_i\right) U_i^E\ (x_i^*) + (1 - d_i)\,(1 - q_i)\, U_i^E\ \ (x_j^*) - \frac{b}{2} d_i^2 \end{aligned} \tag{9}$$

Ejercicio 4. *Demostrar que si la utilidad de los padres es como se representa en 9 la elección óptima de los padres es:*

$$d_i^* = \frac{(1 - q_1)}{b} \left(U_i^E\ (x_i^*) - U_i^E\ \ (x_j^*)\right)$$

$$d_i^* = \frac{(1 - q_i)}{b} (R - \mu_1 K),\ d_2^* = \frac{(1 - q_2)}{b} (\mu_2 K - R)$$

Como puede observarse, la inversión en inculturación[3] ($d_i$) depende (inversamente) del tamaño del grupo, y de la rentabilidad extra percibida por la familia de que el estudiante esté en su grupo.

Ejercicio 5. *Discutir la intuición económica de este resultado y las implicaciones empíricas del mismo.*

[3] No es una palabra frecuente, pero es muy adecuada aquí. El diccionario de la RAE lo define como "Proceso de integración de un individuo o grupo en la cultura y en la sociedad con las que entra en contacto".

### 5.4. Evidencia sobre transmisión cultural

Aunque hay mucha literatura sobre la importancia de factores culturales en cuestiones económicas, la literatura sobre transmisión cultural es más escasa. En sociología, Ozorak (1989) y Hayes y Pittelkow (1993) documentan la importancia de la familia en la transmisión de creencias religiosas. En economía Patacchini y Zenou (2004) encuentran que en el Reino Unido la familia es importante para la transmisión de actitudes respecto a la educación para los estudiantes con padres con un nivel educativo alto, mientras que la comunidad es crucial para las actitudes de los estudiantes de familias con un nivel educativo bajo. Por su parte Cohen-Zada, D. (2006) encuentran que en Estados Unidos los padres religiosos deciden enviar a sus hijos a escuelas religiosas o privadas para protegerlos de las influencias externas y así preservar su identidad religiosa como sugiere el modelo que acabamos de presentar. Fernández et al. (2004) muestran que la presencia de hombres criados en una familia en la que la madre trabajaba ha sido un factor significativo en el aumento de la participación laboral femenina a lo largo del tiempo, mostrando así la importancia de la transmisión cultural de actitudes respecto a esta participación. Finalmente, Bisin et al. (2004), utilizando los datos de la encuesta GSS (General Social Survey) para los Estados Unidos durante el periodo 1972-1996, estiman los parámetros estructurales del modelo de Bisin y Verdier 2000 aplicado a la socialización del matrimonio y sus hijos con una orientación religiosa en los Estados Unidos. Las tasas observadas de matrimonios mixtos y socialización son consistentes con el hecho de que protestantes, católicos y judíos tienen una fuerte preferencia porque sus hijos se identifican con sus propias creencias religiosas y que toman decisiones costosas para influir en las creencias religiosas de sus hijos.

## 6. FAMILIAS Y RECURSOS: ALBORNOZ et al., 2018

La motivación para este modelo nace de que la implementación y evaluación de políticas educativas a menudo ignora que esas reformas pueden generar cambios de comportamiento en los actores involucrados en el proceso educativo. Por ejemplo, cuando se discute el papel de los recursos escolares en el aprendizaje de los estudiantes, no se suele tener en cuenta cómo responderán los padres y los administradores escolares. Si aumentan los recursos docentes, los padres podrían disminuir su involucración en la escuela. Un caso extremo es el de los "boarding schools", las escuelas secundarias residenciales, donde la involucración de los padres es nula. O podrían aumentarla. De manera similar, las propuestas de bonos

escolares generalmente ignoran su influencia en la participación de los padres en la educación.

Para mejorar esta discusión Albornoz et al. (2018), al que llamaremos ABC18, estudian un modelo de educación en el que el esfuerzo y los resultados de aprendizaje de los estudiantes, el comportamiento de los padres y la escuela, y los recursos públicos dedicados a la educación se determinan conjuntamente en un marco teórico integrado. Después, en la sección 6.4, discutimos alguna evidencia relevante para el modelo.

### 6.1. Utilidad de los estudiantes, familias y profesores

El capital humano del estudiante *i*, $H_i$, es una función lineal de su esfuerzo, $e_i$, y de un parámetro de productividad, $v_i$:

$$H_i = v_i e_i \tag{10}$$

Los estudiantes internalizan solo de manera imperfecta el efecto de su esfuerzo en el capital humano. Los padres y profesores dedican algo de su tiempo a inducir a los estudiantes a realizar un esfuerzo costoso utilizando diferentes herramientas de motivación. En particular, la utilidad a corto plazo de los estudiantes está dada por:

$$U_{Si} = c_{1i} e_i + c_{2j} e_i - \frac{1}{2} e_i^2 \tag{11}$$

El esfuerzo $e_i$ tiene un coste cuadrático. La participación de padres y profesores entra positivamente en la utilidad del niño, $c_{1i} e_i + c_{2j} e_i$ con el subíndice *j* denotando la escuela a la que asiste el niño *i*. La variable $c_{1i}$ ($c_{2j}$) expresa la potencia de los incentivos de los padres (profesores) para cada unidad de esfuerzo del niño. Por lo tanto, $c_{1i} e_i + c_{2j} e_i$ es la recompensa total del niño por su esfuerzo. En otras palabras, el papel de los padres (profesores) es ayudar a los niños a internalizar el valor de $e_i$ (y en última instancia $H_i$) en su función de utilidad. La utilidad de la familia depende del capital humano del hijo y de su propio bienestar:

$$U_{Pi} = H_i + W_i$$

El bienestar de los padres $W_i$ depende del tiempo dedicado al trabajo o a la realización de actividades de ocio. Este es el tiempo total ($T$) disponible menos el tiempo pasado con el niño como consecuencia del esquema de recompensas ($t_i$). La recompensa efectiva ($c_{1i} e_i$) depende del tiempo ($t_i$) y la eficiencia de los

padres ($v_{Pi}$) para generar una recompensa efectiva determinada, de modo que $c_{1i}e_i = 2t_i v_P$[4]. El parámetro $v_{Pi}$ captura el hecho de que algunos padres son mejores para motivar que otros y, en consecuencia, generan una mayor recompensa por cualquier cantidad de tiempo dedicada a sus hijos. Es decir, asociamos la productividad de la participación de los padres con algún atributo innato, como el talento o la capacidad de persuasión. Implícitamente, estamos imponiendo que el tiempo dedicado a inducir el esfuerzo, por una recompensa efectiva determinada, decrece en la capacidad del padre para generarla. Por tanto, si $\psi_i$ es el coste de oportunidad de la familia *i*:

$$W_i = (T - t_i)\,\psi_i = \left(T - \frac{c_{1i}e_i}{2v_{Pi}}\right)\psi_i$$

y por tanto,

$$U_{Pi} = v_i\, e_i + \left(T - \frac{c_{1i}e_i}{2v_{Pi}}\right)\psi_i \tag{12}$$

La utilidad de los maestros es muy paralela a la de los padres, excepto porque tienen más estudiantes a su cargo. Depende del capital humano medio que obtienen sus estudiantes y su propio bienestar. Este bienestar viene determinado por la diferencia entre el tiempo para el maestro y el tiempo efectivo para los alumnos, que como para los padres está inversamente relacionada con un parámetro de eficiencia $v_{Tj}$. Si $\gamma$ es el coste de oportunidad de los maestros,

$$U_{Mj} = \frac{1}{N_j}\sum_{i\in j} v_i e_i + \left(T - \frac{n_j}{N_j}\sum_{i\in j}\frac{c_{2j}e_i}{2v_{Mj}}\right)\gamma \tag{13}$$

donde $N_j$ es el número total de estudiantes en la escuela *j* y $n_j$ es el número de estudiantes por aula.

La función objetivo del decisor político (y su variable de control, el tamaño de clase) maximiza la utilidad de la familia (votante mediano $\overline{P}i$) teniendo en cuenta el coste de la política, que depende del número de clases totales: *N/n*. Es decir, el coste es mayor cuantos menos estudiantes hay en cada aula.

$$U_D = U_{\overline{P}i} - \frac{1}{N}\frac{\omega'}{2}\left(\frac{N}{n}\right)^2 \tag{14}$$

[4] El 2 no es necesario, pero ayuda a ahorrar notación en lo que sigue.

Para simplificar notación denotamos $\omega = \omega^J N$, de manera que:

$$U_D = U_{\overline{P}i} - \frac{\omega}{2}\frac{1}{n^2} \tag{15}$$

### 6.2. La estructura y equilibrio del juego de políticas educativas

El decisor político actúa primero y anuncia los recursos escolares, que en este caso es simplemente el tamaño de la clase ($n$). Después, las familias deciden $s_{1j}$ y las escuelas $s_{2j}$, la potencia de los incentivos para los estudiantes, de manera simultánea. Finalmente, los estudiantes deciden su nivel óptimo de esfuerzo, $e_i$.

Dado que nos enfrentamos a un juego secuencial, vamos a resolver el juego usando la "inducción hacia atrás". Esto es, resolvemos primero la decisión del último jugador que decide, tomando como dadas las acciones de los que movieron con anterioridad. Y después procedemos para cada jugador anterior, que toma como dadas las decisiones de los que mueven antes, pero que tiene en cuenta el efecto de sus decisiones en los que siguen.

Por tanto, primero derivamos la decisión óptima de los estudiantes.

Ejercicio 6. *Mostrar que dado que la utilidad de los estudiantes es (11), su reacción óptima a las decisiones de padres y profesores es:*

$$e_i = c_{1i} + c_{2j} \tag{16}$$

El siguiente paso es derivar las decisiones óptimas de padres y profesores, teniendo en cuenta el impacto de su decisión en los estudiantes. Nótese que, de hecho, este subjuego entre padres y profesores es parecido a un juego de Cournot entre los dos decisores.

Ejercicio 7. *Mostrar que sustituyendo la decisión del estudiante (16) en la utilidad de la familia (12) se obtiene que la elección óptima de la familia es:*

$$c_{1i} = máx\left\{\frac{v_{Pi}v_i}{\psi_i} - \frac{c_{2j}}{2}, 0\right\} \tag{17}$$

Para simplificar la notación en lo que sigue, definimos el parámetro de eficiencia relativa de motivación parental:

$$\Upsilon_{Pi} \equiv \frac{v_{Pi}}{\psi_i}$$

Para los profesores, la solución es bastante simétrica a la de los padres.

Ejercicio 8. *Mostrar que sustituyendo la decisión del estudiante (16) en la utilidad de los profesores (1h) su elección óptima es:*

$$c_{2j} = \frac{v_{Mj}\overline{v}_j}{\gamma n_j} - \frac{\overline{c}_{1j}}{2} \tag{18}$$

Definimos también el parámetro de eficiencia relativa de motivación de los profesores:

$$\Upsilon_{Mj} \equiv \frac{v_{Mj}}{\gamma}$$

Y, por tanto, $\Upsilon_{Mj}\overline{v}_j$ mide la calidad de la escuela.

Por último, caracterizamos la decisión óptima del político:

Ejercicio 9. *Mostrar que sustituyendo la decisión del estudiante (16), profesores (18) y padres (17) en la utilidad del político (15) su decisión óptima es:*

$$n = \frac{\omega - \frac{4v^2}{9\Upsilon_p}\Upsilon^2{}_M}{\frac{4v^2}{9}\Upsilon_M} \tag{19}$$

### 6.3. Estática comparada e implicaciones de política del modelo

La primera implicación del modelo que vamos a explorar tiene que ver con uno de los debates de política y académicos más importantes en nuestros días. Sabemos por el capítulo 3a que la mayor parte de los estudios donde se puede hacer una inferencia causal buena indican que los recursos educativos, y en concreto (una reducción de) el tamaño de las clases tiene impactos positivos en el desempeño de los estudiantes. Y, sin embargo, en la mayoría de los países desarrollados los recursos, incluidos los dedicados a reducir el tamaño de las aulas, han aumentado de manera notable en las últimas décadas sin que se pueda observar una gran mejora del desempeño de los estudiantes. ¿Cómo puede explicarse este fenómeno? La estática comparada del mismo nos da una explicación razonable.

A partir de la decisión óptima del político, expresada en la ecuación (19) se puede deducir la siguiente proposición de manera inmediata.

Proposición 1 (Estática comparatiza del tamaño de clase). El tamaño de clase:

- Aumenta con el coste del profesorado ($\omega$) y la eficiencia relativa de motivación de los profesores ($\Upsilon_M$).
- Decrece con la habilidad de los estudiantes ($v$) y la eficiencia relativa de motivación parental ($\Upsilon_p$).

Esto nos permite entender algunas cosas importantes. Para empezar la paradoja aparente de los recursos. El tiempo que los padres pueden dedicar al proceso de aprendizaje de sus hijos tiene un efecto directo (y observable) en los recursos escolares. Recuerde que $\Upsilon_{Pi}$ aumenta en $v_{Pi}$ (talento de los padres) y disminuye en $\psi_i$ (costo de oportunidad del tiempo de los padres). Las décadas posteriores a la Segunda Guerra Mundial se caracterizaron por un gran aumento de la participación femenina en la fuerza laboral. Podría decirse que esto es compatible con un aumento en $\psi_i$. Pero a su vez creo que también implica un aumento de $v_{Pi}$, la eficiencia de los padres, de ahí que el efecto sobre $\Upsilon_{Pi}$ sea ambiguo. Esto daría como resultado en este modelo una disminución de $n$ (probablemente moderada porque la mejora económica general propició un aumento concomitante en el coste de los profesores $\gamma$). De hecho, Flyer y Rosen (1997) encuentran una conexión entre el rápido crecimiento del gasto público asociado con la educación Primaria y Secundaria en los EE. UU. y el aumento del valor del tiempo de las mujeres. Entre 1960 y 1990, los costes reales de la Educación Primaria y Secundaria aumentaron en un 300 por ciento. La causa principal del aumento en los gastos por estudiante fue el aumento del personal escolar, con una reducción a la mitad del número de estudiantes por maestro entre 1950 y 1990. Flyer y Rosen estudian la conexión entre la participación femenina en la fuerza laboral y la ratio de estudiantes por profesor usando un panel de 50 estados de EE. UU. durante 4 décadas. Observan que los aumentos en las tasas de participación femenina en la fuerza laboral explican una parte significativa tanto del nivel como de la reducción en las proporciones de estudiantes por maestro de los estados. Además, los hallazgos son robustos a otras explicaciones plausibles, como aumentos en la sindicalización y cambios en los patrones de fertilidad.

El efecto de $\omega$ es económicamente obvio, si los salarios de los profesores son más elevados el tamaño de clase es mayor. Nótese que, en las universidades, por ejemplo, los grupos magistrales, normalmente impartidos por los profesores más veteranos y mejor pagados, tienen más estudiantes que los grupos reducidos, muchas veces impartidos por estudiantes de doctorado o profesores asociados. La influencia de los maestros en los recursos de las escuelas es sutil. Recuerde

que $\Upsilon_{Mj}$ aumenta en $v_{Mj}$ (el talento docente) y disminuye en $\gamma$ (coste de oportunidad del tiempo de los docentes). El efecto de $\gamma$ conecta el compromiso de los docentes con la inversión del Gobierno en recursos escolares. Si a los maestros les resulta más difícil dedicar su tiempo a las actividades escolares, los beneficios de reducir el tamaño de la clase son marginalmente menores. Por otro lado, el talento docente ($v_{Mj}$) refuerza los beneficios de invertir en clases más pequeñas. Hasta donde sabemos, estos canales que influyen en la inversión pública en recursos escolares no han sido explorados empíricamente.

También encontramos que los recursos aumentan en la capacidad de los estudiantes. Esta es una consecuencia natural del hecho de que, en este modelo, los beneficios marginales tanto para los padres como para las escuelas de fortalecer los incentivos docentes aumentan con el talento de los estudiantes.

Otros resultados de estática comparativa se pueden observar sustituyendo $n$ en los valores óptimos de equilibrio de $s_1$, $s_2$ y $e$.

$$c_1 = \frac{2\Upsilon_{Pv}}{3}\left(\frac{2\omega - 3\left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_p}}{\omega - \left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_p}}\right) \tag{20}$$

$$c_2 = \frac{2\Upsilon_{Pv}}{3}\left(\frac{3\left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_p} - \omega}{\omega - \left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_p}}\right) \tag{21}$$

lo que implica:

$$e = \frac{\frac{2\Upsilon_{Pv}}{3}\omega}{\omega - \left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_p}} \tag{22}$$

Suponemos que

$$\omega - \left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_p} > 0 \tag{23}$$

A partir de estas tres ecuaciones (20), (21) y (22), podemos deducir la siguiente proposición.

Proposición 2 (Estática comparativa de motivación y esfuerzo):

1. ($c_1$): aumenta en $v$, $\omega$ y $\Upsilon_P$ disminuye en $\Upsilon_T$.

2. ($c_2$): aumenta en $\Upsilon_T$ y $v$, disminuye en $\omega$ y $\Upsilon_P$.

3. ($e$): aumenta en $\Upsilon_T$ y $v$, disminuye en $\omega$, y es ambiguo con $\Upsilon_P$.

La estática comparativa respecto a $\Upsilon_T$ es simple de entender. Una mejora en la capacidad de los profesores aumenta la implicación de la escuela ($c_2$) lo cual a su vez aumenta el esfuerzo de los estudiantes ($e$) y reduce la actividad de los padres ($c_1$), a los que "sustituye". El efecto de la capacidad de los estudiantes $v$ es obvio si observamos (20) y (22). Sin embargo, esta expresión esconde algunas respuestas conductuales más sutiles entre padres y escuelas. Por ejemplo, sabemos que el signo de $\partial c_1/\partial v$ es negativo. Esto quiere decir que el talento de estudiantes reduce la motivación parental en equilibrio. Pero un valor menor de $s_1$ induce mayor esfuerzo motivacional de la escuela ($c_2$) porque los esfuerzos de padres y escuelas son sustitutos entre sí, lo que compensa los menores niveles de $c_1$.

El efecto sobre el esfuerzo ($e$) de una caída en $\Upsilon_P$ (que puede surgir, por ejemplo, de un aumento del coste de oportunidad parental $\psi$) es ambiguo. Primero, un menor $\Upsilon_P$ se traduce en un mayor coste de oportunidad de los padres en proveer de actividades motivacionales a los estudiantes. Por tanto, $c_1$ decrece con el menor $\Upsilon_P$. El sistema educativo reacciona disminuyendo $n$ y aumentando $c_2$. El principal factor de este resultado sobre $c_2$ y $n$ es que el decisor político dedica más recursos a la educación, y esto hace decrecer el coste de inducir esfuerzo por los profesores. Y, al contrario, un menor tamaño de clase y una mayor dedicación de los profesores reduce la ganancia para los padres de inducir un elevado esfuerzo en casa.

La decisión de esfuerzo óptimo del estudiante es igual a la suma de los incentivos de los padres y la escuela. Cuando el coste de oportunidad para los padres aumenta, la caída resultante en la incentivación de los padres no siempre es totalmente compensada por el sistema escolar. Por lo tanto, el efecto neto de una disminución de $\Upsilon_P$ en el desempeño de los estudiantes puede ser negativo. Para ver que esto es cierto observen que:

$$\frac{\partial e}{\partial \Upsilon_p} = \frac{\frac{2}{3} v\omega \Upsilon_p (9\omega \Upsilon_p - 8v^2 \Upsilon_T^2}{\left(3\omega \Upsilon_p - 3\left(\frac{2v\Upsilon_T}{3}\right)^2\right)^2}$$

lo que implica que el esfuerzo decrece con la disminución de $\Upsilon_p$ cuando:

$$\omega < \frac{8v^2 \Upsilon_T^2}{9\Upsilon_p} = 2\left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_P}$$

Nótese que a partir de (23) tenemos,

$$3\left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_P} > \omega \frac{3}{2}\left(\frac{2v\Upsilon_T}{3}\right)^2 \frac{1}{\Upsilon_P},$$

por lo que el esfuerzo y el rendimiento escolar pueden ser tanto crecientes como decrecientes con respecto a $\Upsilon_p$ dentro de nuestro rango paramétrico.

Ejercicio 10. *Discuta algunas implicaciones empíricas o de política educativa que le sugieren los resultados de esta sección.*

### 6.4. Evidencia sobre familias y recursos educativos

Una implicación sencilla del modelo es que una disminución del tamaño de la clase, por ejemplo, a causa de una disminución del coste, conduce a un aumento inequívoco en el esfuerzo y una mejora en el desempeño de los estudiantes. Como hemos visto en el capítulo 3a, la mayoría de los estudios creíbles del efecto del tamaño de la clase (Angrist y Lavy, 1999; Jackson et al., 2016) encuentran que tamaños de clase más bajos generalmente conducen a mejores resultados de los estudiantes.

Sin embargo, hay pocas estimaciones creíbles de las reacciones de los profesores y los padres a clases con tamaños más pequeños. Una excepción es el trabajo Fredriksson et al. (2016) que, de acuerdo con las predicciones de ABC18, encuentran que cuando las clases son más grandes, los padres pasan más tiempo ayudando a los estudiantes con la tarea, pero los profesores tienden a poner más responsabilidad en sus estudiantes. Esta evidencia sugiere que, como indica el modelo de ABC18, las respuestas de todos los actores educativos deben tenerse en cuenta en el diseño de intervenciones de política pública.

Otra implicación interesante de ABC18 es que una mejora en la capacidad promedio de los compañeros de la escuela conduce a un mayor esfuerzo escolar. Sin embargo, los padres de niños en escuelas con mejores compañeros reducen su implicación. Bajo los supuestos paramétricos del modelo de ABC18, las reacciones de los padres no son lo suficientemente fuertes como para compensar el esfuerzo que el estudiante realiza en una escuela mejor. Por lo tanto, el modelo de ABC18 predice un efecto positivo para un estudiante que vaya a una escuela de mayor calidad, aunque esto se verá atenuado por las respuestas de los padres.

Hay una gran cantidad de estudios sobre el efecto de ir a una escuela con mejores compañeros como vemos en el capítulo 3c, Cullenet et al. (2006) y Pop-Eleches y Urquiola (2013) encuentran que los padres de estudiantes que van (son asignados aleatoriamente) a escuelas con mejores compañeros pasan menos tiempo ayudando a sus hijos con las tareas que los padres cuyos hijos van a escuelas con peores compañeros. Esto es consistente con el ABC18. Sobre el resultado final en términos de aprendizaje, hay algunos estudios que encuentran efectos positivos de mejores compañeros en los resultados de los estudiantes Pop-Eleches y Urquiola (2013), Deming et al. (2014) o Dobbie y Fryer (2011). Hay, sin embargo, otros que no encuentran efectos, como Cullen et al. (2006), Abdulkadiroğlu et al. (2014) o Clark (2010). Esta discrepancia puede deberse a las restricciones paramétricas y de forma funcional del modelo en ABC18, que si se eliminan pueden conducir a efectos compensatorios mucho más fuertes de los padres. Por lo tanto, el mensaje general y robusto es que, para interpretar el efecto de forma reducida sobre los resultados de asistir a mejores escuelas, es necesario considerar las respuestas conductuales de los diferentes actores. Esto ayudaría a comprender los contextos en los que los resultados encontrados en un artículo en particular se trasladan a otros contextos.

Finalmente, ABC18 sugiere una conexión entre el coste de oportunidad del tiempo de los padres, los compañeros de escuela, las características de los maestros y la determinación de los recursos escolares. Un argumento relacionado aparece en Besley y Case (2000) que argumentan que las políticas educativas no son exógenas al proceso político. Sin embargo, en la literatura no encontramos estimaciones empíricas creíbles sobre cómo estos factores pueden afectar las elecciones que hacen los políticos acerca de la asignación de recursos escolares.

## 7. EVIDENCIA ADICIONAL SOBRE FAMILIAS Y EDUCACIÓN

En la sección 4.2. del capítulo 4 del volumen 3 en el *Handbook of Economic Education* de Hanushek y Woessman (2011) hay un resumen muy extenso de la evidencia, sobre todo de pruebas internacionales, acerca de la influencia de la familia en el desempeño escolar. Dado el tipo de datos, se trata de lo que hemos llamado influencia "pasiva". Es decir, cómo afectan las características de la familia al resultado, y no tanto el efecto de la involucración "activa". Dentro de estas influencias, conviene destacar por su importancia en España, el impacto del desempleo de los padres en la educación de los hijos. Esto está documentado de manera extensa en Ruiz-Valenzuela (2020b, 2021) para contextos internacionales, y específicamente para España en Ruiz-Valenzuela (2020a). Otro aspecto importante tiene que ver con los roles de género. La influencia de la inserción laboral de las madres en la elección por parte de las hijas de estudios técnicos está bien documentada (San Román y de la Rica, 2016; Cheng et al., 2019 y Anaya et al., 2022) o para las habilidades no cognitivas (Méndez y Zamarro, 2018).

En el mismo volumen 3 del *Handbook of Economic Education* antes citado, el capítulo 3 de Björklund y Salvanes (2011) ofrece un estudio más profundo de las influencias parentales. Por un lado, se concentra en presentar evidencia más plausiblemente causal (sección 3.2.) que la de las pruebas internacionales que ofrecen Hanushek y Woessman (2011), y por otro lado la sección 4.3. ilumina los mecanismos de influencia parental, tanto el dinero (en algún sentido lo que llamamos influencia "pasiva") como el tiempo (la influencia "activa"). El capítulo presta especial atención a la importancia de la dedicación en los años más tempranos de la vida de los estudiantes. Asimismo, la sección 5 estudia diferentes opciones de política, como las bajas maternales o la educación prescolar. En este mismo sentido, acerca de los años tempranos, donde la familia es particularmente importante, el capítulo 3 del volumen 6 en el *Handbook of Economic Education*, de Duncan et al. (2023), ofrece un tratamiento muy detallado.

Existe bastante evidencia experimental (RCT) sobre programas para estimular la intervención parental, parecido a la que discutimos en la sección 3. Una revisión de la literatura muy exhaustiva se encuentra en Fryer (2017). El capítulo estudia la formación de capital humano en general, pero la sección 3.2.1. describe experimentos donde se proporciona a los padres información o incentivos. La sección 3.2.2. intervenciones en los que se dan recursos para la educación en el hogar, como ordenadores, por ejemplo. La sección 3.2.3. describe programas

para reducir la pobreza familiar. Y la 3.2.4. otro tipo de ayudas para mejorar el entorno de la familia, ayudándolos a cambiar de barrio. En general los programas tienen un impacto interesante y un coste bajo. Otro resumen de esa evidencia, algo más reciente, es del *Education Endowment Foundation* (2023). Una ventaja de esta fuente es que tienen unos apéndices técnicos muy bien construidos para ordenar los artículos (en este caso 97). Puede resultar interesante a los lectores escoger algunos y profundizar en ellos.

Un artículo reciente que nos parece importante es de Cunha et al. (2023) que estudia, mediante un experimento controlado, la eficacia un programa para reducir el acoso (bullying) en las aulas. La intervención fue dirigida por madres y padres, duró cuatro meses, e intentaba fomentar la empatía en estudiantes de Secundaria en China. La intervención aumentó la empatía y redujo los incidentes de intimidación. Además, los autores miden el coste y observan que cuesta 12,50 dólares en media reducir un incidente de intimidación. Ese bajo coste es una virtud de este tipo de intervención.

Cuando hablamos del artículo de ABC18 discutimos de manera genérica los recursos destinados a la educación, y lo concretamos en el número de estudiantes en el aula. Pero otra forma de aumentar los recursos cuando los padres no pueden dedicarse tanto a los estudiantes es mediante una jornada escolar (o un año escolar), más largo. Ferrero et al. (2022) revisan la literatura sobre jornada escolar en relación con la oferta de trabajo parental. Otro aspecto importante en esta dirección es el de las clases particulares y actividades extracurriculares. Una buena revisión de la literatura del impacto de las clases particulares se encuentra en Nickow et al. (2020), y una descripción de la situación en España en Moreno y Martínez (2023).

## BIBLIOGRAFÍA


Abdulkadiroğlu, A., Angrist, J. y Pathak, P. (2014). The elite illusion: Achievement effects at Boston and New York exam schools. *Econometrica, 82*(1), 137-196.

Albornoz, F., Berlinski, S. y Cabrales, A. (2018). Motivation, resources, and the organization of the school system. *Journal of the European Economic Association, 16*(1), 199-231.

Albornoz, F., Cabrales, A., Calvo, P. y Hauk, E. (2018). Immigrant children's school performance and immigration costs: Evidence from Spain. *Economics Letters, 170*, 27-30.

Anaya, L., Stafford, F. y Zamarro, G. (2022). Gender gaps in math performance, perceived mathematical ability and college STEM education: The role of parental occupation. *Education Economics, 30*(2), 113-128.

Andrew Young School Policy Studies (2010). *Andrew Young School World Tax Indicators.*

Atlanta, GA: Int. Stud. Progr.

Angrist, J. D. y Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics, 114*(2), 533-575.

Barrera-Osorio, F., Gertler, P., Nakajima, N. y Patrinos, H. (2020). *Promoting parental involvement in schools: Evidence from two randomized experiments* (No. w28040). National Bureau of Economic Research.

Besley, T. y Case, A. (2000). Unnatural experiments? Estimating the incidence of endogenous policies. *The Economic Journal, 110*(467), 672-694.

Bisin, A., Topa, G. y Verdier, T. (2004). Religious intermarriage and socialization in the United States. *Journal of Political Economy, 112*(3), 615-664.

Bisin, A. y Verdier, T. (2000). "Beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics, 115*(3), 955-988.

Bisin, A. y Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory, 97*(2), 298-319.

Bisin, A. y Verdier, T. (2023). Advances in the economic theory of cultural transmission. *Annual Rev. Yew of Economics, 15.*

Björklund, A. y Salvanes, K. G. (2011). Education and family background: Mechanisms and policies. *Handbook of the Economics of Education* (Vol. 3, 201-247). Elsevier.

Boyd, R. y Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago Press.

Cabrales, A. y Hauk, E. (2023). *Parenting styles with externalities*. Working paper in preparation.

Cavalli-Sforza, L. L. y Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach* (No. 16). Princeton University Press.

Chen, L., Chung, B. W. y Wang, G. (2023). Exposure to socially influential peer parents: Evidence from cadre parents in China. *China Economic Review, 78*, 101943.

Cheng, A., Kopotic, K. y Zamarro, G. (2019). *Parental Occupational Choice and Children's Entry into a Stem Field.*

Clark, D. (2010). Selective Schools and Academic Achievement. *The B.E. Journal of Economic Analysis Policy, 10*(1 (Advances)), 9.

Cohen-Zada, D. (2006). Preserving religious identity through education: Economic analysis and evidence from the US. *Journal of Urban Economics, 60*(3), 372-398.

Cullen, J. B., Jacob, B. A. y Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica, 74*(5), 1191-1230.

Cunha, F., Hu, Q., Xia, Y. y Zhao, N. (2023). *Reducing bullying: Evidence from a parental involvement program on empathy education* (No. w30827). National Bureau of Economic Research.

De Philippis, M. y Rossi, F. (2021). Parents, schools and human capital differences across countries. *Journal of the European Economic Association, 19*(2), 1364-1406.

De San Román, A. G. y de La Rica, S. (2016). Gender gaps in PISA test scores: The impact of social norms and the mother's transmission of role attitudes. *Estudios de Economía Aplicada, 34*(1), 79-108.

Deming, D. J., Hastings, J. S., Kane, T. J. y Staiger, D. O. (2014). School choice, school quality, and postsecondary attainment. *American Economic Review, 104*(3), 991-1013.

Dobbie, W. y Fryer Jr., R. G. (2011). Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics, 3*(3), 158-187.

Doepke, M. y Zilibotti, F. (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica, 85*(5), 1331-1371.

Duncan, G., Kalil, A., Mogstad, M. y Rege, M. (2023). Investing in early childhood development in preschool and at home. *Handbook of the Economics of Education* (Vol. 6, 1-77). Elsevier.

Education Endowment Foundation (2023). *Parental engagement*. https://educationendowm evidence/teaching-learning-toolkit/parental-engagement

Fernández, R., Fogli, A. y Olivetti, C. (2004). Mothers and sons: preference formation and female labor force dynamics. *Quarterly Journal of Economics 119*, 1249-99.

Ferrero, M., Gortazar, L. y Martínez, Á. (2022). *Jornada escolar continua: Cómo la pandemia está acelerando un modelo social y educativo regresivo*. EsadeEcPol Brief #26

Flyer, F. y Rosen, S. (1997). The new economics of teachers and education. *Journal of Labor Economics, 15*(1, Part 2), S104-S139.

Fredriksson, P., Öckert, B. y Oosterbeek, H. (2016). Parental responses to public investments in children: Evidence from a maximum class size rule. *Journal of Human Resources, 51*(4), 832-868.

Fryer Jr, R. G. (2017). The production of human capital in developed countries: Evidence from 196 randomized field experiments. *Handbook of Economic Field Experiments* (Vol. 2, 95-322). North-Holland.

Hanushek, E. A. y Woessmann, L. (2011). The economics of international differences in educational achievement. *Handbook of the Economics of Education*, 3, 89-200.

Hayes, B. y Pittelkow, Y. (1993). Religious belief, transmission, and the family: an Australian study. *Journal of Marriage and the Family 55*, 755-66.

Jackson, C. K., Johnson R.C. y Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *Quarterly Journal of Economics, 131*, 157-218.

Masey, D. (2023). *For Darmin Day, 6 facts about the evolution debate*. Pew Research Center. https://www.pewresearch.org/short-reads/2019/02/11/darwin-day/ Mendez, I. y Zamarro, G. (2018). The intergenerational transmission of noncognitive skills and their effect on education and employment outcomes. *Journal of Population Economics, 31*, 521-560.

Moreno, J. M. y Martínez, Á. (2023). *Educación en la sombra en España*. EsadeEcPol Brief #35.

Nickow, A., Oreopoulos, P. y Quan, V. (2020). *The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence*.

Patacchini, E. y Zenou, Y. (2011). Neighborhood effects and parental involvement in the intergenerational transmission of education. *Journal of Regional Science, 51*(5), 987-1013.

Pop-Eleches, C. y Urquiola, M. (2013). Going to a better school: Effects and behavior al responses. *American Economic Review, 103*(4), 1289-1324.

Ruiz-Valenzuela, J. (2020a). Job loss at home: children's school performance during the Great Recession. *SERIEs, 11*(3), 243-286.

Ruiz-Valenzuela, J. (2020b). Intergenerational effects of employment protection reforms. *Labour Economics, 62*, 101774.

Ruiz-Valenzuela, J. (2021). The effects of parental job loss on children's outcomes. O*xford Research Encyclopedia of Economics and Finance*.

# Capítulo 3c
# EL EFECTO DE LOS COMPAÑEROS DE CLASE (*PEER EFFECTS*)

José Montalbán Castilla
Institute for Social Research (SOFI)
Stockholm University
Esade Center of Economic Policy
Universidad Ramon Llull

## 1. INTRODUCCIÓN

La comprensión y medición del proceso por el cual las personas afectan nuestro comportamiento, por ejemplo, inspirándonos a mejorar nuestro rendimiento o incentivándonos a desarrollar actividades poco productivas, está considerado en las ciencias sociales como un factor clave para entender diversos problemas y oportunidades. En el ámbito educativo, el rendimiento o las aspiraciones de los estudiantes se ven afectados simultáneamente por la interacción con sus compañeros de clase, por la influencia de su familia, o por los vecinos con los que comparten barrio. El efecto de los compañeros de clase, comúnmente denominado como *peer effect* (PE), se define como una externalidad, en la que el contexto de los compañeros, su comportamiento y/o sus resultados afectan a otro estudiante (Sacerdote, 2001).

Si la composición de un grupo social, ya sea la clase en la que los estudiantes aprenden o el vecindario en el que los estudiantes viven, tiene un impacto significativo en los resultados de estos a través de las interacciones sociales, esto implica que los PE son importantes para la política pública. El impacto de una configuración del grupo determinada puede ser positivo, negativo o neutral, y normalmente no tiene una respuesta evidente. Por ejemplo, dos preguntas clave de política pública en relación a los PE son: ¿es mejor formar colegios mixtos o segregados en términos de la composición étnica o social de los estudiantes? ¿Es mejor tener grupos muy heterogéneos donde los estudiantes en una clase tengan niveles académicos muy diferentes o concentrar a los estudiantes en grupos homogéneos? ¿Cuál sería la edad óptima a la que deberíamos dividir a los

estudiantes en distintos niveles académicos? La respuesta a estas preguntas no es ideológica, sino que debe ser guiada por un estudio teórico y empírico riguroso, y dependerá de varios parámetros. Por ejemplo, dependerá de si se cuenta con una metodología econométrica que permita identificar un impacto causal, de la función objetivo que se considere (ej., rendimiento académico, aspiraciones, o bienestar subjetivo), o del rango temporal que se observe (ej., corto o largo plazo).

La clara ventaja de los PE desde el punto de vista de las políticas públicas y la planificación social es que son políticas relativamente baratas de implementar en comparación a otras políticas. Es decir, sólo hace falta cambiar la distribución de los grupos de estudiantes para mejorar los resultados, y esto es en teoría sencillo y relativamente barato. Mientras que otras políticas, como, por ejemplo, reducir el tamaño de la clase, implican contratar a nuevos profesores, lo que incrementa necesariamente el coste puesto que el 72% del presupuesto de Educación se destina al pago de salarios (ver capítulo 3a), y el beneficio de la política deberá ser mucho mayor para compensarlo. Por supuesto, existen políticas de PE más ambiciosas que otras, como *Moving to Opportunity* (Chetty, Hendren y Katz, 2016), experimento social aleatorio implementado en EE. UU. en la década de los 1990 diseñado para analizar los efectos de ofrecer la oportunidad a las familias más pobres de cambiar de barrio, en comparación con *Busing* (Angrist et al., 2022), política practicada en EE. UU. que consiste en asignar y transportar a los estudiantes a colegios de otros distritos escolares con el ánimo de aumentar la diversidad racial en estos, por ejemplo. Sin embargo, el punto delicado de estas políticas es definir el contrafactual, ya que estimar la eficacia de estas políticas no es sencilla. Las políticas que cambian la composición de los estudiantes pueden tener efectos directos en la reacción estratégica de las familias cuando eligen colegio o en los incentivos entre colegios públicos y privados, con el consecuente impacto final en la segregación escolar, donde la evidencia empírica no está todavía muy desarrollada (ver capítulo 5a sobre elección escolar).

Ejercicio 1. *Discutir ¿por qué deberíamos encontrar PE? Para su caso en particular, pensad qué personas os influencian (a qué personas influenciáis) y de qué manera en vuestras (sus) vidas.*

Desde un punto de vista teórico, ¿por qué deberíamos encontrar PE? La respuesta a esta pregunta es variada y existen diferentes razones. La primera alternativa se relaciona con el hecho de que personas cercanas normalmente se

comportan de manera similar. Esto puede ser porque son similares ex ante, es decir, tendemos a juntarnos con personas que comparten los mismos valores o un entorno social y económico parecido. También se pueden comportar de manera parecida porque les impacta el mismo ambiente. Es decir, viven en el mismo barrio, viven en el mismo hogar con sus padres, o van al mismo colegio. Además, se comportan de manera similar porque interaccionan entre ellos, y todos se influencian mutuamente, ya que son compañeros de clase, amigos o vecinos. La segunda alternativa tiene que ver con que existen tecnologías de capital humano, es decir, en el proceso de interacción el estudiante aprende de sus compañeros de clase, trabajan en equipo, o se contaminan de las malas prácticas de éstos. Una tercera alternativa es que existen externalidades sociológicas, donde estudiantes de diferentes entornos socioeconómicos eligen distintos modelos de rol a seguir.

Desde un punto de vista empírico, si se quiere tratar de identificar PE, ¿qué características del grupo son relevantes? Imaginemos un caso extremo, en el que queremos identificar PE en el rendimiento académico de un estudiante que vive sólo en una isla desierta. Podemos medir los resultados académicos de este estudiante cuando es el único de la isla, después recibe un *shock,* y un segundo estudiante se une a la misma isla desierta. En este caso las características del grupo serían las de este nuevo estudiante, por ejemplo, sus habilidades académicas o su entorno socioeconómico. Mediríamos los PE comparando el rendimiento académico de nuestro primer estudiante antes y después de la incorporación de su único compañero, y veríamos como las características de este compañero impactan a nuestro estudiante. El problema que plantea este ejercicio es que no existe o no es común en la vida real. Normalmente, los estudiantes se distribuyen en clases, colegios, vecindarios y hogares, y están expuestos a varias personas que conforman un grupo social. Por tanto, debemos determinar qué características del grupo, en este caso la clase, son más importantes para identificar PE. Por ejemplo, ¿lo importante es el nivel académico medio de los estudiantes de la clase o por el contrario es más determinante la heterogeneidad/dispersión de niveles lo que es relevante (ej., varianza)? ¿Son más importantes medidas medias del nivel de la clase o son los valores extremos los que importan más (ej., un estudiante brillante o un compañero disruptivo)? Para responder a estas preguntas, es necesario clasificar las distintas hipótesis en una taxonomía de los PE.

Hoxby y Weingarth (2005) se centran en clasificar los distintos PE en una taxonomía de siete categorías que se centran en diferentes suposiciones:

1) Modelo linear en medias (*Linear-in-means model*): solo la media de las características del grupo o su nivel académico importan.
2) Manzana mala (*Bad apple*): un alumno disruptivo daña el rendimiento de todos los demás.
3) Luz inspiradora (*shining light*): un estudiante brillante inspira a todos los demás.
4) Comparaciones envidiosas (*Invidious comparisons*): los estudiantes de mayor nivel académico empeoran los resultados del grupo.
5) *Boutique/Tracking*: los estudiantes rinden mejor cuando están rodeados por otros como ellos.
6) *Focus*: la homogeneidad de la clase es buena, independientemente de la posición relativa con respecto al resto del grupo.
7) Arcoíris (*Rainbow*): la heterogeneidad en la clase es buena para todos los estudiantes.

Ejercicio 2. *¿Qué más configuraciones de PE se os ocurren que echéis en falta en las siete categorías de la taxonomía de Hoxby y Weingarth (2005)?*

La estructura del capítulo se distribuye de la siguiente manera: en la segunda sección, se explican los desafíos que se encuentran para estimar el efecto causal de los PE; en la tercera sección, se ofrecen las diferentes soluciones y métodos que comúnmente se utilizan para identificar el efecto causal de los PE; en la cuarta sección, se realiza un análisis de los resultados de la investigación en PE; en la quinta sección, se exponen las conclusiones y discusión a las que se llegan en este capítulo.

## 2. ¿CÓMO ESTIMAR EL EFECTO DE LOS COMPAÑEROS DE CLASE?

Medir la importancia y naturaleza de los PE de manera fiable es notoriamente difícil. Esto se debe a que el proceso por el cual los estudiantes, vecinos, o los compañeros de trabajo se aglomeran para formar una clase, un vecindario o una empresa, es todo menos aleatorio. Ciertamente, los padres invierten una gran cantidad de tiempo decidiendo donde van a vivir, o seleccionando el colegio al que quieren llevar a sus hijos, de tal manera que sus hijos puedan beneficiarse de un grupo concreto de compañeros, profesores y recursos escolares. Otro ejemplo pueden ser las empresas de alta productividad, que son especialmente atractivas para trabajadores altamente productivos.

Las personas se seleccionan en colegios, barrios o trabajos, en maneras que hacen muy difícil aislar o distinguir el efecto de la auto-selección a ese grupo social del efecto causal de los PE. Por ejemplo, en la revista académica *New England Journal of Medicine*, Christakis y Fowler (2007) estudian la red de amigos y parejas dentro del *Framingham Heart Study*. En su investigación, encuentran que la probabilidad de una persona de convertirse en obesa es un 57% más alta si tiene un amigo obeso. Esto sin duda es una correlación poderosa, sin embargo, ¿Cuánto de este resultado se debe a la influencia de ese amigo en transformarte en obeso y cuanto a la tendencia de las personas de formar redes de amistad con aquellos que tienen similitudes étnicas, contextos socioeconómicos, o hábitos de ejercicio físico similar? Cohen-Cole y Fletcher (2008) encuentran que, efectivamente, el impacto estimado en la obesidad disminuye mucho si se controla por variables que explican parte de la selección en el grupo. Por tanto, para identificar el efecto causal de los PE, nos vamos a encontrar con problemas de identificación.

Los dos problemas de identificación fundamentales, que Manski (1993) explica en detalle, son los siguientes:

1) Problema de endogeneidad o causalidad reversa (reflexión/simultaneidad): el primer problema de identificación surge del hecho de que es muy difícil distinguir el efecto que el grupo tiene en un individuo concreto, del que este individuo tiene en el grupo, debido a que los resultados están determinados conjuntamente (Vigdor y Nechyba, 2007). Este problema, que econométricamente se conoce como "problema de causalidad inversa", normalmente se denomina como problema de endogeneidad o problema de reflexión (Manski, 2991; Sacerdote, 2001). La intuición que proporciona Manski (1993, p. 128) es la siguiente: *El problema de reflexión es similar a interpretar el (casi) movimiento simultaneo de una persona y su reflejo en un espejo. ¿La imagen de la persona causa el movimiento de la persona o la refleja?*

2) Problema de variables correlacionadas no observables:

    2.1) Autoselección (*endogenous sorting*): el segundo problema de identificación se fundamenta en el hecho de que los individuos tienen a autoseleccionarse en los grupos que quieren. En presencia de autoselección, es muy difícil distinguir el efecto de los PE del efecto de autoselección (Sacerdote, 2001).

2.2) Ambiente o contexto compartido: este problema se fundamenta en el impacto común que ejerce el ambiente que compartide cada grupo. Por ejemplo, los estudiantes que constituyen una clase están expuestos al mismo profesor, o los vecinos que viven en un barrio sufren las actividades criminales que se producen en la esquina de su casa.

Para ilustrar estos tres problemas de identificación, una alternativa es utilizar el modelo clásico lineal en medias de Manksi (1993). Este modelo sería la alternativa número uno mencionada anteriormente en la taxonomía de Hoxby y Weingarth (2005), en la que sólo la media de las características o el rendimiento del grupo importan. En el modelo básico, el objetivo es estimar cómo el rendimiento académico medio de mis compañeros afecta mi rendimiento. Para ello se desarrolla la siguiente ecuación:

$$y_{ig} = \alpha + \beta x_{ig} + \phi \overline{y}_{(-i)} g + \delta \overline{X}_{(-i)} g + \mu_g + \varepsilon_{ig}$$

Donde $\overline{y}_{(-i)} g$ es el rendimiento medio del grupo $g$ sin contar con mi rendimiento (conocido como *leave-out mean*), definido como $y_{ig}$; $\overline{X}_{(-i)} g$ son características observadas del grupo (ej., edad, género, estatus socioeconómico o nacionalidad); $\mu_g$ son efectos correlacionados del grupo $g$. En este ejemplo se pueden distinguir los tres problemas de identificación en los tres parámetros siguientes:

1) Efectos endógenos ($\phi$): este parámetro indica el efecto directo que tiene el comportamiento (o el rendimiento) de los compañeros en mi propio comportamiento (o rendimiento). Este es el parámetro de interés para estimar cómo las notas de mis compañeros afectan a las mías.

2) Efectos exógenos o contextuales ($\delta$): este parámetro indica cómo las características exógenas de mis compañeros afectan mi comportamiento (o rendimiento). Este es el parámetro de interés para estimar como el contexto o las características de mis compañeros afectan mis propios resultados. Por ejemplo, los padres de mis compañeros de clase tienen bajos ingresos observando una baja recompensa por la educación.

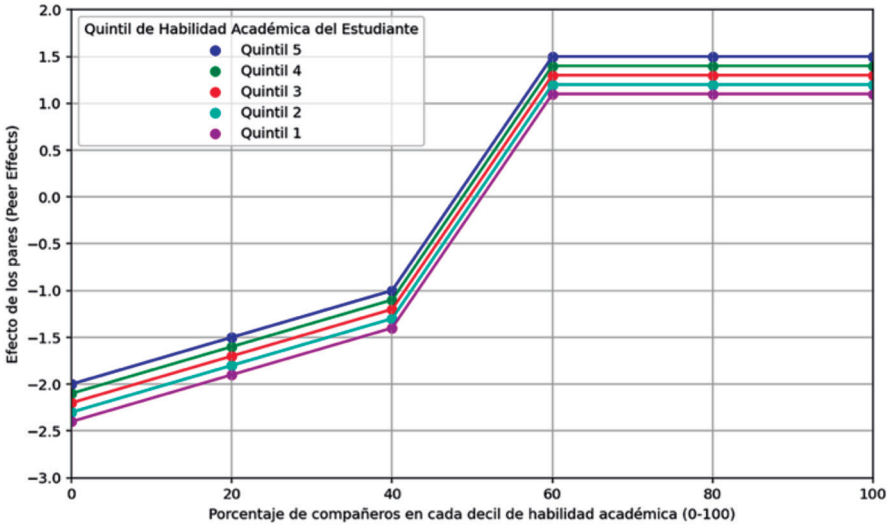
3) Efectos correlacionados ($\mu_g$): este parámetro indica como el ambiente o contexto compartido con mis compañeros de clase afectan mi rendimiento.

El modelo lineal en medias tiene virtudes y defectos. La virtud principal es su sencillez, ya que constituye una manera sencilla de explicar el mundo. La des-

ventaja es que esta manera tan sencilla de entender el mundo, puede no ser una descripción muy convincente de este. Este modelo tiene conclusiones de política pública poco útiles y apasionantes. Esto se debe a que modificar la configuración del grupo en base a los resultados utilizando este modelo no implicaría una mejora universal del rendimiento de los estudiantes, sino que implica necesariamente que unos estudiantes mejoren mientras que otros empeoran. Es decir, por definición, es un juego de suma cero.

Sin embargo, el modelo no-lineal, que ha empezado a utilizarse relativamente recientemente, es más interesante en cuanto a sus aplicaciones de política pública. Como indican Hoxby y Weingarth (2005), los modelos no-lineales abren la posibilidad de que algunos estudiantes puedan mejorar con una nueva configuración del grupo, sin empeorar necesariamente el rendimiento del resto. Por ejemplo, si al cambiar la configuración del grupo encontramos que los estudiantes de alto nivel académico no les afecta la nueva configuración, pero los estudiantes de bajo nivel académico incrementan muy positivamente su rendimiento. Hoxby y Weingarth (2005) construyen un marco coherente para estimar PE que vayas más allá del modelo lineal en medias. Su sugerencia es calcular la fracción de compañeros dentro de un determinado decil de habilidad académica obtenida en el pasado. Esta fracción de compañeros en cada decil de habilidad académica pasada puede ser interaccionada con una variable dicotómica indicando el decil exacto en el que está el propio individuo. En el caso de los deciles, esto llevaría a 100 interacciones (10 variables dicotómicas de la habilidad académica de mis compañeros multiplicadas por 10 propias). Los resultados pueden ser representados gráficamente como en el gráfico 1 para el caso de quintiles. Por supuesto, existen múltiples maneras de llevar a cabo el modelo no-lineal y el descrito aquí es solamente un ejemplo. Este modelo tiene la ventaja añadida de ser relativamente flexible y transparente, y por ello, muchas estimaciones de modelos no-lineales siguen versiones de la especificación presentada.

**Gráfico 1.**




Nota: Elaboración propia.

Ejercicio 3. *¿Qué más modelos no-lineales de PE se os ocurren? Discutir.*

## 3. MÉTODOS PARA ESTIMAR EL EFECTO DE LOS COMPAÑEROS DE CLASE

Los economistas y otros científicos sociales han conseguido abordar algunos de estos problemas de identificación encontrando o creando contextos donde los compañeros son asignados el uno al otro de manera aleatoria (ej., distribuyéndolos en grupos de manera aleatoria).

### 3.1. Variación no experimental

#### *3.1.1. Movimientos exógenos de personas*

Los científicos sociales han utilizado diferentes experimentos naturales o controlados que causan un *shock* exógeno en la localización o colegio, y por ello, a la configuración del grupo, para estudiar el impacto causal de los PE. Se pueden considerar los PE en las personas que se mueven o los PE que ejercen los migrantes en las personas que viven en la localización a la que estos llegan. En principio, los dos tipos de estudios son interesantes de analizar. Por ejemplo, una orden judicial que pretende reducir la segregación de los colegios es un evento

que causa un potencial cambio exógeno en las características de los compañeros en los colegios donde reciben nuevos alumnos. ¿Por qué es exógena? Porque no depende de la decisión directa de las familias, sino de un ente externo que no tiene relación con ellas, como puede ser una orden judicial. Otro ejemplo es utilizar un desastre natural como *shock* exógeno en el movimiento de las personas. ¿Por qué es exógeno? Porque si tu casa se ve arrasada por un huracán, te ves forzado a salir y buscar un nuevo hogar. No elegiste hacerlo por decisión propia, pero el huracán te obligó a hacerlo.

Angrist y Lang (2004) estudian los efectos de ofrecer transporte escolar desde los colegios públicos de Boston a los colegios de Brookline en los estudiantes que viven en Brookline. Para estimar este efecto, se centran en explotar la variación que se produce en el porcentaje de estudiantes que se benefician de esta política a lo largo de diferentes años y cursos escolares dentro del mismo colegio. El trabajo no encuentra efectos significativos de los PE en el rendimiento académico de los estudiantes locales. Otro ejemplo es Imberman et al. (2012), que explotan la variación en la composición del grupo que se produce debido a la evacuación de estudiantes por el Huracán Katrina en colegios de Houston (Texas) y a lo largo de Luisiana. Los autores no encuentran evidencia, de que, en media, la llegada de los evacuados por el Katrina, disminuya el rendimiento de los estudiantes locales. Sin embargo, cuando los autores consideran efectos no-lineales, interaccionando el nivel académico de los estudiantes que llegan con los locales, si encuentran efectos estadísticamente significativos: un incremento de 10 puntos porcentuales en la fracción de "evacuados de bajo nivel académico" se asocia con una reducción del rendimiento académico de los alumnos locales de un 17% de la desviación típica. Así mismo, un incremento de 10 puntos porcentuales en la fracción de "evacuados con alto nivel académico" se asocia a un incremento del rendimiento académico de los alumnos locales de un 9% de la desviación típica.

### *3.1.2. Variación aleatoria entre cohortes*

El concepto de poder identificar PE comparando las características de distintas cohortes dentro de un mismo colegio, no tiene por qué estar limitada a instancias de inmigración, desastres naturales o asilo político. De hecho, una de las primeras aplicaciones de la variación aleatoria de la composición de los compañeros entre cohortes es Hoxby (2000). La visión de este trabajo se basa en el hecho de que las cohortes de los colegios son suficientemente pequeñas para

ofrecer variación significativa en la fracción de chicas y en la composición racial a lo largo de los años. Las chicas tienden a tener mejores resultados en Lectura y registran menos infracciones que los chicos. Por tanto, la variación que se produzca de manera natural en la composición de género (o racial) en un colegio puede proveer de variación exógena en las características de los compañeros, y observar cómo afecta esta al rendimiento académico.

Dos ejemplos de estos métodos son Hoxby (2000) y Lavy y Schlosser (2011). Hoxby (2000) encuentra que el rendimiento académico de los chicos y chicas se incrementa en torno al 2% de la desviación estándar cuando la cohorte tiene 10 puntos porcentuales más de chicas. Lavy y Schlosser (2011) son capaces de ampliar esta estrategia de identificación un paso más allá, pudiendo investigar los mecanismos detrás de estos resultados usando datos educativos de Israel. Estos autores encuentran que incrementar un 10% el número de chicas en la cohorte, aumenta la probabilidad de obtener el diploma de Educación Secundaria en 0,5% para los chicos y 1% para las chicas.

### *3.1.3. Asignación aleatoria de compañeros de cuarto*

Una alternativa a las estrategias mencionadas anteriormente es el uso de procesos sistemáticos donde se asignan aleatoriamente los compañeros de cuarto o clase dentro de un colegio o Universidad concreta. Aunque estos estudios se enfocan en un contexto muy particular, tienen la enorme ventaja de poder identificar grupos de estudiantes que interaccionan muy estrechamente, es decir, es posible identificar muy de cerca cual es el grupo relevante (ej., tu compañero de cuarto).

Algunos ejemplos de estudios que utilizan este diseño son Sacerdote (2001), Zimmerman (2003) o Carrel et al. (2009). Sacerdote (2001) es el primero que analiza PE en el rendimiento académico de los compañeros de cuarto aprovechando su asignación aleatoria en el *Darmouth College* de EE. UU. Sacerdote no encuentra efectos significativos en media, sin embargo, cuando utiliza un modelo no-lineal más flexible, muestra que hay efectos significativos. Todos los estudiantes se benefician académicamente de tener un compañero de cuarto de alto nivel académico, en especial, los estudiantes que también tienen un alto nivel. Los compañeros de cuarto de bajo nivel académico son los que más reducen el rendimiento de aquellos que están en la misma distribución de notas. Carrel et al. (2009) explotan la asignación aleatoria de soldados a diferentes escuadrones

de la *US Air Force Academy*. El diseño de este trabajo ofrece un escenario ideal para estimar PE, ya que los soldados nuevos se asignan aleatoriamente a los escuadrones, y estos tienen muy limitada la interacción con otros escuadrones, de tal manera que están totalmente expuestos al escuadrón asignado. En este ambiente controlado, Carrel et al. (2009) encuentran PE a nivel de escuadrón que un incremento de una desviación estándar en el nivel académico medio de los estudiantes del escuadrón incrementa las notas en un 5% de la desviación estándar.

Ejercicio 4. *Lea el trabajo de Carrel et al. (2009), y discuta: ¿Importa más el nivel académico medio o el no-lineal de los compañeros de escuadrón? ¿Explique en qué medida difiere y que consecuencias de política pública tendría?*

## 3.2. Variación experimental

### *3.2.1. Experimentos aleatorios*

No todos los estudios que están bien identificados utilizan experimentos naturales. Existen varios trabajos influyentes que emplean experimentos aleatorios controlados, donde normalmente estudian el impacto de un tratamiento discreto que manipula la composición del grupo.

Uno de los trabajos más influyentes en este aspecto es Duflo, Dupas y Kremer (2011). Estos autores implementan un experimento controlado para analizar el efecto de componer los grupos con "tracking" con respecto a "no-tracking" en el rendimiento de los estudiantes en Kenia. Los autores cuentan con 121 colegios en los cuales hay dos clases, y seleccionan los siguientes grupos de tratamiento y de control (60 colegios en tratamiento y 61 control): (a) Tratamiento ("tracking"): dentro de cada uno de los colegios, los autores asignan dos grupos en base al nivel académico de los estudiantes, uno donde asignan a la mitad baja de la distribución, y otro la mitad alta; (b) Control ("no-tracking"): dentro de cada uno de los colegios, los autores asignan los estudiantes a los grupos aleatoriamente. Este trabajo encuentra que todos los estudiantes se benefician del "tracking", tanto los de bajo como alto nivel académico, reportando un rendimiento de un 14% de la desviación estándar más alto que los de los colegios sin "tracking". Los estudiantes de bajo nivel académico se beneficiaron del "tracking" gracias a que los profesores ajustaron la instrucción a su nivel.

En otro trabajo seminal, Carrel et al. (2013) realizan un experimento aleatorio controlado en el que asignan a los soldados de la *US Air Force Academy*

en tres grupos: (a) Tratamiento 1 (Escuadrones bimodales): mezcla estudiantes con nivel académico alto con los de bajo; (b) Tratamiento 2 (Escuadrones homogéneos): estudiantes con nivel académico medio; (c) Control: la mitad de los estudiantes fueron asignados aleatoriamente a los escuadrones mezclando niveles académicos. El objetivo de este estudio fue tratar de crear un diseño óptimo de grupos para que, sobre todo, los estudiantes con bajo nivel académico se beneficiaran en mayor medida de los de alto. Los resultados de este estudio no son lo que los autores pensaron en un principio. La división de grupos en tres modos tuvo un efecto negativo en el rendimiento académico de los estudiantes de bajo nivel académico, efectos positivos para los de la mitad de la distribución, y efecto neutral para los de alto nivel académico. Los mecanismos del trabajo indican que los resultados negativos de los estudiantes de bajo nivel académico se produjeron debido a que no se relacionaron con aquellos de alto nivel académico, segregándose cada uno en su subgrupo particular, y dañando el rendimiento de los primeros. Este trabajo nos enseña una importante lección: las respuestas endógenas al comportamiento son un claro obstáculo para predecir los efectos de manipular grupos con el objetivo de obtener un resultado socialmente deseable. Las respuestas endógenas al comportamiento pueden resultar para el planificador social un desafío crucial, ya que la mencionada respuesta endógena aumenta la probabilidad de serias consecuencias no intencionadas.

### *3.2.2. Método de tratamiento parcial* (Partial Population Approach)

En hilo con los resultados de Carrel et al. (2013), el *Partial Population Approach* (PPA) aprovecha las redes y conexiones de amistad ya existentes, posiblemente bastante robustas puesto que se han creado de manera natural, para llevar a cabo una intervención aleatoria a una fracción de esta red. Es decir, primero se identifican las distintas redes de amistad, después se implementa un tratamiento de manera aleatoria a porcentaje de personas dentro de la red, y finalmente, se estiman las externalidades de ese tratamiento en los compañeros no tratados, lo que nos da el PE. Esta idea, se comenzó a utilizar por Moffit (2001), y como indica Sacerdote (2014), es ideal en muchos sentidos. En el mundo real, las oportunidades de asignar aleatoriamente los grupos son muy escasas, mientras que la viabilidad de implementar una intervención aleatoria es mucho más alta.

Un ejemplo es Babcock y Hartman (2010), que implementa esta estrategia en el contexto de ejercicio en el gimnasio en la *UC Santa Barbara*. En lugar de asignar aleatoriamente a los compañeros a un grupo, estos autores implementan una inter-

vención aleatoria, que en este caso es un incentivo para hacer más ejercicio, a una fracción de las redes de amistad. Los autores encuentran que los amigos de los estudiantes que recibieron el tratamiento realizan más ejercicio si una mayor parte de su red recibió el tratamiento. Por tanto, los amigos influenciaron a sus amigos para hacer ejercicio. Avvisati et al. (2014) es otro excelente ejemplo de este método. Los autores implementan un experimento controlado de amplia escala en los colegios de una zona de bajos ingresos en París (Francia). El programa ofreció información a los padres sobre el funcionamiento del colegio, y consejo sobre como apoyar y monitorizar a sus hijos con el trabajo escolar. En cada colegio y clase, se identifican familias que fueron voluntarias y no voluntarias para participar en el programa antes de la implementación del mismo. La mitad de las clases-colegio se seleccionan aleatoriamente para formar parte del programa. Comparando familias voluntarias en clases tratadas con las familias voluntarias en clases en el grupo de control, los autores capturan el efecto directo del programa en los padres tratados y en sus hijos. Comparando familias no voluntarias en clases tratadas y no tratadas, los autores capturan el efecto indirecto (o PE) de tener familias tratadas en la clase. El programa registró efectos muy positivos. Al final del curso, las familias tratadas incrementaron considerablemente su implicación con el trabajo escolar de sus hijos. Los estudiantes de clases tratadas mostraron un mejor comportamiento y actitudes con el colegio, además de tener un impacto positivo indirecto adicional en los resultados de sus compañeros.

### 3.3. Reconciliando la variación natural y experimental

La variación natural, que normalmente explota variación continua en la composición del grupo, tiene ventajas y limitaciones. La ventaja fundamental es que nos permite realizar una estimación flexible de la función de respuesta de los compañeros. Es decir, nos ayuda a estar más preparados para estimar como serían los resultados con configuraciones alternativas del grupo. Sin embargo, tiene dos limitaciones fundamentales. La primera es que esta variación casi siempre ocurre de manera muy limitada. Es decir, afortunadamente los desastres naturales, *shocks* de inmigración o asilo político, no son eventos muy frecuentes en los países desarrollados, y las situaciones en las que se asignan de manera aleatoria a los compañeros de grupo suelen ser escasas. La segunda es que estos estudios tienen claros problemas de soporte. Cuando hay configuraciones del grupo que no se cubren en los datos, la extrapolación confía sólamente en la forma funcional, e incurre en el riesgo de ser invalida. Es decir, lo normal es que la clase de un co-

legio tenga aproximadamente un 50% de chicos y de chicas. La variación natural a la Hoxby (2000) explota ligeros desvíos de esta regla (ej., situaciones en las que hay 48% *vs* 52% de uno de los géneros). Sin embargo, clases en las que haya un 90% *vs* 10% de uno de los géneros son poco frecuentes, y, por tanto, no hay datos suficientes para confiar en su extrapolación. Este ejemplo se ve claramente en la figura 2, de Booij et al. (2017), en la que se observa que el 99% de los grupos que se forman aleatoriamente, no varían mucho entre sí.

Los experimentos aleatorios, que típicamente exploran el efecto de intervenciones discretas (aleatorias), y no continuas, tienen también ventajas y limitaciones. La clara ventaja de estos métodos es que proporciona una estimación muy creíble de una configuración dada del grupo. Sin embargo, la limitación principal es que los tratamientos discretos limitan claramente el alcance de la extrapolación de los resultados, ya que es muy difícil extrapolar los resultados a configuraciones diferentes a la del tratamiento realizado. Por ejemplo, Carrel et al. (2013) obtiene estimaciones creíbles de los efectos de agrupar a los estudiantes del tercio más bajo de la distribución de nivel académico con los del tercio más alto, en comparación con una configuración que mezcle alumnos con diferentes niveles académicos. Duflo et al. (2011) obtiene estimaciones creíbles de los efectos de agrupar alumnos de la mitad más baja con los de la más alta de la distribución de nivel académico, en comparación con una configuración que mezcle alumnos con diferentes niveles académicos. Sin embargo, ninguno de los dos estudios es capaz de identificar el impacto de la configuración de compañeros del otro trabajo en el contexto de su estudio. Es decir, Carrel et al. (2013) no puede estimar el efecto de agrupar alumnos de la mitad más baja con los de la más alta de la distribución en su población de cadetes de la *US Air Force Academy*, ni Duflo et al (2011) pueden estimar el efecto de agrupar a los estudiantes del tercio más bajo de la distribución con los del tercio más alto en su población de estudiantes de Primaria en Kenia.

Por tanto, ¿que se necesita para poder extrapolar configuraciones alternativas del grupo para la misma población y contexto? Principalmente dos ingredientes:

1) Ser capaz de superar el problema de reflexión y selección (idealmente a través de asignación aleatoria al grupo).

2) Asegurar una estimación consistente de la función de respuesta del grupo a lo largo de un soporte suficientemente grande. Es decir, que haya configuraciones del grupo en toda la distribución.

Booij, Leuven y Oosterbeek (2017) diseñan explícitamente un experimento para superar las limitaciones mencionadas. Realizan un experimento aleatorio que manipula la composición y habilidades de los estudiantes en diferentes grupos para los alumnos de primer curso de Economía en la Universidad de Ámsterdam para que existan configuraciones del grupo en toda la distribución posible. La aleatoriedad de esta configuración soluciona el problema de selección y reflexión. El hecho de que los grupos están manipulados para que exista suficiente variación en su composición (ver puntos azules de la figura 2, de Booij, Leuven y Oosterbeek (2017), permite estimar el impacto de diferentes escenarios en el rendimiento total de los estudiantes.

Es importante resaltar que, en la investigación sobre PE, existe una compensación o *trade-off* entre la validez externa y la robustez de la identificación empírica. La mayoría de los trabajos de PE se basan en contextos muy específicos, donde la variación exógena está extraída de manera muy creativa y sólida. Por ejemplo, la asignación aleatoria de compañeros de cuarto en un *college* concreto de una Universidad de EE. UU. Sin embargo, esto tiene el coste de que sea potencialmente difícil extrapolar a otros contextos más generales. También hay contextos que engloban contextos más generales y extrapolables, pero suelen tener el coste de una estrategia de identificación menos clara. Por tanto, a la hora de interpretar los resultados de los distintos trabajos y sacar conclusiones de política pública, se deben tener muy en cuenta estas dinámicas.

## 4. ¿QUÉ CONOCEMOS SOBRE EL EFECTO DE LOS COMPAÑEROS DE CLASE EN EDUCACIÓN?

Esta sección resume los principales resultados de la investigación de PE en educación. En todas las subsecciones se distinguen entre los PE que ejercen los compañeros de clase, la familia y los vecinos.

### 4.1. Rendimiento académico

#### *4.1.1. Compañeros de clase*

La mayoría de los estudios que investigan la influencia de los PE en el rendimiento académico se centran en los compañeros de clase. Existen numerosos estudios que se dedican a capturar el efecto de las características de los compañeros (ej., habilidad académica, género, raza, origen socioeconómico). Estos estudios

normalmente encuentran efectos no significativos, o si lo son, estos son muy modestos en magnitud. La mayoría de los que exploran efectos heterogéneos encuentran que hay algunos grupos de estudiantes que se ven afectados más que otros por las características de los estudiantes.

Los estudios que se enfocan en identificar en el efecto de tener distintas proporciones de compañeros de distinto género, están los ya mencionados Hoxby (2000) y Lavy y Schlosser (2011), a los que se suma Black et al. (2013). Black et al. (2013) encuentra que el cociente intelectual de los chicos no está afectado por la exposición a compañeras chicas durante el instituto en Noruega. Además, muestran que características como la edad media, la educación e ingresos de los padres de los compañeros de clase no afectan prácticamente al cociente intelectual de sus compañeros. También hay investigación que se centra en averiguar si la raza de los compañeros de clase tiene algún impacto en el rendimiento académico. Angrist y Lang (2004) es un ejemplo ya mencionado. Además, Figlio y Özek (2019) realizan un estudio explotando la gran afluencia de migrantes provenientes de Haiti, no nativos en inglés y de bajos recursos, hacia los colegios públicos de Florida debido al terremoto de 2010. Estos autores encuentran efectos de prácticamente cero en el rendimiento académico de los estudiantes floridanos, independientemente de su estatus socioeconómico, la edad, raza o lugar de nacimiento.

Otro bloque de investigación se centra en observar los efectos de estar expuestos a compañeros con distintos niveles académicos. Estos trabajos, como en el caso de las características de los estudiantes, encuentran resultados mixtos. El ya mencionado Imberman et al. (2012), es un ejemplo. Abdulkadiroglu et al. (2014) explotan variación cuasi-experimental generada por los umbrales de admisión en los colegios más selectivos en Boston y Nueva York. A pesar de que estos umbrales ofrecen variación significativa en el nivel académico de los estudiantes, esta investigación encuentra que el nivel académico de los compañeros de clase no parece afectar demasiado al rendimiento académico de los estudiantes. El ya mencionado Booij et al. (2017), encuentra que los estudiantes en la parte más alta de nivel académico, no ven su rendimiento académico afectado por la composición del grupo. Sin embargo, aquellos estudiantes que están en lo más bajo y en el medio de la distribución de nivel académico, si se les agrupa con estudiantes de su mismo nivel académico, mejoran su rendimiento en un 19% de la desviacion estándar y reducen el abandono universitario en 12 puntos porcentuales.

Además, hay otras maneras en las que el nivel académico de los estudiantes puede impactar en el rendimiento de sus compañeros. Un ejemplo es Murphy y Weinhardt (2020), que usan datos del universo de estudiantes de Inglaterra, donde muestran que independientemente del nivel académico, el orden académico relativo en la clase que tenía un estudiante durante Educación Primaria, tiene consecuencias a medio plazo en el rendimiento académico en Educación Secundaria. Es decir, que, a igual nivel académico absoluto del estudiante, la posición relativa que uno tuviera en su antigua clase tiene consecuencias en el largo plazo. Estos efectos son particularmente pronunciados para los hombres, y se observan impactos en el nivel de confianza de los estudiantes, e impactos añadidos en la elección de continuar en la vía académica.

Ejercicio 4. *¿Qué implicaciones de política pública tienen estos resultados? Si fuera usted el Ministro de Educación de España, discuta si llevaría a cabo alguna política educativa concreta extraída de dicho trabajo, el porqué de la decisión, además de recalcar los potenciales costes y beneficios de dicha política.*

### *4.1.2. Familia*

Existe también otro bloque de investigación destinado a estudiar cómo nos afectan nuestros hermanos en el rendimiento académico. Estos estudios suelen encontrar efectos más grandes que los que se documentan para compañeros de clase. Sin embargo, existe evidencia que señala que, dependiendo del contexto general y familiar, el signo de los efectos puede cambiar. Este es un elemento importante a tener en cuenta, especialmente cuando se intenta incorporar este tipo de PE en el análisis coste-beneficio y en el diseño de nuevas políticas públicas (ver capítulo 3b sobre Familias).

Qureshi (2018b) encuentra que los hermanos pequeños se benefician de que su hermano mayor tuviera profesores con mayor experiencia en Carolina del Norte. De acuerdo a este estudio, una mayor experiencia del profesor de lengua mejora los resultados académicos de lengua de los hermanos mayores en 4,2% de la desviación estándar, y sucesivamente, la de los hermanos pequeños en 1%. Sin embargo, este efecto no es simétrico, ya que los hermanos mayores no están afectados por la experiencia de los profesores de los hermanos pequeños. Nicoletti y Rabe (2019) estudian las externalidades entre hermanos en los colegios públicos de Inglaterra. Este trabajo encuentra que un incremento de una desviación estándar en el rendimiento académico de los hermanos mayores mejora el de los pequeños en un 11% de la desviación estándar.

### *4.1.3. Vecinos*

El tercer tipo de PE que se presenta en este capítulo son los vecinos, una línea de investigación relativamente reciente. Identificar el efecto del barrio es complicado ya que la gente que vive cerca normalmente comparte otras características que están correlacionadas con el rendimiento académico en el colegio. Además, identificar a los vecinos que viven muy cerca, que son los que más efectos (a priori) podrían tener, es difícil ya que la mayoría de las bases de datos con información del barrio, definen esta categoría de una manera muy amplia. Existe evidencia robusta de que la exposición a un barrio mejor desde pequeño mejora los ingresos futuros, incrementa la probabilidad de ir a la Universidad y reduce la probabilidad de tener un embarazo siendo adolescente (Chetty et al., 2014, 2016; Chetty y Hendren, 2018ab). Sin embargo, los resultados de estos trabajos presentan la duda de hasta qué punto el efecto de los barrios se deriva de una mayor exposición a vecinos de mayor nivel educativo o bienes locales de mayor calidad (ej., colegios, servicios de salud, infraestructura o seguridad). Las implicaciones de política pública son muy diferentes dependiendo de cuál sea la respuesta.

Ejercicio 5. *Detalle y discuta cuáles son las implicaciones de política pública a las que se llegaría dependiendo de cuál sea el efecto que domine en el párrafo anterior.*

Algunos ejemplos de PE de vecinos son Goux y Maurin (2017) o Aslund et al. (2011). Goux y Maurin (2007) es uno de los primeros estudios que investigó los PE entre vecinos. Los autores explotan algunas características únicas de Francia, encontrando que el rendimiento de un adolescente al final del instituto está influenciado fuertemente por el nivel académico de los otros adolescentes del barrio. De acuerdo con este estudio, la fecha de nacimiento de los vecinos que viven cerca, y la proporción de familias que son asignadas a vivienda social, impactan significativamente el rendimiento académico en la adolescencia. Aslund et al. (2011) estudian cómo la exposición a adultos de alto nivel educativo en el barrio impacta el rendimiento académico. Para identificar este efecto, los autores utilizan un análisis cuasi-experimental explotando la localización a la que se asigna a los refugiados recientemente llegados a Suecia.

## 4.2. Trayectoria educativa

Existe evidencia que se centra en la influencia que tienen las otras personas en nuestra trayectoria educativa. Este es un margen importante de estudio, so-

bre todo en contextos con altos niveles de segregación, donde la influencia de otras personas puede estar detrás de las desigualdades persistentes entre diferentes grupos sociales. Las otras personas pueden influenciar nuestras trayectorias educativas transmitiendo información relevante, proporcionando apoyo en las solicitudes a los distintos colegios o la Universidad, ayudándonos en la preparación de exámenes competitivos, o incluso afectando nuestras preferencias.

### *4.2.1. Compañeros de clase*

El ya mencionado estudio de Black et al. (2013), también encuentra que una mayor proporción de chicas en el curso mejora los resultados a largo plazo en el mercado de trabajo para las chicas. Sin embargo, una mayor presencia de chicas parece dañar los resultados de los chicos, al disminuir la probabilidad de escoger la vía académica y completar menos años de educación. Carrel et al. (2010, 2018) exploran la hipótesis de la *bad apple* o cual es la influencia de los alumnos disruptivos en el resto de compañeros de clase (mencionada anteriormente en la sección 1). En este estudio, los alumnos disruptivos se aproximan como aquellos estudiantes que tuvieron experiencias de violencia doméstica en su hogar, ya que estos estudiantes son más propensos a registrar comportamientos relacionados con la violencia y los problemas para relacionarse con sus compañeros (Carlson, 2000). Este trabajo encuentra que estar expuesto a un estudiante disruptivo adicional en una clase de 25 estudiantes, en especial si este compañero es chico, reduce la futura matriculación y graduación universitaria de los compañeros, además de registrar menores ingresos laborales medidos cuando estos tienen entre 24 y 28 años.

Bursztyn y Jensen (2015) y Bursztyn et al. (2019) escogen un ángulo diferente al estudiar como la presión y las normas sociales de los compañeros influencian la inversión de los estudiantes en educación. Ambos estudios realizan experimentos aleatorios controlados. Estos estudios muestran que los compañeros de clase ejercen una externalidad amplificadora en la presión social y normas sociales, influenciando las elecciones individuales de los estudiantes. En otro trabajo reciente, Barrios-Fernández et al. (2022) muestra que los compañeros de Universidad pueden tener efectos intergeneracionales en las trayectorias educativas. Usando una regresión de discontinuidad, este estudio muestra que los compañeros de Universidad de los padres chilenos, influenciaron en gran medida el tipo de instituto y los círculos sociales a los que pertenecieron sus hijos en el futuro.

### *4.2.2. Familia*

Varios estudios se han centrado en la influencia de los hermanos en los años de educación completados. Qureshi (2018a) estudia el efecto que tienen las hermanas mayores en los hermanos pequeños en Pakistán. Este estudio encuentra que cuantos más años de educación tiene la hermana mayor, el hermano pequeño sigue su estela aumentándolos también. Gurantz et al. (2020) también documentan externalidades positivas de los hermanos en la probabilidad de presentarse a exámenes avanzados que normalmente son requisito para acceder a la Universidad en EE. UU. Joensen y Nielsen (2018) explotan variación cuasi-experimental en un estudio piloto a nivel colegio en Dinamarca, y muestran que una mayor matriculación del hermano mayor en cursos avanzados de Matemáticas y Ciencias durante el instituto, aumenta la probabilidad del hermano pequeño de matricularse en dichos cursos. Dahl et al. (2020) muestran que el campo de especialización educativa en la que los hermanos mayores deciden especializarse durante el instituto tiene un gran impacto en aquella que eligen sus hermanos pequeños en Suecia. Esto indica que los hermanos mayores no sólo impactan la cantidad de educación adquirida, sino el tipo de educación que los hermanos pequeños tratan de conseguir. Altmejd et al. (2021) investigan el impacto de los hermanos en la transición del instituto a la Universidad en un análisis utilizando datos administrativos de Chile, Croacia, Suecia y EE. UU. Los autores encuentran que los hermanos mayores afectan en gran medida la decisión de matricularse en la Universidad, así como la decisión de en qué Universidad matricularse.

Ejercicio 6. *Discuta cuales son los mecanismos más relevantes que llevarían a los efectos que se muestran en esta sección. Con los resultados mostrados ¿cuál o cuáles cree usted que dominan?*

### *4.2.3. Vecinos*

La investigación que se centra en el impacto de los vecinos en las trayectorias educativas es bastante escasa, y relativamente reciente, por lo que aún es necesaria mucha más evidencia.

Bobonis y Finan (2009) estudian el impacto de los vecinos en la matriculación en Educación Secundaria en zonas rurales de México, utilizando la asignación cuasi-aleatoria del programa de transferencias condicionales Progresa. Estos autores encuentran que Progresa incrementa la matriculación en Educación Secundaria para los hijos de familias que recibieron el programa, y para los hijos

que no lo recibieron. Este último resultado indica la influencia de los vecinos en la decisión de matricularse en Educación Secundaria.

Barrios-Fernández (2022) documenta unos efectos similares, pero en la decisión de matricularse en la Universidad. Este estudio explota la variación generada por las normas que determinan la concesión de préstamos universitarios en Chile. Muestra que los estudiantes del último curso del instituto son más propensos a asistir y completar una carrera universitaria cuando su vecino más cercano (definido como la persona que vive más cerca que solicitó entrar en la Universidad un año antes) obtuvo un préstamo y se matriculó en la Universidad. Este incremento en la matriculación se da a través de un aumento en la probabilidad de hacer el examen final de admisión y solicitar una carrera Universidad.

Michelman et al. (2022) estudian PE en la Universidad explotando asignación aleatoria de estudiantes a dormitorios en Harvard entre los años 1920 y 1930. Estos autores encuentran que los individuos de alto estatus social se benefician de interactuar con compañeros de cuarto que tienen un alto estatus social, es decir, aumentan su probabilidad de formar parte de clubs exclusivos y de completar carreras exitosas en el mundo de las finanzas, mientras que los individuos de bajo y medio estatus social no se benefician en absoluto. Este trabajo muestra que los PE no afectan necesariamente nuestras trayectorias por razones académicas. De hecho, los efectos documentados en este trabajo están más relacionados con dimensiones sociales que académicas.

### 4.3. Resultados no académicos

Las dos secciones previas se centran en el impacto de los PE en el rendimiento académico y las trayectorias educativas. Sin embargo, los compañeros de clase, la familia y los vecinos, también pueden influenciar otras variables de interés.

#### *4.3.1. Compañeros de clase*

Existe una gran cantidad de estudios que se han dedicado a investigar la influencia de los compañeros de clase en el colegio y la Universidad en los resultados que van más allá de lo académico.

Gaviria y Raphael (2001) es quizás el estudio que mira al mayor número de variables no educativas como el uso de drogas, consumo de alcohol y tabaco, asistencia a la iglesia, o el abandono escolar. Utilizan una muestra de estudiantes

de unos 15 años en EE. UU., y encuentran una relación muy fuerte en la influencia que ejercen los compañeros de colegio en las cinco variables comentadas.

Otros estudios analizan PE en el mal comportamiento en el colegio. Por ejemplo, Benett y Bergman (2021) usan datos administrativos del estado de West Virginia en EE. UU. y desarrollan un nuevo método para identificar las redes de amistad en base a estudiantes que practican juntos el absentismo escolar. Este estudio muestra que realizar una intervención a los padres en la que se les da información con respecto al absentismo de sus hijos, tiene une externalidad positiva adicional en los compañeros de estos hijos no tratados por la intervención, ya que al ser amigos interactúan, y este tiene un efecto multiplicador. Los autores muestran que ignorar esta externalidad, infraestima el coste-efectividad del programa en un 43%. Carrel et al. (2008) también investiga PE en la probabilidad de copiar en los exámenes en las tres mayores academias militares de EE. UU. Estos autores encuentran que cuanto más copian los compañeros, mayor es la probabilidad que el estudiante acabe copiando también.

En un margen más dramático, también existe evidencia disponible que los compañeros de colegio pueden influenciar la participación en actividades criminales. Deming (2011) estudia el efecto del colegio e instituto al que van los estudiantes en el crimen futuro de adultos en Charlotte-Mecklenburg en EE. UU. El estudio presenta evidencia sugestiva de que mientras que en el colegio la influencia de los compañeros es un factor decisivo del efecto en crimen, en el instituto el factor más importante es la calidad de la educación recibida. Eren et al. (2022) explotan umbrales de repetición para alumnos de 13 años, encontrando que la repetición escolar incrementa la probabilidad de llevar a cabo un crimen violento en 1 punto porcentual (58%). Como ocurre en Deming (2011), los compañeros parecen tener una influencia decisiva para explicar estos resultados.

Ejercicio 7. *¿Cree usted que este caso es extrapolable a España? Base su respuesta en múltiples datos concretos sobre la criminalidad, repetición escolar y calidad de la educación en España en comparación con otros países.*

### *4.3.2. Familia*

Existen múltiples estudios que documentan la alta correlación de comportamientos de riesgo entre hermanos, sugiriendo que los hábitos pueden ser influenciados por los hermanos. Altonji et al. (2017) investigan si las correlaciones entre el consumo y venta de drogas de los hermanos se explican, al

menos parcialmente, por un efecto causal. Concluyen que, aunque el consumo de tabaco, alcohol, y mariguana parece estar influenciado por el ejemplo que dan los hermanos mayores, la mayoría de las correlaciones observadas entre hermanos provienen de las influencias comunes que tienen ambos (i.e., efectos correlacionados). Bingley et al. (2021) documentan PE entre hermanos en la selección del sector laboral que se elige en Dinamarca. Encuentran que la asignación aleatoria del hermano mayor a servir en el Ejército incrementa la probabilidad del hermano pequeño de servir también en un 7%. Heissel (2017) muestra que los hermanos que tienen hermanas que fueron madres durante la adolescencia presentan un menor rendimiento académico en el colegio, tienen mayores probabilidades de abandonar el instituto, y de interaccionar con el sistema judicial juvenil, al nacer su sobrino. El incremento de la necesidad de apoyo parental de atención y recursos después del nacimiento de un nuevo hijo en la familia puede explicar las externalidades negativas documentadas en este estudio.

#### *4.3.3. Vecinos*

Existen estudios que documentan consistentemente efectos de los vecinos en el crimen y la delincuencia. Kling et al. (2005), por ejemplo, estudia la iniciativa de *Moving to Opportunity* (MTO) en EE. UU. en la delincuencia y el crimen juvenil. El MTO asignó aleatoriamente cheques de vivienda para familias que podían usarlos para vivir en un barrio mejor en cinco ciudades de EE. UU. El estudio muestra que realojarse a áreas de menor pobreza, reduce los arrestos por robo de la propiedad y crimen violento entre las chicas jóvenes. También reduce la participación en el crimen violento de los chicos, pero aumenta la participación en robos contra la propiedad. Billings et al. (2019) muestran que el barrio y la segregación escolar incrementan el nivel de crimen al fomentar las interacciones sociales de jóvenes en riesgo. El estudio muestra que agrupar a más estudiantes desaventajados en el mismo colegio incrementa el crimen total. Además, estos estudiantes desaventajados aumentan la probabilidad de convertirse en "compañeros de crimen", y acabar cometiendo crímenes juntos.

Los vecinos también pueden afectar otras variables. Corno y Burns (2022) explotan la asignación aleatoria de compañeros de cuarto en la Universidad en Sudáfrica, mostrando que ser asignado a una habitación interracial reduce los prejuicios y mejora las actitudes entre los estudiantes de distintas razas. Además, incrementa la probabilidad de crear amistades interraciales.

## 5. CONCLUSIONES Y DISCUSIÓN

La evidencia discutida en este artículo pone de manifiesto la importancia de los muchos márgenes en los cuales las otras personas, ya sean compañeros de clase, familia o vecinos, influencian los resultados de los estudiantes. Cuando se analiza en conjunto toda la evidencia de PE, se sacan cuatro conclusiones fundamentales:

1) Como indica Sacerdote (2014), sólo la mitad de los estudios que utilizan un modelo lineal en medias para explicar la realidad de los PE encuentran efectos modestos en el rendimiento académico, y la otra mitad, no encuentra efectos significativos.

2) El efecto de los compañeros de clase se incrementa en magnitud y en poder estadístico cuando relajamos lo supuestos del modelo linear en medias, y aplicamos un modelo no lineal para describir la realidad.

3) Los resultados de los PE en variables no académicas y sociales muestran efectos muy grandes en comparación con los encontrados en el rendimiento académico. Es decir, es más sencillo que los compañeros de clase influencien variables como el comportamiento en clase, el consumo de drogas o la actividad física, que el rendimiento académico (ver también capítulo 4.a. sobre Educación Infantil).

4) Todavía no conocemos suficiente la naturaleza de los PE como para implicarnos en experimentos de ingeniería social (Sacerdote, 2014), ya que: (i) los resultados no son consistentes entre estudios; (ii) varían mucho dependiendo del contexto; (iii) existen claras respuestas endógenas a la manipulación de los grupos, como, por ejemplo, en Carrel et al. (2013).

Ejercicio 8. *¿Cómo cree usted que se pueden solucionar las limitaciones mencionadas en el punto (4)?*

Confirmar empíricamente la existencia de PE es importante ya que afecta al análisis coste-beneficio de diferentes políticas públicas, y pueden ser incorporados en el diseño de nuevos programas para maximizar su impacto. En el contexto educativo, como indica Barrios (2023), pueden ayudar a explicar la persistencia en la desigualdad de distintos grupos sociales. Sin embargo, realizar recomendaciones generales de política pública en este respecto es complicado, ya que el tamaño de los efectos documentados en este capítulo a lo largo de los distintos estudios varía sustancialmente entre contextos y dependiendo de cómo

se caracterizan y definen las personas de influencia. Comprender qué hay detrás de esta heterogeneidad es clave para extraer lecciones de política pública más generales, que tengan una visión más completa sobre la naturaleza y la relevancia de las interacciones sociales en los contextos educativos.

Un primer paso consiste en identificar los mecanismos que hay detrás de los PE, en el que todavía hay mucho que hacer. Por ejemplo, los efectos positivos en el rendimiento académico asociado con tener una fracción más alta de chicas en la clase han sido atribuida a la mejora del ambiente de aprendizaje asociada con esto, es decir, menos disrupciones en la clase, mejor relación entre estudiantes, entre estudiantes y profesores, y profesores menos fatigados. Si un mejor ambiente en la clase es en efecto el principal mecanismo detrás de estos resultados, esto podría explicar por qué en algunos contextos una mayor proporción de compañeras chicas parece no tener ningún efecto (ej., Black et al., 2013). Hay otras razones por las cuales los compañeros pueden impactar los resultados de los estudiantes de manera diferencial. Por ejemplo, los compañeros de alto rendimiento académico pueden mejorar el ambiente de aprendizaje en la clase generando menos disrupciones y ayudando a otros estudiantes, pero al mismo tiempo, pueden hacerlo más difícil para que los estudiantes de bajo rendimiento hagan preguntas y participen activamente en la clase. Hay evidencia que apoya que ambos mecanismos existen, y no está claro que hace un mecanismo más relevante que el otro en los distintos contextos. La posición académica relativa de un estudiante en la clase se ha demostrado que afecta su confianza y rendimiento futuro (ej., Murphy y Weinhardt, 2020). Por tanto, estar rodeado por alumnos de alto rendimiento puede tanto mejorar como empeorar los resultados académicos de los compañeros.

La comprensión de los mecanismos detrás de los resultados, además de ayudarnos a racionalizar las diferencias en PE que surgen en diferentes contextos, son claves para tomar decisiones de política pública. Si la mejora del ambiente de aprendizaje en la clase es un mecanismo fundamental de los PE, entonces tendríamos que intentar asignar los estudiantes a grupos, tratando de minimizar las disrupciones en clase (ver capítulo 3a donde se expone el modelo de Lazear de los efectos de las disrupciones en clase), y maximizar las interacciones positivas. Sin embargo, otra alternativa es generar un mejor ambiente de aprendizaje modificando la cultura del colegio o reduciendo el tamaño de la clase. Otro ejemplo es la gran evidencia que documenta PE en las trayectorias educativas, que no ha sido capaz de concluir qué mecanismo es más relevante: transmisión de la información, apoyo académico

directo, cambio en preferencias (ej., cambios en las aspiraciones o la motivación). Comprender esto es clave para entender qué herramientas de política deberían ser empleadas. Por ejemplo, si el mayor mecanismo es la transmisión de la información, entonces los gobiernos podrían diseñar estrategias para sustituir la ausencia de esta en los compañeros, y proveerla directamente a los estudiantes, aunque diseñar intervenciones efectivas de este tipo no es fácil (ej., Dynarski et al., 2021). Por tanto, comprender los mecanismos que hay detrás de los PE es fundamental para saber cuáles son las políticas alternativas o complementarias que pueden ayudar a mejorar los resultados de los estudiantes.

Independientemente de los mecanismos, la existencia de externalidades sociales en educación son un hecho relevante para entender las desigualdades en el rendimiento educativo. Esto implica que las consecuencias de las barreras y desafíos que afectan desproporcionadamente a ciertos grupos de estudiantes, se amplifican a través de las interacciones sociales. Sin embargo, desde un punto de vista más optimista, estas externalidades sociales implican que programas diseñados para ampliar el acceso a las oportunidades educativas de estudiantes desaventajados, muy probablemente tengan efectos más grandes que aquellos que sólo miden sus efectos en estudiantes concretos e ignoran los efectos en el grupo social más amplio (ej., Bennet y Bergman, 2021).

Todos estos resultados tienen implicaciones en las políticas públicas que se deben aplicar. Como sugiere Sacerdote (2014), las políticas públicas que implementen intervenciones como la de Duflo y Saez (2003) o Avvisati et al. (2014), en la que simplemente se intervenga a una parte de los compañeros de la red de amistad (esperando que estos influencien a los compañeros no tratados), pero que no intenten manipular activamente quién está en la red o cuanto interaccionan los diferentes individuos, serán más seguras y tendrán mayor probabilidad de alcanzar el objetivo marcado. Es decir, el *Partial Population Approach* es el método potencialmente más seguro para conseguir implementar políticas con el ánimo de influenciar a los compañeros a través de los efectos multiplicadores de las intervenciones.

En general, no hay duda de que se ha avanzado mucho en la investigación sobre PE, y que queda demostrado que estos ejercen una influencia muy poderosa en el comportamiento humano. A pesar de las limitaciones y problemas de identificación, se ha producido un sólido progreso hacia la comprensión de en qué dominios los compañeros importan más, y cómo de grandes son estos efectos. Lo que queda por descubrir en el futuro sobre este tema, es apasionante.

## BIBLIOGRAFÍA


Abdulkadiro˘glu, A., J. Angrist y P. Pathak (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica, 82*(1), 137-196.

Altmejd, A., A. Barrios-FernÅLandez, M. Drlje, J. Goodman, M. Hurwitz, D. Kovac, C. Mulhern, C. Neilson y J. Smith (2021). O Brother, where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries. *The Quarterly Journal of Economics, 136*(3), 1831-1886.

Altonji, J. G., S. Cattan y I. Ware (2017). Identifying Sibling Influence on Teenage Substance Use. *Journal of Human Resources, 52*(1), 1-47.

Angrist, J. D. y Lang, K. (2004). *Does school integration generate peer effects? Evidence from Boston's Metco Program. American Economic Review, 94*(5), 1613-1634.

Angrist, J., Gray-Lobe, G., Idoux, C. M. y Pathak, P. A. (2022). *Still Worth the Trip? School Busing Effects in Boston and New York* (No. w30308). National Bureau of Economic Research.

Aslund, O., P.-A. Edin, P. Fredriksson y H. GrÅNonqvist (2011). Peers, Neighborhoods, and Immigrant Student Achievement: Evidence from a Placement Policy. *American Economic Journal: Applied Economics, 3*(2), 67-95.

Avvisati, F., Gurgand, M., Guyon, N. y Maurin, E. (2014). Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies, 81*(1), 57-83.

Babcock, P. S. y Hartman, J. L. (2010). *Networks and workouts: Treatment size and status specific peer effects in a randomized field experiment* (No. w16581). National Bureau of Economic Research.

Barrios-Fernández, A. (2022). Neighbors' Effects on University Enrollment. *American Economic Journal: Applied Economics, 14*(3), 30-60.

Barrios- Fernández, A., C. Neilson y S. Zimmerman (2022). *Elite universities and the intergenerational transmission of human and social capital.*

Barrios- Fernández, A. (2022). *Peer Effects in Education*.

Bennett, M. y P. Bergman (2021). Better Together? Social Networks in Truancy and the Targeting of Treatment. *Journal of Labor Economics, 39*(1), 1-36.

Billings, S. B., D. J. Deming y S. L. Ross (2019). Partners in Crime. *American Economic Journal: Applied Economics, 11*(1), 126-50.

Bingley, P., P. Lundborg y S. V. Lyk-Jensen (2021). Brothers in Arms Spillovers from a Draft Lottery. *Journal of Human Resources, 56*(1), 225-268.

Black, S. E., P. J. Devereux y K. G. Salvanes (2013). Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics, 31*(1), 119-153.

Bobonis, G. J. y F. Finan (2009). Neighborhood Peer Effects in Secondary School Enrollment Decisions. *The Review of Economics and Statistics, 91*(4), 695-716.

Booij, A. S., Leuven, E. y Oosterbeek, H. (2017). Ability peer effects in university: Evidence from a randomized experiment. *The Review of Economic Studies, 84*(2), 547-578.

Bursztyn, L. y Jensen, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics, 130*(3), 1329-1367.

Bursztyn, L., G. Egorov y R. Jensen (2019). Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education. *The Review of Economic Studies, 86*(4), 1487-1526.

Carlson, B. E. (2000). *Children exposed to intimate partner violence: Research findings and implications for intervention*. Trauma, Violence & Abuse, 1(4), 321-342.

Carrell, S. E., Malmstrom, F. V. y West, J. E. (2008). Peer effects in academic cheating. *Journal of Human Resources, 43*(1), 173-207.

Carrell, S. E., M. Hoekstra y E. Kuka (2018). The long-run effects of disruptive peers. *American Economic Review, 108*(11), 3377-3415.

Carrell, S. E. y M. L. Hoekstra (2010). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics, 2*(1), 211-28.

Carrell, S. E., Fullerton, R. L. y West, J. E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics, 27*(3), 439-464.

Carrell, S. E., Sacerdote, B. I. y West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica, 81*(3), 855-882.

Chetty, R., Hendren, N., Kline, P. y Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics, 129*(4), 1553-1623.

Chetty, R., Hendren, N. y Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review, 106*(4), 855-902.

Chetty, R. y Hendren, N. (2018a). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics, 133*(3), 1107-1162.

Chetty, R. y Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility II: Childhood exposure effects. *The Quarterly Journal of Economics, 133*(3), 1163-1228.

Christakis, N. A. y Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine, 357*(4), 370-379.

Cohen-Cole, E. y Fletcher, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of health economics, 27*(5), 1382-1387.

Corno, E. L. F. y J. Burns (2022). Interaction, Stereotypes, and Performance: Evidence from South Africa. *American Economic Review, 112*(12), 3848-3875.

Dahl, G. B., D.-O. Rooth y A. Stenberg (2020). *Family Spillovers in Field of Study*.

Deming, D. J. (2011). Better Schools, Less Crime? *The Quarterly Journal of Economics, 126*(4), 2063-2115.

Duflo, E., Dupas, P. y Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review, 101*(5), 1739-1774.

Dynarski, S., Libassi, C. J., Michelmore, K. y Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing

on the choices of low-income students. *American Economic Review, 111*(6), 1721-1756.

Eren, O., M. F. Lovenheim y H. N. Mocan (2022). The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy. *Journal of Labor Economics, 40*(2), 361-395.

Figlio, D. y U. ÅN Ozek (2019). Unwelcome Guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students. *Journal of Labor Economics, 37*(4), 1061-1096.

Gaviria, A. y S. Raphael (2001). School-Based Peer Effects and Juvenile Behavior. *The Review of Economics and Statistics, 83*(2), 257-268.

Goux, D. y E. Maurin (2007). Close Neighbours Matter: Neighbourhood Effects on Early Performance at School. *The Economic Journal, 117*(523), 1193-1215.

Gurantz, O., M. Hurwitz y J. Smith (2020). Sibling Effects on High School Exam Taking and Performance. *Journal of Economic Behavior & Organization, 178*, 534-549.

Hoxby, C. M. (2000). *Peer effects in the classroom: Learning from gender and race variation.*

Hoxby, C. M. y Weingarth, G. (2005). *Taking race out of the equation: School reassignment and the structure of peer effects* (No. 7867). Working paper.

Imberman, S. A., Kugler, A. D. y Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review, 102*(5), 2048-2082.

Joensen, J. S. y H. S. Nielsen (2018). Spillovers in Education Choice. *Journal of Public Economics 157,* 158-183.

Kling, J. R., Ludwig, J. y Katz, L. F. (2005). Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. *The Quarterly Journal of Economics, 120*(1), 87-130.

Karbownik, K. y U. ÅN Ozek (2019). *Setting a Good Example? Examining Sibling Spillovers in Educational Achievement using a Regression Discontinuity Design*.

Lavy, V. y Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics, 3*(2), 1-33.

Manski, C. F. (1993*).* Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies, 60*(3), 531-542.

Michelman, V., J. Price y S. D. Zimmerman (2022). Old Boys' Clubs and Upward Mobility among the Educational Elite. *The Quarterly Journal of Economics, 137*(2), 845-909.

Murphy, R. y F. Weinhardt (2020). Top of the Class: The Importance of Ordinal Rank. *The Review of Economic Studies, 87*(6), 2777-2826.

Moffitt, R. A. (2001). Policy interventions, low-level equilibria, and social interactions. *Social Dynamics, 4*(45-82), 6-17.

Nicoletti, C. y B. Rabe (2019). Sibling Spillover Effects in School Achievement. *Journal of Applied Econometrics, 34*(4), 482-501.

Qureshi, J. A. (2018a). Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital. *The Economic Journal, 128*(616), 3285-3319.

Qureshi, J. A. (2018b). Siblings, teachers, and spillovers on academic achievement. *Journal of Human Resources, 53*(1), 272-297.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly journal of economics, 116*(2), 681-704.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: two steps forward? *Annu. Rev. Econ., 6*(1), 253-272.

Vigdor, J. y Nechyba, T. (2007). *Peer effects in North Carolina public schools. Schools and the equal opportunity problem*, 73-101.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics, 85*(1), 9-23.

# Capítulo 3d
# PROFESORES: MUCHO MÁS QUE UNA PROFESIÓN

Lucas Gortazar
Esade Center of Economic Policy
Universitat Ramon Llull
ORCID: http://orcid.org/0000-0002-6785-1059

**Resumen**

Se dice que ser profesor es tener la profesión más importante del mundo. ¿Hasta qué punto es eso verdad? El objetivo de este capítulo es mostrar la enorme importancia que tiene el profesorado en la educación, el aprendizaje y desempeño de los alumnos, y su impacto a largo plazo en el mercado de trabajo y la economía. Repasamos factores como la formación del profesorado, su capital cognitivo y socioemocional, su desempeño y los canales a través de los cuales se produce su influencia sobre el alumnado, así como la importancia de sus interacciones con alumnos, familias y compañeros para su propio desempeño presente y futuro. Analizamos el papel de las políticas docentes para mejorar su atracción, retención y desempeño a lo largo de su carrera. Esto se desarrolla sobre todo con evidencia de Estados Unidos y otros países: lamentablemente la investigación económica sobre profesores en España es prácticamente inexistente. El capítulo informa por tanto acerca de uno de los factores, en particular el factor escolar más importante, que más afectan al rendimiento estudiantil y de las escuelas. Por otro lado, se presentarán algunas técnicas econométricas recientemente aplicadas a estas preguntas. También se presentan estrategias noveles de explotación y unión de datos administrativos e instrumentos de medición de la calidad y el desempeño del profesorado.

## 1. INTRODUCCIÓN: LA PROFESIÓN MÁS IMPORTANTE DEL MUNDO

Los profesores son los arquitectos de las sociedades y las economías futuras. Transmiten en la escuela los conocimientos, habilidades y valores a las siguien-

tes generaciones: son la base del capital humano y el conocimiento futuro de las economías. Los países que son capaces de hacer la profesión docente mucho más que un trabajo son los que han logrado grandes cambios en la educación y en la acumulación de capital humano y conocimiento y, por tanto, en la generación de riqueza y trabajo. Y no es tarea fácil. Ser profesor no se trata solo de estar académicamente preparado y, por tanto, de haber sido un buen alumno. Es mucho más que eso: implica observar, diagnosticar, comunicarse, organizar, ser ejemplo, facilitar, gestionar o liderar. Según una investigación de los años 60 del siglo pasado que tuvo una gran popularidad, los profesores toman de media entre 1.200 y 1.500 decisiones al día, entendiendo decisiones como elecciones no rutinarias e improvisadas (Jackson, 1968). Ser profesor se trata sobre todo de una profesión asociada a una forma de vocación social y transcendencia más allá de uno mismo.

La profesión docente ha sido un foco de interés creciente en el campo de la economía de la educación durante las últimas décadas. Un repaso de lo que ha ocurrido en Estados Unidos nos permite entender por qué ese foco ha aumentado, por qué ha fortalecido a la investigación económica y por qué esta nos arroja luz y resultados relevantes sobre la importancia del profesorado para la educación de millones de niños.

En 1983, una comisión nombrada por el Secretario de Educación Terrel Bell, en la administración de Ronald Reagan, publicó "*A nation at risk: the imperative for education reform*" (National Commission on Excellence in Education, 1983). El informe es considerado uno de los documentos más relevantes en la historia de la educación de Estados Unidos por las repercusiones que tuvo posteriormente. El informe, que interpelaba al sentido de urgencia de las administraciones educativas como si de una situación de guerra y amenaza nacional se tratara, ponía el foco en varios temas, pero uno de ellos era el profesorado. Hablaba de salarios, carrera profesional e incentivos, y de cómo esos incentivos eran clave para que el sistema escolar mejorara en su conjunto. Pocos años después, el gobierno federal y varios estados ya estaban comenzando a poner en marcha programas de incentivos para atraer, evaluar y mejorar el desempeño de los profesores en las escuelas de todo el país.

El año 2001 marcó un antes y un después con la aprobación de la ley "No Child Left Behind" ("Que ningún niño se quede atrás"), con apoyo de Republicanos y Demócratas y promovido por la administración de George Bush. Esta

reforma supuso el espaldarazo definitivo para promover incentivos más fuertes para escuelas y docentes, una de las prioridades históricas del partido Republicano, entonces en el gobierno federal. Pero también puso el foco en que dichos cambios supusieran una mejora los resultados de los alumnos, especialmente minorías y alumnado de renta más baja, una de las prioridades históricas del partido Demócrata, y que hizo que este también rubricara su firma en la ley. También supuso un espaldarazo a la medición de los aprendizajes con herramientas rigurosas, a la promoción de la investigación educativa y más en general, al uso de investigación educativa basada en métodos empíricos causales para la toma de decisiones de escuelas y administraciones.

Estados Unidos es el ejemplo paradigmático de lo ocurrido, pero no es el único lugar donde poner el foco. En otros países como Reino Unido, Canadá, Holanda, países nórdicos como Finlandia o países asiáticos como Japón, China, Singapur o Vietnam ha ocurrido algo semejante. Cada vez hay más ojos puestos en la profesión docente y cada vez más economistas se han ocupado de estudiarla. En 2007, la prestigiosa empresa de consultoría americana McKinsey publicó un informe (Barber y Mourshed, 2007) sobre los sistemas educativos internacionales que causó un gran impacto en los medios de comunicación y los gobiernos. Se titulaba "*Cómo hicieron los sistemas educativos con mejor desempeño del mundo para alcanzar su objetivo*" y apuntaba a la experiencia de algunos de estos países: todos ellos tenían un denominador común, y es el de haber apostado por la profesión docente como una profesión de excelencia y como pilar de cambio de sus sistemas escolares. Una frase de un gobernante surcoreano, recogida en el informe, tuvo enorme repercusión y se convirtió en un mantra para muchos: "*la calidad de un sistema escolar no puede exceder a la calidad de sus docentes*". El techo de un sistema escolar lo marcan los docentes. Aunque es una frase que cuestionaremos en el capítulo con algunos resultados recientes de la investigación (¿Es la suma mayor o no a las partes?), sí dejó un mensaje que caló entre muchos. El informe trasladó a empresas, organismos multilaterales, grandes fundaciones y gobiernos la prioridad total en enfocarse en los docentes. Y fue otro ejemplo del empujo para situar a la profesión docente en el centro de la investigación económica.

Desgraciadamente, la investigación económica sobre profesores en España es prácticamente inexistente. Lo es por la falta de tradición de la economía de la educación en España, con una comunidad investigadora todavía incipiente, a la que este libro tratará hacer crecer. Pero también por la falta de acceso a datos

administrativos y de encuestas relacionadas con profesores, que se enmarcaría muchas veces en proyectos de unificación de bases de datos no preparadas para investigación: esto además aumenta los costes de entrada (y vuelve la investigación más costosa y cara), unos costes difíciles de reducir hasta que los datos no se vuelven accesibles, tratables y fáciles de manipular con fines investigadores.

En este capítulo se tratará de mostrar la importancia de la profesión docente desde el punto de vista de la economía. Algunos podrían preguntarse: ¿Por qué la economía debe ocuparse de estudiar y pensar la profesión docente y no dejarla a otras disciplinas, como la pedagogía, la psicología o la sociología? Hay poderosas razones para hacerlo.

La primera, de la que se ocupa todo este libro, es que la educación es un factor que impacta en la economía, el mercado de trabajo y las desigualdades; si la profesión docente es uno de los factores (sino el factor más importante) de la función de producción de la educación, entonces merece ser abordada en toda su profundidad. Y eso no es solo una frase o eslogan, sino que lo reflejan el uso de los recursos públicos y también privados (de las familias y las empresas). En concreto, buena parte del presupuesto público destinado a educación va dirigido a pagar los salarios de los docentes. En los países desarrollados, el 91% del gasto público en educación va dirigido a gasto recurrente, del cual un 74% de ese gasto va destinada a pago de salarios (esto es, casi un 70% del gasto público educativo se destina a los docentes). Esta cifra que es todavía mayor (81% del gasto recurrente, cuando este supone un 95% del gasto total, esto es un 77% del total) para el caso de España (OCDE, 2022), una cuestión también mencionada en el capítulo 3 de este libro. Y lo es todavía superior en los países, superando fácilmente el 80% del gasto educativo público. Por tanto, hay razones económicas para pensar en los docentes como una parte central en el devenir de la economía en el largo plazo.

La segunda es que los sistemas escolares son una fuente de empleo público y privado que puede llegar al 5% del mercado de trabajo en países desarrollados y ser todavía superior en países en desarrollo: según datos más recientes del Ministerio de Educación y Formación Profesional, en el curso escolar 2022-2023 hubo 811.049 docentes contratados entre centros públicos y privados en enseñanzas no universitarias, lo que representa un 4,3% de todo el mercado de trabajo español ese año. Entender dicho mercado de trabajo y sus características es clave. Si los economistas se ocupan del mercado de trabajo

como una de sus áreas de mayor interés para comprender el funcionamiento de la economía, entonces el mercado de trabajo de la profesión docente cobra un protagonismo enorme.

La tercera razón tiene que ver con la economía política de las reformas educativas. Los docentes son un colectivo habitualmente organizado (tanto sindical como profesionalmente), con capacidad de movilización e influencia y, por tanto, con un margen importante para influir en las políticas y los resultados educativos (Lovenheim y otros, 2019; Biasi y Sarsons, 2022). Entender sus incentivos, así como sus creencias, formas de organización y participación política es clave cuando se elaboran políticas públicas. Son siempre necesarios en toda reforma educativa ambiciosa; y, de la misma forma, tienen el poder y la capacidad de frenar los cambios políticos en educación si así lo consideran.

El capítulo repasa toda la investigación económica reciente que se ha venido ocupando del profesorado. Los ejemplos utilizados son sobre todo de países desarrollados, aunque una literatura incipiente también se está llevando a cabo en países en desarrollo. La estructura del capítulo es la siguiente. El capítulo 2 sitúa al profesorado como un elemento central en la construcción y creación de capital humano a través de la educación, mostrando su efecto en las habilidades cognitivas y socioemocionales de los alumnos (y las escuelas) y más en perspectiva, en el futuro del mercado laboral y la economía. El capítulo 3 aborda la pregunta sobre qué es ser un buen docente y cómo podemos identificarlo a través de las Ciencias Sociales en general y de la investigación económica en particular; también intenta comprender por qué unos profesores son más eficaces que otros y finalmente aborda la pregunta de cómo podemos potencial una mejor selección, atracción, formación y desarrollo profesional docente en las escuelas. El capítulo 4 aborda de manera breve el papel que juegan los equipos directivos y directores de escuela en el aprendizaje de los alumnos.

## 2. PROFESORADO Y CAPITAL HUMANO

### 2.1. El profesorado desde una perspectiva del capital humano

Pensar en los profesores desde una perspectiva de capital humano tiene sentido económico. Desde el modelo de Becker (Becker, 1962; Becker y Chiswick, 1966), se ha desarrollado un amplio campo de investigación y pensamiento económico para estudiar el capital humano como factor de producción de la

economía, medido habitualmente a través del nivel educativo, las credenciales académicas y la experiencia profesional.

En educación, los modelos econométricos relacionaron las características docentes como input o variable explicativa y el aprendizaje de los alumnos como output o resultado. Muchas de las investigaciones de la segunda parte del siglo XX y el comienzo del siglo XXI en el campo de la economía de la educación trataron de identificar la relación entre habilidades docentes y aprendizaje de los alumnos: mostraron claramente que, en el contexto de un país desarrollado y con educación universal y gratuita como Estados Unidos, las características habituales (título universitario, por ejemplo) que miden el capital humano del profesorado no eran capaces de explicar diferencias de resultados de aprendizaje y graduación entre los alumnos. Esto ocurría a pesar de que en el campo de la econometría los modelos causales cobraban fuerza y en un contexto en el que los investigadores tenían cada vez mayor acceso a mejores datos administrativos longitudinales.

### *Einstein: Valor añadido docente*

Por ejemplo, Rivkin, Hanushek y Kain (2005) llevaron a cabo un modelo de valor añadido (que entonces ya estaba popularizando en economía) del aprendizaje de los alumnos, por primera vez con una gran base de datos de algo más de 80.000 observaciones de alumnos del estado de Texas que incluía información rica y precisa de escuelas y profesores emparejados a estos alumnos. Estos modelos toman habitualmente la siguiente forma (ecuación 1):

$$A_{ijgs} = A_{ij(g-1)s} + X_i \beta_X + T_{ij} \beta_T + S_{si} X_S + f_i + \varepsilon_{cijgs}$$

Donde, el aprendizaje $A_{ijgs}$ del alumno $i$, con el docente $j$, en la escuela $s$ y el curso $g$ depende del aprendizaje de ese mismo alumno en el curso $g-1$, un vector $X$ de características socioeconómicas y personales del alumno, un vector $T$ de características del profesor $j$ del alumno $i$, un vector $S$ de características de la escuela $s$ del alumno $i$, un parámetro $f$ que refleja las habilidades inherentes y características no observables del alumno $i$ y un error $\varepsilon_{cijgs}$. La posibilidad de aplicar diferencias entre los aprendizajes presentes y pasados permite eliminar los sesgos introducidos por las características que no observamos del alumno $i$, lo cual supone una mejora metodológica significativa: esto permite, por ejemplo, controlar por diferencias en esfuerzo, la predisposición al aprendizaje, el aprendizaje previo, o factores como el capital genético o cultural transmitido por la familia.

Los autores identifican una importante variación de la calidad docente, medida a partir del modelo, pero que dicha variación no parece deberse al nivel educativo (en particular, tener un Máster en comparación con tener solo un Grado universitario) o a los años de experiencia en la profesión (siendo solo el efecto de los primeros años ligeramente significativo). Con datos más recientes de la evaluación de competencias en adultos (PIAAC) de la OCDE, Hanushek y otros (2019) introdujeron métricas de habilidades cognitivas (medidas por pruebas estandarizadas de Matemáticas y comprensión lectora) en la estimación de la relación entre capital humano del profesorado y aprendizaje de los estudiantes para 31 países: encontraron que estas medidas explicaban parte de la variación de los niveles de aprendizaje de los alumnos entre países y qué parte de esas diferencias de habilidades se debían al acceso a ocupaciones de altas habilidades en el mercado laboral por parte de las mujeres.

### *Más allá de las habilidades cognitivas*

Hasta cierto punto, los resultados pueden parecer evidentes. El capital humano del profesorado no va solo de diplomas, certificados o años en la profesión. Va también de lo que uno sabe y puede enseñar. Pero ¿es solo conocimiento cognitivo o hay más? En los últimos años, el acceso a información individual y anonimizada de alumnos y docentes se ha multiplicado en muchos países, especialmente en Estados Unidos. En particular, a los resultados de aprendizaje de los alumnos se han sumado otro tipo de outputs o resultados como el comportamiento, las ausencias o faltas en la escuela, el nivel de absentismo, y también, a través de datos posteriores, la graduación y continuación en estudios postobligatorios y finalmente la participación en el mercado de trabajo. Varios trabajos recientes han venido a resaltar la importancia de las habilidades no-cognitivas en el profesorado (algo que recoge el capítulo 5d del libro, firmado por Ildefonso Méndez y Gemma Zamarro) y su capacidad de transmitirlas al alumnado, y cómo eso tiene un impacto en el largo plazo.

La reflexión para comprender mejor el papel de las habilidades no cognitivas en los docentes surgió a partir estudios recientes que mostraron que incluso la correlación entre habilidades cognitivas y efectividad docente (medida como el valor añadido que hemos explicado previamente) era también débil. Por ejemplo, Zamarro y Cheng (2018) muestran que las habilidades no cognitivas del profesorado como el esfuerzo, el liderazgo o la resiliencia, también tienen un impacto sobre el aprendizaje. Realizan este estudio a partir del proyecto

Measures of Effective Teaching (MET), de la fundación Bill y Melinda Gates, otra gran Fundación que apoyó la investigación sobre profesores, en la que miles de profesores y alumnos eran emparejados de forma aleatoria y monitorizados a lo largo del tiempo. En otros contextos, por ejemplo Suecia, Grönqvist y Vlachos (2016) encontraron que detrás de un impacto nulo de las habilidades docentes sobre el aprendizaje del alumnado, se encontraban resultados muy dispares por tipo de alumnado: mientras que los docentes con habilidades cognitivas elevadas tenían un impacto académico mayor entre el alumnado con mayores aptitudes académicas, el profesorado con habilidades sociales altas tenían un impacto académico mayor sobre el alumnado con peores aptitudes académicas. Para el caso de Finlandia, Izadi (2019) muestra, a partir de una base de datos administrativos de todos los universitarios que acceden a la profesión docente, que los resultados académicos de los profesores no están relacionados con otras habilidades (no cognitivas) que también impactan en los alumnos. En conclusión, este conjunto de resultados hace necesario salir de la premisa de que, como anticipábamos en la introducción, un buen docente no es solo un buen alumno, sino que las habilidades cognitivas parecen ser una condición necesaria, pero no suficiente. Reducir, por tanto, la medición de la calidad docente a partir de sus resultados académicos en Secundaria o Universidad parece ser insuficiente para capturar su desempeño en la escuela y el aula y su impacto en los alumnos en el corto y largo plazo.

### 2.2. ¿Por qué el profesorado es tan importante para el desempeño de los alumnos?

Hemos visto que el profesorado importa para el aprendizaje, que la manera en la que lo hace es múltiple, y que las habilidades requeridas son variadas. La pregunta siguiente es saber cómo de importante es ese efecto en relación con otros factores de la función de producción educativa y en qué se materializa el efecto en el **corto** (aprendizaje y desarrollo de habilidades sociales y emocionales), **medio** (graduación y continuación de estudios) y **largo** (participación en el mercado de trabajo, indicadores de salud o criminalidad) plazo para los alumnos.

En esta sección hemos presentado tanto las nuevas métricas de calidad docente como las de valor añadido. ¿Qué capacidad tienen de predecir el impacto futuro sobre el alumnado? Estudios recientes muestran que tienen un efecto económicamente grande en el aprendizaje futuro y retornos en el mercado de

trabajo. En un trabajo original, Chetty, Friedman y Rockoff (2014), con medidas de calidad docente, encuentran un impacto relevante y positivo a largo plazo, esto es, en graduación, acceso a la Universidad y retornos en el mercado laboral. Los autores miden el impacto de las métricas de valor añadido (del aprendizaje a los 14 años) sobre indicadores de largo plazo: encuentran que una desviación estándar de valor añadido docente aumenta la probabilidad de estar en la Universidad a los 20 años en 0,82 puntos porcentuales (con respecto a una media de 37%). También aumentan los ingresos futuros ya que, a los 28 años, un aumento en una desviación estándar de valor añadido aumenta en un 1,3% los ingresos anuales, el equivalente a 39.000 dólares en ingresos agregados a lo largo de la vida. También se observa un impacto positivo reduciendo la probabilidad de tener un hijo en edades adolescentes, un aumento la calidad del barrio en que se vive en vida adulta y la participación en planes de ahorro para pensiones.

En un caso real del Distrito de Columbia en Washington, se puso en marcha un sistema de evaluación docente a partir de estimaciones de valor añadido: los docentes con valor añadido más bajo eran sustituidos por otros docentes si no lograban mejorar su desempeño. Esta política supuso una mejora de entre 0,14 (Lectura) y 0,21 (Matemáticas) desviaciones estándar de aprendizaje (Adnot y otros, 2016). Como explica Hanushek (2011), este tamaño es superior al que puede lograrse con alternativas habituales, mucho más costosas en términos de presupuesto público, como, por ejemplo, la reducción del tamaño de clase.

Tres artículos de investigación muy reciente han complementado las métricas de valor añadido sobre el aprendizaje habitualmente utilizadas: lo han hecho con métricas de valor añadido docente en resultados no cognitivos, esto es, sociales y emocionales, y han sido capaces de demostrar que eso conlleva enormes efectos en la vida adulta. Por ejemplo, Jackson (2018) muestra de manera causal que la influencia del profesorado también se produce en indicadores no estandarizados, como absentismo, faltas, notas de clase o repetición de curso. Además, el efecto de los docentes sobre esos indicadores en Secundaria (a los 15 años) predice resultados posteriores, como la finalización de la Educación Secundaria (a los 18 años), y lo hacen con mayor fuerza que el valor añadido sobre los resultados de aprendizaje. Incorporar al efecto de los docentes no solo resultados de tests sino otro tipo de resultados como los presentados en el estudio implica una capacidad predictiva a largo plazo mayor al doble de la capacidad predictiva cuando solo se tienen en cuenta los resultados de los tests en la medición del valor añadido docente.

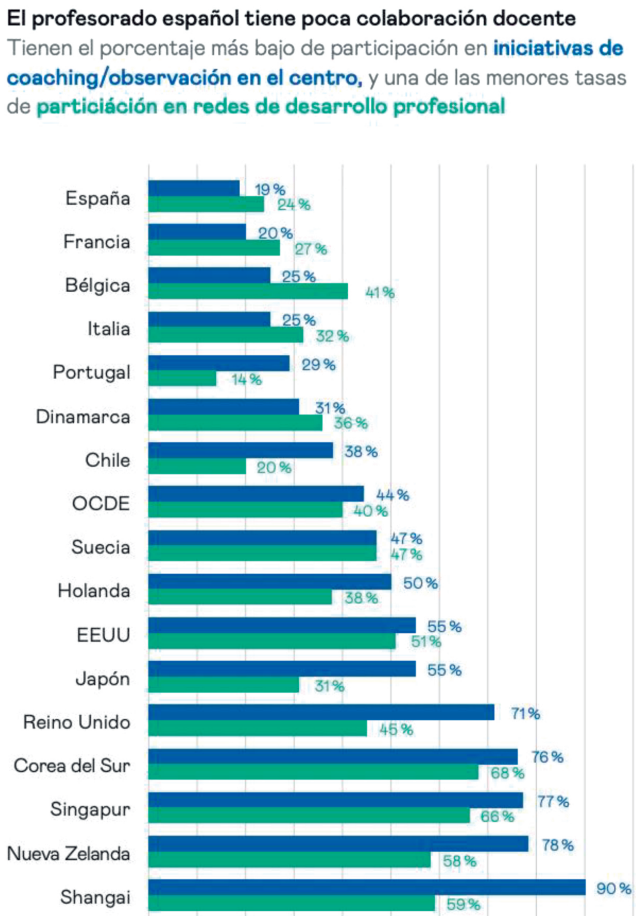
En un estudio todavía más reciente, Petek y Pope (2023) responden a una pregunta similar a partir de los datos administrativos de más de 100.000 alumnos de las escuelas públicas de la ciudad de Los Ángeles, con información muy específica de habilidades no cognitivas de docentes y alumnos. Encuentran algo semejante a lo aportado por Jackson (2018), pero añaden tres resultados más: la efectividad docente en los tests de aprendizaje no correlaciona con la efectividad docente en resultados de comportamiento y no cognitivos, por lo que plantean una métrica de efectividad docente (vía un mayor impacto a largo plazo) mucho más precisa que aquellas que solo usaban exámenes: apuntan a la importancia de las políticas para favorecer docentes con una combinación de habilidades, no solo aquellos que son capaces de lograr lo mejor de sus estudiantes desde el punto de vista académico. También encuentran que los modelos de valor añadido para las asignaturas de Lectura y Matemáticas son mucho mejores predictores de resultados futuros, comparado con otras asignaturas. Finalmente, encuentran que los docentes de Secundaria con alto valor añadido (tanto en tests como en resultados no cognitivos) son mucho más decisivos que los de Primaria en cuanto a los resultados de largo plazo. Los resultados de graduación y mercado laboral no son los únicos a través de los cuales los docentes impactan en sus estudiantes. Por ejemplo, Rose et al. (2022) encuentran un efecto positivo de los docentes en la reducción de la criminalidad (asociada a menos absentismo y más graduación), un efecto no relacionado con el efecto en términos cognitivos, lo que apunta de nuevo a las complementariedades en la efectividad docente y sus dispares efectos a largo plazo.

Esta complementariedad de efectos puede también observarse cuando trasladamos el argumento de la efectividad docente a la escuela. Resaltamos aquí, por su relevancia para este libro, el trabajo de Loviglio (2019) que muestra, con datos administrativos de la ciudad de Barcelona, que en efecto, parece haber esa especialización en valor añadido (aunque a nivel escuela): las escuelas capaces de mejorar el rendimiento académico de los alumnos (condicionado a que estos se gradúen de la Secundaria obligatoria) son las que reciben alumnado de rentas altas; en cambio, las escuelas que maximizan la probabilidad de graduación y continuar los estudios atienden habitualmente al alumnado de renta baja. Estos resultados muestran que evaluar y comparar escuelas usando solo resultados de exámenes y pruebas externas no es suficiente para identificar y mejorar el éxito escolar del alumnado de rentas más bajas.

## 2.3. Profesores, escuelas y su impacto más allá del aula

Hasta ahora hemos identificado el impacto de los docentes en el aprendizaje y el desarrollo socioemocional de los alumnos y hemos tratado de comprender cómo esto afecta a los resultados a largo plazo, incluyendo resultados económicos. Detrás de este planteamiento hemos asumido que el impacto de los docentes se materializa en el espacio físico de la clase o el aula y que la organización escolar habitual es la de un profesor en un aula con un grupo de alumnos ("*egg-crate classroom*" en inglés –ver Corbett (1982)–, "*aula huevera*" en castellano –ver Fernandez Enguita (2018)–), en vez como un individuo en una institución (la escuela) donde su impacto va más allá de su propia aula.

**Gráfico 1. Colaboración entre el profesorado.**




Fuente: OCDE (2019a), tabla 1.5.7 a partir de TALIS 2018.

Mientras que los estudios de sociología y pedagogía han abordado la cuestión de la profesión docente como un fenómeno en red, esta cuestión llegó al campo de la economía hace no tanto tiempo. Sin embargo, las aportaciones de la última década han sido determinantes, siendo esto especialmente relevante en España, un país donde, como vemos en el gráfico 1 a continuación, la interacción y colaboración entre docentes es muy baja en comparación con otros países.

### *Einstein: Impacto de los docentes en sus pares*

El trabajo, de nuevo de Jackson, esta vez con Bruegmann (2009), supone el estudio más importante en el campo de la economía donde se aborda el impacto de los profesores más allá de los alumnos (pero con un impacto indirecto importante). Siguiendo el ejemplo de valor añadido, ¿cómo puede medirse el impacto docente sobre los pares? En particular, ¿cómo podemos controlar por el hecho endógeno de que los docentes no se interaccionan con sus pares de manera aleatoria? Para evitar incurrir en dichos errores de autoselección, los autores incorporan información de docentes pares en una ecuación como la que hemos descrito anteriormente (ecuación 1), pero calculan

$$A_{ijgs} = A_{ij(g-1)s} + X_i \beta_X + T_{ij} \beta_T + S_{si} X_S + \mu P_{ij} + f_i + \varepsilon_{cijgs}$$

donde $P_{ij}$ es una medida de calidad docente de los pares. Por tanto, plantean el siguiente modelo:

$$A_{ijgs} = A_{ij(g\text{-}1)s} + X_i \beta_X + T_{ij} \beta_T + S_{si} X_S + f_i + \xi_{js} + \varepsilon_{cijgs}$$

donde $\xi_{js}$ es un efecto fijo de docente y escuela, que elimina los potenciales sesgos y desaparece al calcular el efecto en el cambio de pares sobre el cambio de resultados de los alumnos. Gracias a lo anterior, los autores identifican, a partir de datos administrativos de Carolina del Norte, y de forma causal, las **externalidades positivas en el valor añadido docente** cuando sus escuelas reciben docentes efectivos. También encuentran que estos efectos son mayores para docentes con poca experiencia, llegando a explicar el 20% del efecto de un docente en el largo plazo. Mientras que esta cuestión había sido ampliamente estudiada en la economía de las aglomeraciones y ciudades, este artículo cambió la perspectiva y amplió la visión sobre la profesión docente a algo mucho mayor que lo que un docente hace con sus alumnos. De nuevo, volvemos al principio: estamos hablando de una profesión de alto valor añadido, un valor que se transmite cuando hay interacción entre pares.

Decíamos en la introducción que la evidencia de la economía contestaría la frase del político surcoreano sobre los profesores como el techo de los sistemas educativos. Quizás los resultados de este estudio nos obligan a refutar la frase y complementarla de tal manera que "*la calidad de un sistema escolar no puede exceder a la calidad de* ***la suma*** *sus docentes en* ***una escuela***". Después de este artículo ha llegado toda una línea de investigación para ampliar. Sun, Loeb y Grissom (2017) encuentran algo semejante a los resultados de Jackson y Bruegmann en la ciudad de Miami.

En un proyecto experimental donde docentes de alto y bajo desempeño (valor añadido) eran emparejados a trabajar conjuntamente, Papay y otros (2020) encuentran un impacto positivo en 0,12 desviaciones estándar para los docentes de bajo valor añadido, asociadas a la mejora en las habilidades específicas en las que el docente tenía peores resultados. Por su parte, Burguess, Rawal y Taylor (2021) exploran los mecanismos de dicha interacción en un experimento aleatorio controlado donde los docentes observan y son observados por sus pares: encuentran que los alumnos mejoraron un 7% de una desviación estándar, siendo los efectos mayores para los docentes de bajo desempeño. En un estudio similar, Murphy y otros (2021) encuentran un efecto nulo, pero al desagregarlo por características de la escuela, muestran que el efecto es heterogéneo por tamaño de la escuela: en las escuelas con más de una clase por curso, el efecto es positivo (y semejante al de Burguess, Rawal y Taylor, 2021), mientras que en las escuelas con una sola clase por curso, el efecto es negativo y significativo: estos resultados tienen sentido desde el punto de vista de oferta, pues el tamaño correlaciona con una mayor varianza de calidad docente dentro de la escuela, lo cual estaría asociado a mayores posibilidades de aprendizaje entre pares en promedio. En resumen, parece que el campo de actuación de los profesores no puede circunscribirse al aula, sino que pueden tener un impacto importante en sus pares a través de sus interacciones y eso puede incrementar su propia calidad con los alumnos.

## 3. PROFESORADO: QUIÉNES SON, DÓNDE ESTÁN Y EL PAPEL DE LAS POLÍTICAS PÚBLICAS

### 3.1. ¿Cómo identificar un buen profesor?

Se dice que el profesor perfecto no existe. Que cada maestrillo tiene su librillo. Y que tratar de medir una profesión tan compleja es osado y probablemente re-

duccionista. Quizás haya algo de verdad en todo ello, pero la investigación, y en particular la ciencia económica, han tratado de aportar y arrojar luz para mejorar nuestra comprensión acerca de quién es y cómo identificar un buen docente. Los resultados de los últimos años son verdaderamente importantes. A continuación, hacemos un repaso de lo que hemos aprendido sobre cómo se puede medir la efectividad docente con distintos instrumentos, qué información específica y matices nos aporta cada instrumento y cómo se complementan entre sí.

1) Valor añadido docente cognitivo y no-cognitivo: en la sección anterior, hemos visto las distintas formas de estimar el valor añadido (esto es, el progreso a partir de un punto concreto de aprendizaje y otros indicadores, controlando por factores del alumno, docente y escuela) a partir de datos longitudinales de desempeño cognitivo y no-cognitivo del alumnado y características individuales. Las métricas de valor añadido más recientes a partir del trabajo Chetty, Friedman y Rockoff (2014) se han extendido, como hemos explicado, a resultados no cognitivos (Jackson, 2018) y también se han llevado a cabo en países en desarrollo, por ejemplo, con el caso de Pakistan (Bau y Das, 2020). Tienen la propiedad estadística de ser estimadores no sesgados (Chetty y otros, 2017). A pesar de ser aceptada como métrica en el campo la economía, su aplicación para la evaluación docente ha tenido algunas dificultades que comentaremos más adelante.

2) Observación de aula: mientras que las métricas de valor añadido aportan estimaciones precisas del resultado de aprendizaje derivado de los profesores, no dicen nada del proceso que transcurre dentro del aula para explicarlo. Esta ausencia de información ha venido a popularizar el término de "la caja negra de la educación", es decir, toda la relación entre inputs y outputs de la función de producción educativa. Desde el ámbito de la pedagogía y los métodos de investigación en educación, en las últimas décadas se han venido desarrollando instrumentos **fiables** (el instrumento da los mismos resultados independientemente del observador) y **válidos** (el instrumento está asociado a buenos resultados de los alumnos en una determinada competencia de aprendizaje). Los más conocidos son el Classroom Assessment Scoring System (CLASS), centrado en las interacciones entre docentes y alumnos, el Framework For Teaching (FFT), inspirado en una visión constructivista del aprendizaje, o los instrumentos específicos para Lectura (PLATO) o Matemáticas (MQI). La investigación más reciente ha mostrado que estos métodos son capaces de predecir aprendizaje futu-

ro y correlacionan positivamente con las medidas de valor añadido (Tyler et al., 2010; Araujo et al., 2016) aunque son caros de implementar. Todos estos instrumentos tienen la enorme ventaja de desengranar lo que ocurre en el aula y traducir lo que se observa en cambios en la enseñanza mediante formación al profesorado.

3) Encuestas de alumnado: estas encuestas, completadas por los alumnos, recogen información sobre su opinión acerca del trabajo de sus docentes. Tienen la ventaja de ser más baratas que los instrumentos de observación. Sin embargo, al evaluarlas y compararlas con las medidas de valor añadido, muestran que no tienen una correlación fuerte con estas métricas (Bacher-Hicks et al., 2019).

4) Impacto en los pares y en el centro: tal y como hemos explicado anteriormente, la calidad de los docentes no solo debe reducirse al aula, pues la evidencia muestra que los docentes efectivos también son capaces de tener un impacto indirecto en otros alumnos a través de sus pares en la escuela (Jackson y Bruegman, 2009; Papay y otros, 2020).

¿Cómo podemos construir medidas completas, fiables y con alto valor predictivo de la calidad docentes? Los estudios de Kane y Staiger (2012) y Bacher-Hicks et al. (2019) muestran que la combinación de medidas, especialmente observaciones de aula y métricas de valor añadido, permitirían aumentar la capacidad predictiva de los instrumentos por separado y a su vez ayudar a incrementar la precisión entendiendo lo que ocurre dentro del aula. A su vez, continuar explorando el impacto que los docentes tienen en sus pares y sus escuelas debería estar en la ecuación para definir "qué es ser un buen docente".

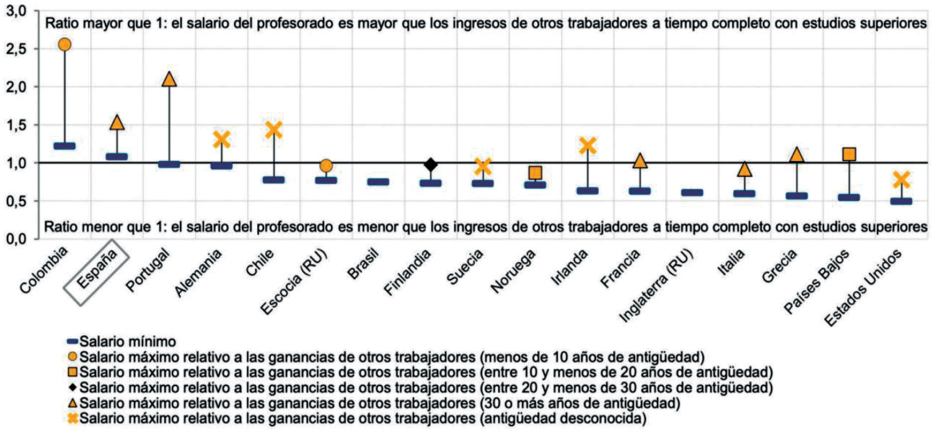
### 3.2. ¿Qué impacto tienen las condiciones laborales de los docentes sobre su efectividad?

En la subsección anterior hemos explorado qué define un buen docente y cómo identificarlo. En esta parte nos centramos en factores contextuales de la profesión docente independientes de sus cualidades y características individuales. En particular, repasamos la estructura del mercado de trabajo de los profesores, las condiciones laborales que tienen en las escuelas y el poder de negociación de los sindicatos docentes: el objetivo es comprender de qué forma afectan a la efectividad docente, a su desarrollo profesional y, en último término, promueven un mayor aprendizaje de los alumnos.

## *El mercado de trabajo del profesorado*

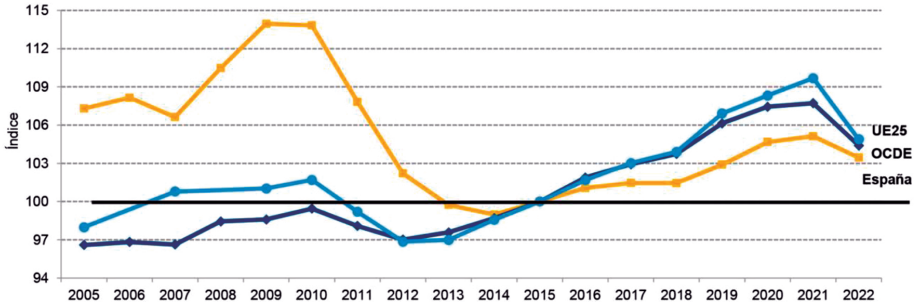
Como hemos apuntado al comienzo del capítulo, el mercado de trabajo docente es amplio en volumen y abarca habitualmente a entre el 4% y el 5% del empleo en países desarrollados, pudiendo ser superior en países en desarrollo. Uno de los indicadores más obvios para comprender el mercado de trabajo es el de los salarios docentes de las escuelas públicas, habitualmente fijados por los gobiernos mediante negociación con los sindicatos de profesores. En España, el salario docente es comparativamente superior al de otros trabajadores con estudios superiores, como puede apreciarse en el gráfico 2 a continuación (véase también el capítulo 1 de Indicadores), ya desde el salario de entrada a la profesión, aunque es cierto que el salario máximo al que pueden acceder no crece demasiado, lo cual refleja una carrera profesional más bien plana en comparación con otros países. Pero el reverso de esa realidad, como muestra el gráfico 3, es que los docentes españoles han experimentado una pérdida de poder adquisitivo desde la Gran Recesión, como consecuencia de una moderación de sus salarios por debajo de la inflación y sin llegar a recuperar los niveles de 2010. A esto se suma el añadido del aumento de los precios de 2022 debido a la guerra de Ucrania, algo que puede continuar en 2023 y 2024.

**Gráfico 2. Relación entre el salario del profesorado y el de quienes trabajan a tiempo completo teniendo estudios superiores.**




Fuente: Ministerio de Educación y Formación Profesional (2023). Salarios iniciales y Máximos reglamentarios relativos del profesorado de la segunda etapa de Educaciónsecundaria y años necesarios para alcanzar el máximo de la escala (2022).

**Gráfico 3. Evolución del salario (descontando inflación) del profesorado de Secundaria con la cualificación más frecuente, entre 2005 y 2022.**




Fuente: Ministerio de Educación y Formación Profesional (2023). Índice de cambio entre 2005 y 2022 del salario reglamentario del profesorado de la segunda etapa de Educación Secundaria con 15 años de experiencia (2015=100), convertido a precios constantes utilizando deflactores para consumo privado.

¿Qué dice la investigación sobre los mercados laborales docentes en relación con los salarios? Desde el punto de vista de la oferta, parece claro que, al tratarse de una carrera profesional igual para todos los docentes, y, por tanto, con condiciones laborales de partida similares, esas condiciones van a influir en la oferta de empleo por parte de los trabajadores que se incorporan en el mercado de trabajo. Eso es lo que encuentran Leigh (2012) y Gjefsen (2020). El trabajo de Leigh (2020) con datos de profesores en Australia es muy relevante, pues permite separar los efectos de oferta y demanda a la hora de identificar el efecto de los salarios en la calidad de los docentes (medida por los resultados de pruebas de los admitidos a las universidades australianas y en particular los programas universitarios para acceder a la profesión docente). El autor encuentra que el aumento de un 1% en el salario de entrada de los docentes mejora la aptitud académica de los docentes en 0,6 percentiles. También encuentra que, a mayor variación en el mercado laboral fuera de la docencia, menor es la calidad de los docentes potenciales. En un estudio para la ciudad de Oslo (Noruega), el trabajo de Gjefsen (2020) muestra de nuevo a través de una metodología de Diferencias en Diferencias que, tras un aumento de los salarios del 5%, se produjo un incremento en la entrada de profesores con mayores credenciales académicas (título de Máster) pero sobre todo, y más importante, una mejora en los resultados académicos de los alumnos: esto último sugiere que los nuevos docentes tenían una efectividad mayor, además de mayores credenciales académicas.

El salario de entrada parece ser un factor importante, aunque es probable que el efecto sobre la oferta docente tenga rendimientos decrecientes y otros factores entren en juego una vez alcanzado un determinado nivel salarial. ¿Qué otros factores importan? Como apuntaba el estudio de Leigh (2012), los salarios no deben analizarse en términos absolutos, sino relativos. El contexto del mercado de trabajo es relevante: por ejemplo, Bacolod (2017) documenta, a partir del modelo de Roy (1951) de autoselección en ocupaciones, la caída de la calidad docente en Estados Unidos de las mujeres jóvenes que se incorporan a la profesión debido al aumento de las oportunidades en empleos de alto valor añadido. Bajo esa misma lógica, Nagler y otros (2020) encuentran, con datos del caso del estado de Florida, y explotando la variación ciclo económico como fuente de exogeneidad que altera las alternativas de los docentes potenciales, que los profesores que entran en la profesión durante las recesiones son significativamente más efectivos en mejorar los resultados de sus estudiantes. Marchand y Weber (2020) apuntan al mismo argumento con otro ejemplo: el shock positivo que experimentó el estado de Tejas con la perforación de petróleo y gas redujo los resultados de aprendizaje y la asistencia a clase, siendo el mecanismo la mayor rotación docente y la menor experiencia de los nuevos docentes, cuyos salarios (en términos relativos) no se vieron mejorados por el shock positivo de la economía, a diferencia de otros sectores.

Otros elementos del mercado de trabajo también influyen los resultados de aprendizaje y la calidad docente. La falta de liquidez para los alumnos que provienen de hogares de rentas más bajas puede ser un freno para sumarse a la profesión. El estudio de Coffman y otros (2019) muestra que las barreras de acceso a una profesión peor pagada que otras supone a corto plazo una falta de liquidez imposible de solventar para algunos candidatos. Finalmente, las condiciones laborales en el centro influyen. El estudio de Ronfeldt y otros (2013) muestra que la excesiva rotación docente en las escuelas perjudica el aprendizaje.

Finalmente, cabe preguntarse los condicionantes de las condiciones laborales, pues en la medida en que estas últimas influyen en los resultados académicos, son relevantes. En los últimos años se ha ampliado el marco de investigación para incorporar el papel que tiene la negociación entre gobierno, escuelas y sindicatos sobre salarios y condiciones laborales. Por ejemplo, Biasi (2021) encuentra que descentralizar la posibilidad de fijar condiciones laborales al nivel de las escuelas atrae a mejores docentes y mejora los resultados de los alumnos. Además, la evidencia presentada por Lovenheim y Willem (2019) muestra que la negociación

colectiva (esto es, la fijación de salarios por parte de los sindicatos y el gobierno) tiene un impacto en el largo plazo sobre los alumnos (especialmente minorías raciales y hombres) a través de un empeoramiento de sus habilidades no cognitivas y una reducción de su participación en el mercado de trabajo y sus ingresos a largo plazo (una reducción que estiman en 213 mil millones de dólares anuales en Estados Unidos). Pero otra parte de la investigación también identifica externalidades negativas, como una mayor rotación docente en las escuelas (Baron, 2018) y una mayor brecha de género entre profesores (Biasi y otros, 2022).

### 3.3. El continuo de la carrera profesional docente: desde la selección a la jubilación

Ya hemos aprendido cómo son los buenos docentes, qué impacto tienen sobre sus alumnos y cómo el contexto económico y político afecta su desempeño y su interés en la profesión. A continuación, vamos a la que es probablemente la pregunta más importante de este capítulo: ¿Qué políticas funcionan para que los docentes mejoren? ¿En qué momento y de qué manera debemos intervenir para atraer, seleccionar, potenciar, perfeccionar y evaluar a los mejores docentes? A continuación, repasamos todo lo que se denomina como el "continuo" de la carrera profesional docente, y que incluye la selección, formación inicial, inducción o iniciación a la docencia, desarrollo profesional continuo y jubilación.

***Selección.*** A la pregunta de cómo seleccionar a los docentes, la respuesta debe obviamente responderse teniendo en cuenta lo que hemos aprendido de "qué es un buen docente" en la sección anterior. Y eso es lo que precisamente ha mostrado toda la investigación al respecto de la selección docente, que habla de una selección con criterios variados que valoren tanto el expediente académico como por otras habilidades, algo que encuentran tanto Izadi (2019) para el caso de Finlandia (conocida por sus políticas docentes de excelencia) como Jacob y otros (2018) en el caso de Estados Unidos. Finalmente, es relevante el estudio de Schaede y Mankki (2023) para Finlandia, donde encuentran que al retirar un umbral mínimo de un 40% de docentes masculinos en Primaria tuvo un impacto negativo en el largo plazo de alumnado de ambos géneros, lo que sugiere la importancia de combinar talentos y habilidades (en este caso, la complementariedad de género en el aula) no necesariamente recogidos en procesos de selección, normalmente basados en indicadores cognitivos y no cognitivos.

***Experiencia inicial.*** ¿Hasta qué punto es relevante la experiencia en la mejora de la efectividad docente? ¿Cuántos años son necesarios para llegar a la efectividad óptima? ¿Y cómo se desarrolla esa experiencia y de qué depende para llegar a ella? Un resumen de lo que la investigación económica ha aportado en los últimos años nos invita a concluir lo siguiente: la experiencia es relevante, pero lo es especialmente en los primeros años y tiene un efecto marginal decreciente. En comparación con otros factores, parece ser un factor menos relevante que el punto de partida (la calidad inicial del docente) o que otros factores (Rockoff y otros, 2011; Kane y otros, 2008), aunque otros estudios muestran que los retornos a la docencia pueden seguir creciendo (Kraft y Papay, 2014). En todo caso, y quizás esto es lo más relevante, la experiencia no impacta de manera automática en todos los lugares por igual, sino que está asociada con la calidad del contexto organizativo y de desarrollo profesional de la escuela: el trabajo de Kraft y Papay (2014) es ilustrativo, pues muestran que en función del contexto de desarrollo profesional de la escuela y de cómo estas apoyan a los docentes (mediante medidas que recogen el clima disciplinario, la colaboración docente, el liderazgo del director, el desarrollo profesional, el respeto y confianza profesional entre docentes, y la existencia de evaluación docente) las diferencias del crecimiento de la efectividad docente entre un docente en el percentil 25 de efectividad inicial y un docente en el percentil 75 pueden ser de casi el 40%.

***Formación continua, aprendizaje entre iguales y coaching.*** Como ya hemos apuntado, una forma con importante potencial de mejorar el desempeño docente es fomentando la interacción entre docentes, algo que requiere espacios y tiempos adecuados para que esto ocurra, pero sin duda con capacidad para provocar mejoras importantes en las escuelas y las aulas. De la misma forma, los programas de acompañamiento o coaching docente, donde un docente experimentado acompaña durante varias sesiones a otro docente y le observa y da información sobre sus estrategias de aula tiene un potencial enorme, aunque los costes son importantes, pues los programas suelen ser muy individualizados. El meta-análisis de Kraft y otros (2018) encuentra un efecto de 0,18 desviaciones estándar sobre el aprendizaje, aunque con estudios de pequeña escala que no han sido replicados a gran escala.

***Expectativas y sesgos docentes***. La forma en que los docentes miran, diagnostican y observan a sus alumnos importa y tiene un impacto enorme, algo prometedor, pues se trata de cambios que tienen que ver más con creencias y sesgos que con programas costosos que requieren de recursos. Por ejemplo, el trabajo de Pa-

pageorge y otros (2020) muestra que la elasticidad de la finalización de estudios universitarios con respecto a las expectativas docentes es de 0,12. Además, la investigación más reciente ha estudiado la existencia de sesgos de género por renta u origen que afectan de manera desigual al alumnado: los estudios muestran la existencia de importantes sesgos de género en Italia y Francia (Carlana, 2019; Terrier, 2020). En España existen tres estudios que muestran resultados muy semejantes en Andalucía (Marcenaro-Gutiérrez et al., 2015), Cataluña (Calsamiglia y Loviglio, 2019) y País Vasco (Gortazar et al., 2022): los tres muestran cómo los docentes evalúan a sus alumnos respecto a las evaluaciones externas y evidencian los sesgos inducidos en detrimento de chicos, rentas bajas e inmigrantes.

***Incentivos y motivación docente.*** Una de las preguntas que más ha tratado de responder el campo de la microeconomía en las últimas décadas es la de cómo los agentes individuales responden a los incentivos. En educación, lógicamente, esto se ha trasladado con especial intensidad a comprender si funcionan y cómo los incentivos sobre los docentes. En este sentido, es importante volver a lo que hemos comentado en la introducción: la complejidad de la profesión docente no puede ignorarse cuando hablamos de incentivos (monetarios y no monetarios) y su impacto en los profesores. Si los profesores tienen la capacidad y las habilidades para reaccionar a los incentivos de manera positiva, tiene sentido que los incentivos pueden funcionar; pero en muchas ocasiones, no se trata solo poner los incentivos adecuados, sino que estos tengan en cuenta las capacidades de los profesores para reaccionar a ellos: en ocasiones la complejidad de la profesión hace que los profesores no reaccionen a los incentivos, no porque no quieren, sino porque no saben. Es por tanto central comprender la pregunta. ¿Qué incentivos funcionan y qué papel juega la motivación docente? A continuación, repasamos los resultados principales de la cuestión:

1) *Informar a los docentes* sobre su propio desempeño puede suponer un incentivo en sí mismo activando la motivación. Infomar sobre el desempeño puede suponer un incentivo a través de la motivación intrínseca (Taylor y Tyler, 2012).

2) *Los incentivos económicos* por desempeño (pagar a los profesores por sus resultados) funcionan en algunos contextos, afectando positivamente tasas de graduación y resultados en el mercado laboral a largo plazo a través de mejoras en la enseñanza, más horas de refuerzo y adaptación a las necesidades de los estudiantes (ver Lavy, 2009; y Lavy, 2020 para el caso de

Israel). Pero otros programas han sido poco o nada efectivos al respecto: por ejemplo, las reformas escolares en Estados Unidos que han promovido una mayor rendición de cuentas y evaluación docente mediante incentivos económicos al profesorado en función de su desempeño en Estados Unidos muestran evidencia mixta (Dee y Wyckoff, 2015; Dee, James y Wickoff, 2021) siendo políticamente muy costosas (por su impopularidad) y a la vez reduciendo la satisfacción docente y la oferta en el mercado laboral docente (Kraft, 2020). En países en desarrollo la evaluación por resultados puede funcionar en contextos de bajo desempeño (Ganimian y Murnane, 2014; Muralidharan et al., 2011) especialmente para prevenir el absentismo docente.

***Asignación de docentes a escuelas y alumnos ("Matching policies").*** Una cuestión central a la hora de pensar en los docentes y sus efectos sobre los estudiantes es de qué manera debemos emparejar y asignar profesores a escuelas y alumnos. La investigación ha mostrado que las características sociodemográficas del profesorado influyen e impactan en el alumnado de manera diferente en función de las características de estos últimos. Por ejemplo, un resultado habitual en la literatura es que el género influye en los resultados: Lim y otros (2020) muestran que para el caso de Corea del Sur, los docentes de género femenino tienen un impacto sobre las niñas positivo, tanto a corto plazo (resultados académicos) como largo plazo (aumentando su probabilidad de escoger asignaturas científicas en la Educación Secundaria). A su vez, como hemos visto, el estudio de Schaede y Mansski (2023) para Finlandia habla de la importancia del equilibrio de género entre el profesorado, ya que a partir de un 60% de mujeres docentes los resultados pueden resentirse. Finalmente, existe evidencia de mejoras en los resultados cuando docente y alumno comparten raza, algo que se produce sobre todo con alumnado de nivel socioeconómico bajo (Dee, 2005). En la misma dirección apuntan las políticas que promueven llevar docentes experimentados en centros que atienden a alumnos de renta baja (Steele et al., 2010; Swain et al., 2019; Bobba y otros, 2021) en zonas remotas o en asignaturas con poca oferta docente (Bueno et al., 2018; Feng et al., 2018). Estos programas se articulan mediante incentivos económicos *ex-ante* por trabajar en este tipo de escuelas.

## 4. LOS EQUIPOS DIRECTIVOS Y SU IMPACTO EN LA ESCUELA

En esta última sección nos ocupamos del trabajo que algunos profesores acaban realizando a lo largo de sus carreras profesionales: dirigir escuelas. ¿Qué

dice la investigación sobre gestión de escuelas desde una perspectiva económica? ¿Tienen impacto y marcan la diferencia los directores? ¿De qué forma lo hacen? Leyendo este capítulo y comprendiendo que la escuela y su contexto explican una parte de la efectividad docente, entonces quienes son los responsables últimos de ello (los equipos directivos) deberían tener algo que ver en todo lo que hemos contado a lo largo de estas páginas.

¿Qué evidencia disponemos y cómo medimos esa gestión? ¿En qué consiste una buena gestión? ¿Se refiere a la gestión económica y financiera de la institución o nos referimos también al impacto educativo que los directores tienen? Un estudio importante, publicado por Bloom y otros (2015) adaptó la metodología de la economía de la empresa sobre gestión organizacional para aplicarla a educación: los autores llevaron a cabo 1.800 entrevistas telefónicas a directores de escuelas de 8 países para construir un indicador de gestión escolar basado en 20 dimensiones de la gestión de los colegios centradas en operaciones, seguimiento, fijación de objetivos y gestión de personas. Muestran que esos indicadores están asociados a mejores resultados académicos.

Algunos canales bajo los cuales operan buenos directores es el de la estabilidad: dos estudios recientes muestran que el impacto de buenos directores (medidos a través de diversos sistemas de evaluación) es alto sobre los alumnos y además está asociado a la estabilidad en las escuelas (Grissom et al., 2019; Miller, 2013). Sin embargo, aun siendo una cuestión muy importante desde el punto de vista de la investigación y las políticas públicas, es un campo en proceso de exploración con críticas metodológicas recientes (Bartanen, 2022).

Por otro lado, hay dos dimensiones más acerca de los directores de escuela que merecen especial atención: la manera en que son seleccionados y la forma en que podamos mejorar su desempeño. Respecto a lo primero, el trabajo de Akhtari y otros (2022) muestra que un modelo de gestión directiva de escuelas demasiado dependiente del poder político tiene un impacto negativo sobre aprendizaje: a partir de un modelo de regresión en discontinuidad con los resultados de elecciones locales en Brasil, encuentran que en los municipios con cambio de partido en la alcaldía aumenta el reemplazo de directores y profesores en escuelas que dependen del poder local, y eso, a su vez, provoca una caída de resultados de aprendizaje de entre un 5% y un 8% de una desviación estándar. Por otro lado, el trabajo de De Hoyos y otros (2021) en Argentina muestra una vía de investigación importante, que enlaza con lo que ya hemos visto acerca de

cómo la información y feedback sobre su propio desempeño mejora la efectividad docente: en este caso, mediante un experimento aleatorio controlado, los autores asignan a directores de escuelas de forma aleatoria a tres grupos: uno donde estos recibieron información sobre el desempeño de sus escuelas, otro donde los directores recibieron formación y visitaron las escuelas y un grupo de control. Los autores muestran que las escuelas donde los directores recibieron feedback e información sobre su desempeño tuvieron resultados mejores un año y dos años después, en una magnitud relevante: los directores de estas escuelas eran más proclives a usar datos para tomar decisiones y sus alumnos reportaron que sus docentes usaban estrategias de aprendizaje eficaces.

## 5. CONCLUSIONES

Este capítulo ha tratado de contribuir a una de las preguntas más importantes que debemos responder los economistas interesados por la educación: ¿Cuánto importan los docentes en la educación? En particular, hemos repasado la relevancia que tienen los profesores desde el punto de vista económico, es decir, tanto del lado del presupuesto público sobre su impacto en la economía a través de la creación y construcción del capital humano. Si un sistema escolar y una escuela es capaz de lograr traducir escolarización en aprendizaje y capital humano, esto se debe, sobre todo, aunque no solo ni tampoco siempre de manera aislada, al papel que juegan los profesores.

Este capítulo se beneficia de la revolución de la causalidad que ha ocurrido en los últimos 20 años en la economía, pues buena parte de los resultados presentados y las preguntas respondidas se deben a ello. También a que se ha multiplicado la disponibilidad de datos administrativos, así como la colaboración con otras disciplinas como la sociología o la psicología. Llama la atención cómo la investigación económica se está volviendo enormemente sofisticada y multiplicando sus preguntas y modelos para comprender el universo de los profesores, como por ejemplo la importancia del factor escuela, la incorporación de factores no-cognitivos en el aprendizaje o el estudio de las políticas clave del profesorado y los equipos directivos.

La ampliación de datos y encuestas disponibles, la aplicación de nuevas metodologías econométricas, así como el incremento de la relevancia de las preguntas a cuestiones económicas de primer calado (como por ejemplo el papel que juegan los sindicatos, el funcionamiento del mercado laboral de los docentes o el funcionamiento de los incentivos para motivar de manera efectiva) muestran

que estamos ante una cuestión central en la economía de la educación y la economía laboral, que con toda seguridad contribuirá a la mejora del conocimiento en las próximas décadas y a las decisiones que tomen los gobiernos. A nivel global, cuestiones como la falta de atracción de la profesión y la escasez de profesores, las preferencias de los docentes en el mercado de trabajo y sus condiciones laborales y los desequilibrios de género parecen cobrar importancia y lo seguirán haciendo en los próximos años.

Todo ello debe llegar y extenderse al caso español donde, como hemos visto a lo largo del capítulo, los retos son importantes, pero la investigación es todavía muy escasa, por falta de datos, apoyo económico a proyectos complejos, y la todavía escasa comunidad investigadora de economía de la educación. La investigación sobre profesores requiere de tres ingredientes clave: primero, hacer accesible y tratable datos administrativos relacionados con los profesores; segundo, priorizar esta cuestión en las agencias de financiación públicas y privadas, pues hablamos clave para el conocimiento acumulado y el modelo de crecimiento económico para España a largo plazo; tercero, colaborar de manera más activa con las facultades de educación, encargadas de la formación del profesorado; cuarto, promover un mayor interés por la cuestión en las facultades de economía, algo que este libro y este capítulo pretenden lograr.

El sistema educativo español no está exento de retos, siendo la profesión docente un espacio donde se han producido pocos cambios en las últimas décadas. Algunas preguntas importantes deberían ser respondidas por nuestra disciplina e incluyen, entre otras, la falta de rendición de cuentas del profesorado, la medición de valor añadido, la colaboración y aprendizaje entre pares, la existencia sesgos o el papel de las expectativas sobre los alumnos. Otra vía de trabajo debería incorporar el efecto del mercado laboral docente post-crisis y los sistemas de asignación docente, teniendo en cuenta la alta interinidad y rotación y su impacto en aprendizaje, el papel de otros sistemas de asignación de docentes a escuelas, y el progresivo envejecimiento de la profesión. Finalmente, desde el punto de vista de la selección docente, es necesario comprender, mediante datos de panel, la capacidad predictiva de los sistemas de selección (y sus posibles alternativas) en Grados, Másteres y oposiciones docentes en el desempeño laboral futuro. También es necesario comprender mejor la efectividad de programas de formación continua docente, así como el papel que pueden y están jugando los incentivos a docentes por trabajar en zonas rurales o centros con alumnado de bajo nivel socioeconómico, ambos muy relevantes para el caso español.

## BIBLIOGRAFÍA


Adnot, M., Dee, T., Katz, V. y Wyckoff, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis*, *39*(1), 54-76.

Akhtari, M., Moreira, D. y Trucco, L. (2022). Political turnover, bureaucratic turnover, and the quality of public services. *American Economic Review*, *112*(2), 442-493.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y. y Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, *131*(3), 1415-1453.https://academic.oup.com/qje/article-abstract/131/3/1415/2461200

Bacher-Hicks, A., Chin, M. J., Kane, T. J. y Staiger, D. O. (2019). An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys. *Economics of Education Review*, *73*, 101919.

Bacolod, M. P. (2007). Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality. *The Review of Economics and Statistics*, *89*(4), 737-751.

Barber, M. y Mourshed, M. (2007). *How the world's best-performing school systems come out on top*. McKinsey & Company.

Bardach, L. y Klassen, R. M. (2020). Smart teachers, successful students? A systematic review of the literature on teachers' cognitive abilities and teacher effectiveness. *Educational Research Review*, *30*, 100312.

Baron, E. J. (2018). The effect of teachers' unions on student achievement in the short run: Evidence from Wisconsin's Act 10. *Economics of Education Review*, *67,* 40-57.

Bartanen, B., Husain, A. N. y Liebowitz, D. D. (2022). *Rethinking principal effects on student outcomes*. Ed Working Paper: 22-621. Retrieved from Annenberg Institute at Brown University, Providence, RI, USA. doi: 10.26300/r5sf-3918.

Bau, N. y Das, J. (2020). Teacher value added in a low-income country. *American Economic Journal: Economic Policy*, *12*(1), 62-96.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, *70*(5, Part 2), 9-49.

Becker, G. S. y Chiswick, B. R. (1966). Education and the Distribution of Earnings. *The American Economic Review*, *56*(1/2), 358-369.

Biasi, B. y Sarsons, H. (2022). Flexible wages, bargaining, and the gender gap. *The Quarterly Journal of Economics*, *137*(1), 215-266.

Biasi, B. (2021). The labor market for teachers under different pay schemes. *American Economic Journal: Economic Policy*, *13*(3), 63-102.

Biasi, B. y Sarsons, H. (2022). Flexible wages, bargaining, and the gender gap. *The Quarterly Journal of Economics*, *137*(1), 215-266.

Bloom, N., Lemos, R., Sadun, R. y Van Reenen, J. (2015). Does management matter in schools? *The Economic Journal*, *125*(584), 647-674.

Bobba, M., Ederer, T., Leon-Ciliotta, G., Neilson, C. y Nieddu, M. G. (2021). *Teacher compensation and structural inequality: Evidence from centralized teacher school choice in Perú* (No. w29068). National Bureau of Economic Research.

Bueno, C. y Sass, T. R. (2018). The effects of differential pay on teacher recruitment and retention. *Andrew Young School of Policy Studies Research Paper Series* (18-07).

Burgess, S., Rawal, S. y Taylor, E. S. (2021). Teacher peer observation and student test scores: Evidence from a field experiment in English secondary schools. *Journal of Labor Economics*, *39*(4), 1155-1186.

Calsamiglia, C. y Loviglio, A. (2019). Grading on a curve: When having good peers is not good. *Economics of Education Review*, *73*, 101916.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, *134*(3), 1163-1224.

Cheng, A. y Zamarro, G. (2018). Measuring teacher non-cognitive skills and its impact on students: Insight from the Measures of Effective Teaching Longitudinal Database. *Economics of Education review*, *64*, 251-260.

Chetty, R., Friedman, J. N. y Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review*, *104*(9), 2633-2679.

Chetty, R., Friedman, J. N. y Rockoff, J. E. (2017a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593.

Coffman, L. C., Conlon, J. J., Featherstone, C. R. y Kessler, J. B. (2019). Liquidity affects job choice: Evidence from Teach for America. *The Quarterly Journal of Economics*, *134*(4), 2203-2236.

Corbett, H. D. (1982). To Make an Omelette You Have to Break the Egg Crate. *Educational Leadership*, *40*(2), 34-35.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, *95*(2), 158-165.

Dee, T. S. y Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267-297.

Dee, T. S., James, J. y Wyckoff, J. (2021). Is effective teacher evaluation sustainable? Evidence from District of Columbia Public Schools. *Education Finance and Policy*, *16*(2), 313-346.

Escardíbul, J. O. (2015). ¿Es recomendable implantar incentivos salariales para el profesorado vinculados al rendimiento académico de los estudiantes? Alegre, M.À. y Escardíbul, J.O. *¿Qué funciona en educación*?

Feng, L. y Sass, T. R. (2018). The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects. *Journal of Policy Analysis and Management*, *37*(1), 112-135.

Fernández Enguita, M. (2018). *Más escuela y menos aula: la innovación en la perspectiva de un cambio de época.*

Ganimian, A. J. y Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, *86*(3), 719-755.

Gjefsen, H. M. (2020). Wages, teacher recruitment, and student achievement. *Labour Economics*, *65*, 101848.

Gortazar, L., de Lafuente, D. M. y Vega-Bayo, A. (2022). Comparing teacher and external assessments: Are boys, immigrants, and poorer students undergraded? *Teaching and Teacher Education*, *115*, 103725.

Grissom, J. A. y Bartanen, B. (2019). Principal effectiveness and principal turnover. *Education Finance and Policy*, *14*(3), 355-382.

Grönqvist, E. y Vlachos, J. (2016). One size fits all? The effects of teachers' cognitive and social abilities on student achievement. *Labour Economics*, *42*, 138-150.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education review*, *30*(3), 466-479.

Hanushek, E.A., Piopiunik, M. y Wiederhold, S. (2019). The value of smarter teachers international evidence on teacher cognitive skills and student performance. *Journal of Human Resources*, 2019, vol. 54, no 4, p. 857-899.

Izadi, R. (2019). Teacher Selection in Finland. Does academic ability predict teacher aptitude? IZA *Working Paper*.

Jacob, B. A., Rockoff, J. E., Taylor, E. S., Lindy, B. y Rosen, R. (2018). Teacher applicant hiring and teacher performance: Evidence from DC public schools. *Journal of Public Economics*, *166*, 81-97.

Jackson, P. W. (1968). *Life in classrooms*. Teachers College Press.

Jackson, C. K. y Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, *1*(4), 85-108.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non-test score outcomes. *Journal of Political Economy*, *126*(5), 2072-2107.

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.

Kraft, M. A. y Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational evaluation and policy analysis*, *36*(4), 476-500.

Kraft, M. A., Blazar, D. y Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of educational research*, *88*(4), 547-588.

Kraft, M. A., Brunner, E. J., Dougherty, S. M. y Schwegman, D. J. (2020). Teacher accountability reforms and the supply and quality of new teachers. *Journal of Public Economics*, *188*, 104212.

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, *99*(5), 1979-2011.

Lavy, V. (2020). Teachers' pay for performance in the long-run: The dynamic pattern of treatment effects on students' educational and labour market outcomes in adulthood. *The Review of Economic Studies*, *87*(5), 2322-2355.

Leigh, A. (2012). Teacher pay and teacher aptitude. *Economics of Education Review*, *31*(3), 41-53.

Lim, J. y Meer, J. (2020). Persistent effects of teacher-student gender matches. *Journal of Human Resources*, *55*(3), 809-835.

Lovenheim, M. F. y Willén, A. (2019). The long-run effects of teacher collective bargaining. *American Economic Journal: Economic Policy*, *11*(3), 292-324.

Loviglio, A. (2019). *Schools and their multiple ways to impact students: a structural model of skill accumulation and educational choices*. Mimeo.

Marcenaro-Gutierrez, O. y Vignoles, A. (2015). A comparison of teacher and test-based assessment for Spanish primary and secondary students. *Educational Research*, *57*(1), 1-21.

Marchand, J. y Weber, J. G. (2020). How local economic conditions affect school finances, teacher quality, and student achievement: evidence from the Texas shale boom. *Journal of Policy Analysis and Management*, *39*(1), 36-63.

Miller, A. (2013). Principal turnover and student achievement. *Economics of Education Review*, *36*, 60-72.

Muralidharan, K. y Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of political Economy*, *119*(1), 39-77.

Nagler, M., Piopiunik, M. y West, M. R. (2020). Weak markets, strong teachers: Recession at career start and teacher effectiveness. *Journal of Labor Economics*, *38*(2), 453-500.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. *The Elementary School Journal*, *84*(2), 113-130.

OECD (2022). Education at a Glance 2022: OECD Indicators*, OECD Publishing,* Paris, https://doi.org/10.1787/3197152b-en

Papageorge, N. W., Gershenson, S. y Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics*, *102*(2), 234-251.

Papay, J. P., Taylor, E. S., Tyler, J. H. y Laski, M. E. (2020). Learning job skills from colleagues at work: Evidence from a field experiment using teacher performance data. *American Economic Journal: Economic Policy*, *12*(1), 359-388.

Petek, N. y Pope, N. G. (2023). The multidimensional impact of teachers on students. *Journal of Political Economy*, *131*(4), 1057-1107.

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

Rockoff, J. E., Jacob, B. A., Kane, T. J. y Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education finance and Policy*, *6*(1), 43-74.

Ronfeldt, M., Loeb, S. y Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4-36.

Rose, E. K., Schellenberg, J. T. y Shem-Tov, Y. (2022). The Effects of Teacher Quality on Adult Criminal Justice Contact (No. w30274). *National Bureau of Economic Research*.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers, 3*(2), 135-146.

Schaede, U. y Mankki, V. (2022). Quota vs quality? Long-term gains from an unusual gender quota. *Working Paper*.

Steele, J. L., Murnane, R. J. y Willett, J. B. (2010). Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California. *Journal of Policy Analysis and Management*, *29*(3), 451-478.

Sun, M., Loeb, S. y Grissom, J. A. (2017). Building teacher teams: Evidence of positive spillovers from more effective colleagues. *Educational Evaluation and Policy Analysis*, *39*(1), 104-125.

Swain, W. A., Rodriguez, L. A. y Springer, M. G. (2019). Selective retention bonuses for highly effective teachers in high poverty schools: Evidence from Tennessee. *Economics of Education Review*, *68*, 148-160.

Taylor, E. S. y Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, *102*(7), 3628-3651.

Terrier, C., 2020. Boys lag behind: How teachers' gender biases affect student achievement. *Economics of Education Review, 77*, 101981.

Tyler, J. H., Taylor, E. S., Kane, T. J. y Wooten, A. L. (2010). Using student performance data to identify effective classroom practices. *American Economic Review, 100*(2), 256-260.

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, *100*, 61.

# Capítulo 3e
# LOS CONOCIMIENTOS DE IDIOMAS EXTRANJEROS

Brindusa Anghel
Departamento de Análisis Estructural y Estudios Microeconómicos
Banco de España
ORCID: https://orcid.org/0000-0002-3504-0235

**Resumen**

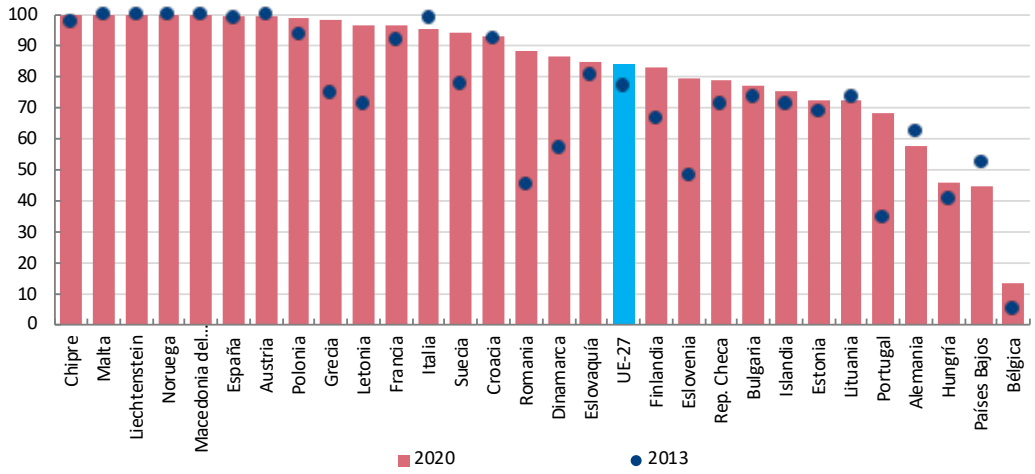
Los idiomas extranjeros son actualmente una parte fundamental de la educación de una persona. Los gobiernos invierten muchos recursos en el aprendizaje de los idiomas extranjeros, cada vez desde edades más tempranas. Este capítulo documenta en primer lugar la importancia y los beneficios económicos de los conocimientos de idiomas extranjeros. Segundo, realiza un repaso de algunos programas escolares en España y en otros países que implementan el aprendizaje de un segundo idioma a través de su uso como idioma de la enseñanza. Tercero, presenta la evidencia empírica sobre la evaluación de estos programas, explicando los problemas metodológicos que esta supone y las soluciones que se han propuesto en la literatura.

## 1. INTRODUCCIÓN

Aprender y conocer un idioma extranjero es una parte fundamental de la educación de un individuo en la actual sociedad y es esencial para que los individuos puedan desarrollar sus actividades en un mundo cada vez más globalizado e interconectado. Además, es cada vez más frecuente que el aprendizaje de los idiomas extranjeros se produzca desde edades muy tempranas. Según Foreign Language Statistics (Eurostat), en el año 2021, más de un 80% de los alumnos de Educación Primaria estaban estudiando el inglés como un idioma extranjero y en muchos países de la Unión Europea (UE) aprender inglés es obligatorio en el nivel de Educación Primaria (véase gráfico 1). Casi la práctica totalidad de los alumnos de Educación Primaria (un 99%-100%) en países como Austria, Chipre, España o Malta estaban estudiando inglés. En cuanto a los alumnos de Edu-

cación Secundaria en la UE, el 98,3% de los alumnos de la Secundaria inferior y el 88% de los alumnos de la Secundaria superior estaban estudiando inglés[1] (véase gráfico 2). Aparte del inglés, los idiomas más estudiados en los países de la UE son el francés (un 30,5% y un 19% de los alumnos de Educación Secundaria inferior y de la superior, respectivamente), el alemán (aproximadamente un 20% de los alumnos de Educación Secundaria inferior y de la Secundaria superior) y el español (aproximadamente un 18% de los alumnos de Educación Secundaria inferior y de la superior).

**Gráfico 1. Proporción de alumnos en Educación Primaria que estudian inglés.**

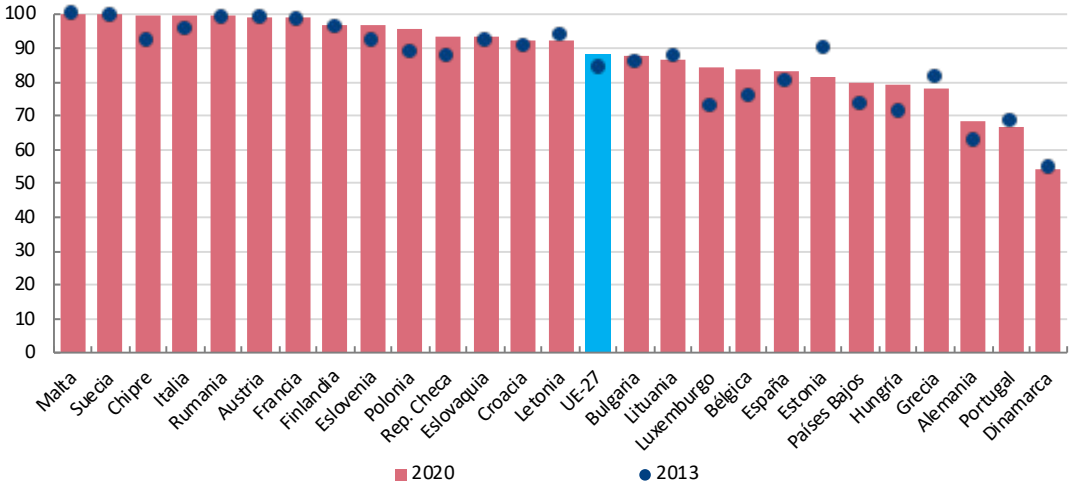



Los datos de la prueba internacional Programme for International Student Assessment (PISA) del año 2018[2] muestran que las expectativas de los alumnos de Secundaria con respecto a graduarse de una Universidad o desarrollar una carrera como manager o especialista en algún campo son más altas entre aquellos estudiantes que hablan más de un idioma que entre los estudiantes monolingües. También son más altas entre aquellos estudiantes que están aprendiendo un idioma extranjero en la escuela relativo a los que no lo hacen. Estas asociacio-

[1] Véase Foreign Language Statistics (Eurostat): https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_learning_statistics&oldid=379813

[2] El último año disponible en la fecha de publicación de este libro.

**Gráfico 2. Proporción de alumnos en Educación Secundaria Superior que estudian inglés.**




nes se observan incluso cuando se tiene en cuenta el desempeño académico, el género y el estatus de inmigrantes de los estudiantes, así como el nivel socioeconómico de los alumnos y de las escuelas. Este resultado se podría interpretar de dos formas: los conocimientos de otro idioma podrían llevar a los estudiantes a tener expectativas más altas en cuanto a su desarrollo educativo o profesional o los estudiantes con expectativas altas simplemente eligen estudiar más idiomas extranjeros. La información cualitativa que se recoge a través de entrevistas personales en algunos países de la OCDE que participaron en PISA[3] indican que se podría tratar de ambos mecanismos.

El aprendizaje de uno o más idiomas puede ocurrir en contextos muy diferentes (Costa y Sebastián-Gallés, 2014). Los individuos pueden aprender al mismo tiempo dos idiomas desde el nacimiento (bilingües simultáneos). Alternativamente, los individuos pueden aprender un segundo idioma más tarde en la vida dentro de un sistema educativo formal, por ejemplo, como un idioma extranjero en la escuela o en un ambiente de inmersión lingüística donde el idioma extranjero se utiliza como el idioma de la enseñanza (los

[3] Dentro del marco de la prueba PISA 2018, la OCDE organizó entrevistas cualitativas personales con estudiantes de algunos de los países participantes, uno de ellos siendo España (para más información véase Covacevich y Vargas, 2021).

programas de educación bilingües). Asimismo, un contexto social importante que podría generar el aprendizaje de un idioma distinto al idioma materno es la inmigración (bilingües sucesivos). Este capítulo se centrará sobre todo en el aprendizaje de idiomas dentro de programas educativos formales (como asignaturas en el currículo escolar o como programas de educación bilingüe), pero también se analizará brevemente el caso de los individuos bilingües como resultado de la inmigración.

En primer lugar, en el capítulo se presentan los beneficios económicos derivados de los conocimientos de idiomas extranjeros para los individuos, las empresas y los países y se proporciona una revisión de los artículos académicos más destacados que analizan esta cuestión en distintos países. Los conocimientos de idiomas extranjeros teóricamente mejorarían el capital humano de un individuo, por tanto, podrían tener un impacto positivo sobre su productividad, y en consecuencia, sobre sus resultados en el mercado laboral.

En segundo lugar, el capítulo realiza una descripción de varios programas que implementan el aprendizaje de un idioma extranjero a través de su uso como idioma de la enseñanza (denominados a veces programas de educación bilingüe) en España y en otros países de Europa, así como en los EE. UU.

En tercer lugar, el capítulo hace especial hincapié en la importancia de la investigación sobre la adquisición de un segundo idioma y la evaluación de estos programas y políticas, dado que tanto la enseñanza como el aprendizaje de otro idioma por parte de los individuos tienen no solamente beneficios, pero también costes que hay que tener en cuenta a la hora de implementar dichas políticas. En cuanto a la enseñanza, los más evidentes son los costes financieros: por ejemplo, las escuelas tienen que contratar profesores especializados para esos idiomas, tienen que ofrecerles formación en el idioma extranjero o tienen que contratar profesores ayudantes nativos para clases de conversación. En cuanto al aprendizaje, por parte de los niños requiere invertir tiempo en la escuela en detrimento de la instrucción en otras materias en el idioma nativo, y por parte de los adultos requiere invertir recursos económicos, así como tiempo en detrimento de otras actividades laborales o de formación. Por tanto, una parte importante del capítulo se dedicará a estudiar detenidamente algunos trabajos académicos que han evaluado programas y políticas que promueven el aprendizaje de un idioma extranjero.

Tal y como se verá a lo largo del capítulo, los resultados de los estudios existentes difieren dependiendo de muchos factores como el país, el idioma, la legislación, el sistema educativo, etc. Según Hulstijn (1997), una de las mayores limitaciones en la investigación sobre la adquisición de un segundo idioma es que hay muchas variables que afectan la comparabilidad de los grupos involucrados en distintas experiencias de aprendizaje y que pueden confundir cuando se realiza el análisis de estas experiencias. Según este autor, "Uno de los mayores retos metodológicos es mantener estas variables constantes. Esto es casi imposible en clases "normales" con alumnos reales que aprenden un segundo idioma. Por tanto, no es ninguna sorpresa, que los resultados de los estudios llevados a cabo en entornos de aprendizaje naturales, muy a menudo son el objeto de discrepancias considerables".

La estructura del capítulo es la siguiente: en la segunda sección se analizan los beneficios de los conocimientos de idiomas extranjeros, en la tercera sección se describen varios programas que implementan el uso de un segundo idioma como idioma de la enseñanza en los colegios en España y en otros países, en la cuarta sección se analizan algunos artículos académicos que realizan evaluaciones de estos programas y políticas y en la quinta se analizan otros ejemplos de programas o reformas que han implementado el uso de un idioma extranjero como idioma de la enseñanza en las escuelas. La última sección contiene las conclusiones.

## 2. LOS BENEFICIOS ECONÓMICOS DE CONOCER OTROS IDIOMAS

Desde la perspectiva de un trabajador, las habilidades derivadas de conocer un idioma extranjero son parte de su capital humano, por tanto, podrían mejorar su situación en el mercado laboral, en cuanto a encontrar mejores oportunidades laborales, tener más movilidad geográfica, tener más oportunidades de ser promocionados, o tener unos salarios más altos. Desde la perspectiva de las empresas y los países, los conocimientos de idiomas extranjeros de sus trabajadores serían un elemento esencial para abrir mercados, comercializar productos o invertir en otros países.

### 2.1. Beneficios económicos a nivel individual

Un estudio de la Comisión Europea realizado por Araújo et al. (2015) encuentra que, en 25 estados miembros de la UE, los individuos que conocen uno

o más idiomas extranjeros tienen más probabilidades de estar empleados que los individuos monolingües, independientemente de su nivel de competencias. Varios trabajos que utilizan datos sobre los inmigrantes en Alemania o en el Reino Unido (Aldashev, Gernandt y Thomsen, 2009; Dustmann y Fabbri, 2003) encuentran que las oportunidades de empleo de los inmigrantes aumentan con las competencias en el idioma del país. En concreto, en el caso de Alemania, un alto uso del idioma alemán reportado por los inmigrantes aumenta las probabilidades de estar empleado, mientras que, en el Reino Unido, las competencias en inglés de los inmigrantes están asociadas con un aumento de un 22% en la probabilidad de estar empleados.

En cuanto a los salarios, la evidencia empírica existente muestra resultados diferentes en los distintos países. En los Estados Unidos (EE. UU.), las estimaciones indican que los rendimientos salariales de las habilidades en otros idiomas que no sean el inglés son relativamente modestos, de aproximadamente un 2%-3% (Saiz y Zoido, 2005). Fry y Lowell (2003) argumentan que los salarios más altos de los trabajadores que conocen otro idioma que no sea el inglés reflejan en realidad su nivel educativo más alto y una vez que controlan por características socioeconómicas observables, no parece que haya un rendimiento salarial de los idiomas extranjeros estadísticamente significativo.

En Europa, los rendimientos salariales de los idiomas extranjeros son considerablemente más altos que en los EE. UU., el inglés siendo un idioma que tiene un valor económico indiscutible en el mercado laboral. Por ejemplo, Ginsburgh y Prieto-Rodríguez (2011) muestran que en nueve países europeos las competencias en un idioma extranjero que se usa en el trabajo tienen un impacto positivo sobre los salarios de los trabajadores. Así, el conocimiento del inglés aumenta los salarios de forma muy heterogénea en los distintos países analizados: en Austria en un 11%, en Italia en un 18%, en Finlandia en un 23%, en Francia en un 29% y en España en un 39%. En algunos países, el impacto sobre los salarios es aún mayor cuando se trata de otros idiomas que no sean el inglés: conocer el francés aumenta los salarios en España en un 49% y en Portugal en un 34%, conocer el alemán en Francia aumenta los salarios en un 46% y conocer el español en Portugal en un 45%. Asimismo, en la mitad de los países analizados los efectos son más grandes en las decilas superiores de la distribución de los salarios, puesto que en esa parte se encuentran las ocupaciones más cualificadas donde es más probable que se usen los idiomas extranjeros.

Williams (2011) estima el retorno laboral del uso de los idiomas extranjeros en el trabajo para muestras de trabajadores en catorce países de Europa occidental y obtiene resultados similares a los de Ginsburgh y Prieto-Rodríguez (2011): unos retornos entre 5% y 20%, dependiendo del país. El impacto más relevante es para el idioma inglés, mientras que para el francés, italiano o alemán también se obtienen rendimientos salariales en algunos países. Resultados similares se obtienen para una muestra de inmigrantes de España en el trabajo de Isophording (2013): unos rendimientos salariales muy altos para el inglés, pero también para el francés.

Hahm y Gazzola (2022) estiman la relación entre las habilidades en inglés como idioma extranjero de los individuos y sus rentas laborales en Alemania. Encuentran que, para nativos alemanes, los conocimientos del inglés pueden llegar a tener un rendimiento salarial de hasta un 13%. Además, los autores hacen un análisis del impacto sobre los salarios de distintos niveles de competencias en inglés utilizando el método de estimación con variables instrumentales y encuentran que cada nivel adicional de mejora de competencias en inglés incrementa los salarios en un 11%, en promedio.

### 2.2. Beneficios económicos a nivel de empresa y de país

Los beneficios económicos de los idiomas extranjeros van más allá del impacto sobre los salarios o sobre la situación en el mercado laboral de un individuo.

A nivel de empresa, los conocimientos de idiomas extranjeros de sus empleados podrían estar asociados con un aumento o una reducción en los ingresos de exportaciones. La Comisión Europea realizó en 2006 un estudio llamado ELAN[4] ("The Effects on the European Economy of Shortages of Foreign Language Skills in Enterprise", Comisión Europea, 2006) que recogió información sobre las competencias en idiomas extranjeros de los trabajadores de 2.000 pequeñas y medianas empresas. Los resultados indican que la falta de conocimientos de idiomas extranjeros en empresas de este tamaño podría estar asociada con una pérdida de oportunidades de negocios. El estudio también muestra que una proporción muy grande de empresas reclutadoras eligen candidatos nativos en un idioma extranjero o con conocimientos de idiomas extranjeros; que muchas

[4] https://ec.europa.eu/assets/eac/languages/policy/strategic-framework/documents/elan_en.pdf

empresas ofrecen cursos de formación en otros idiomas a sus empleados; que además del inglés, se usa también el francés, alemán o ruso en el ambiente de negocios de Europa, y que los candidatos con experiencias de trabajo fuera de su país de origen se consideran más competitivos debido a sus habilidades y conocimientos de idiomas y a su más amplia cultura de negocios. Otra conclusión del estudio es que las pequeñas y medianas empresas con capacidad exportadora que no disponen de trabajadores con conocimientos de idiomas extranjeros podrían perder una cantidad relevante del negocio hasta el punto de que algunas empresas podrían haber perdido contratos de más de 8 millones de euros.

A nivel de país, los idiomas son un importante determinante de su comercio bilateral. Hutchinson (2002) analiza el papel del inglés en las relaciones de comercio de 33 países con los EE. UU. y concluye que la proporción de la población que habla inglés en esos países es un factor significativo tanto para las exportaciones como para las importaciones entre los EE. UU. y esos 33 países. Melitz (2008) y Melitz y Toubal (2014) analizan, aparte de los idiomas oficiales, el papel de los idiomas indígenas o establecidos que se hablan por un 4% de la población y encuentra que ambas categorías de idiomas aumentan el comercio bilateral. Firdmuc y Fidrmuc (2016) demuestran a través de un modelo de gravedad del comercio y utilizando datos de 29 países europeos, que los conocimientos de idiomas extranjeros de la población de un país son un importante determinante de su comercio exterior, el inglés jugando un papel especialmente importante.

## 3. PROGRAMAS QUE IMPLEMENTAN EL USO DE UN IDIOMA EXTRANJERO COMO IDIOMA DE LA ENSEÑANZA (PROGRAMAS BILINGÜES) EN ESPAÑA Y EN OTROS PAÍSES

La enseñanza de un idioma distinto del idioma nativo es actualmente una parte fundamental de los sistemas educativos nacionales. Programas de educación bilingüe en los colegios (en el idioma del país y en un segundo idioma, generalmente el inglés), la introducción de uno o más idiomas extranjeros en los planes de estudio de los colegios desde edades tempranas, programas de intercambio de estudiantes en el extranjero, escuelas privadas de enseñanza de idiomas extranjeros, etc., son algunos ejemplos de herramientas educativas que pueden proporcionar ayuda para el aprendizaje de otros idiomas. A continuación, se describen las características y el diseño de algunos programas que implementan el uso de un idioma extranjero como idioma de la enseñanza en España y tam-

bién en otros países. Comprender estas características es muy importante para que luego se puedan evaluar estos programas.

### 3.1. La educación bilingüe en España

En Europa apoyar la diversidad lingüística y el aprendizaje de los idiomas ha sido una línea política constante. La educación bilingüe se empezó a ofrecer en los colegios públicos a través de programas que generalmente se denominan Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (CLIL, según sus iniciales en inglés –Content and Language Integrated Learning–). El objetivo es el aprendizaje de un idioma extranjero de forma simultánea con el idioma nativo y dentro de los planes de estudios correspondientes a cada curso. Se trata de programas de inmersión lingüística, donde algunas asignaturas, como la geografía, ciencias naturales, física, música, etc., se imparten en una lengua que es diferente a la de la escolarización. Estos programas han proliferado en España en las últimas décadas al menos por dos razones: primero, España es un territorio donde se hablan otros idiomas aparte del castellano (catalán, vasco), por tanto, este tipo de programas podrían ser adecuados para promover el trilingüismo, y segundo, en España hay una necesidad grande de mejorar las competencias en idiomas extranjeros, sobre todo en inglés, después de que esto se identificó como una de las debilidades del sistema educativo español: según el Eurobarómetro 386 (Comisión Europea, 2012) o los datos de Eurydice sobre la enseñanza de los idiomas en las escuelas (Eurydice, 2023), España se coloca consistentemente detrás de los países de la Unión Europea en cuanto al dominio de los idiomas extranjeros.

Un ejemplo de programa es la educación bilingüe y bicultural que ofrece el Ministerio de Educación conjuntamente con el British Council a través de un acuerdo firmado en el año 1996, que sigue el modelo del British Council School en Madrid, llamado Programa Educativo Bilingüe (PEB)[5]. El programa alcanza actualmente 90 centros de Educación Infantil y Primaria y 58 institutos de Educación Secundaria, de 10 Comunidades Autónomas. El PEB dedica aproximadamente un 40% del tiempo lectivo semanal a la enseñanza en inglés de distintas asignaturas tales como Ciencias, Historia o Geografía.

---

[5] Véanse más detalles en la página web del programa: https://www.britishcouncil.es/programas/educacion/educacion-bilingue

Otro ejemplo de programa que es uno de los más ambiciosos en España se puso en marcha en la Comunidad de Madrid en el año 2004. Consistió en la introducción de la educación bilingüe (en español y en inglés) en el primer curso de educación Primaria, en algunos colegios públicos seleccionados. Posteriormente, este programa se extendió también a la educación Secundaria (desde 1º hasta 4º de la ESO). En estos colegios denominados bilingües es obligatorio que al menos un tercio de las clases se desarrollen en inglés y que las clases de inglés se den cinco veces a la semana. Además, es obligatorio que las clases de lengua castellana y matemáticas se den en castellano.

Para entrar en el programa bilingüe, un colegio tiene que hacer una solicitud y cumplir ciertos requisitos: a) tener el apoyo de los profesores del colegio y del consejo escolar (compuesto por el director y algunos profesores y padres); b) la viabilidad de la solicitud que tiene en cuenta la experiencia previa del colegio, los profesores y en particular los profesores con la especialización en inglés, los recursos del colegio y el número de clases y alumnos; y c) la existencia de una distribución equilibrada de los colegios seleccionados entre las distintas áreas geográficas. Dentro de este programa, los alumnos aprenden el inglés como un idioma extranjero, pero también estudian otras asignaturas (ciencias, historia, geografía, educación física, artes, o religión o valores) en inglés. De esta forma, entre al menos a un tercio y un 40% de la enseñanza se desarrolla en inglés (entre 10 y 12 horas del total de 25 horas de clases a la semana).

### 3.2. La educación bilingüe en otros países

#### *Los Países Bajos*

En el sistema educativo holandés, la educación bilingüe consiste en una inmersión parcial en el idioma extranjero, que suele ser el inglés, enseñándose aproximadamente un 50% de las asignaturas en inglés. El contenido de algunas materias se enseña en el idioma extranjero, pero el enfoque de la enseñanza es sobre los conocimientos de la asignatura y no tanto sobre el idioma en el que se enseña (Admiraal, Westhoff y de Bot, 2006). La primera escuela que empezó a usar el inglés como idioma de enseñanza en algunas materias fue en 1989 y actualmente hay unas 120 escuelas. Normalmente estos programas se ofrecen en el nivel de educación Secundaria inferior que correspondería al grupo de edad de 12-15 años. Los colegios tienen cierta libertad para diseñar su currículo, pero

tienen que cumplir unos requisitos importantes del Ministerio de Educación holandés: el 50% de las clases se pueden enseñar en inglés, se tiene que seguir el currículo pre-universitario holandés y, además, se exige que estos programas no puedan afectar el nivel de competencias de los alumnos en holandés. Las asignaturas que se enseñan en inglés son, en su gran mayoría, historia, geografía y biología. Además, para alcanzar el 50% del currículo en inglés, los colegios pueden ofrecer en inglés otras asignaturas como artes dramáticas, dibujo, manualidades, educación física, matemáticas, y en los cursos más avanzados, economía, física y química. En los últimos dos años de Educación Secundaria superior, el holandés es el idioma más utilizado en las clases, puesto que los exámenes de final de Secundaria son en holandés.

***Estados Unidos***

Un caso particular de educación bilingüe es el de los EE. UU., donde hay una proporción importante de la población formada por inmigrantes. Un número considerable de niños inmigrantes con competencias bajas en inglés tienen que atender a clases en escuelas Primarias y Secundarias públicas donde se les enseña en inglés. Como consecuencia, esta barrera lingüística se traslada muchas veces en unos resultados educativos bajos de estos alumnos. Por ello, en varios estados de los EE. UU. las escuelas ofrecen programas que entran en las siguientes dos categorías: aquellos que usan parcialmente el idioma nativo del alumno como idioma de la enseñanza (educación bilingüe) y aquellos donde el inglés se usa como idioma de la enseñanza y también se enseña como un segundo idioma.

La gran mayoría de los programas de educación bilingüe en los EE. UU. están disponibles sobre todo para los inmigrantes hablantes de español, dado que representan una proporción importante de los inmigrantes en este país. Por tanto, la participación de un alumno en un programa bilingüe no es aleatoria y depende de muchos factores como su idioma materno, curso escolar, competencias en inglés, los padres (nivel de renta, educación, si aceptan el programa bilingüe que se les ofrece), el distrito de residencia (a veces la disponibilidad de estos programas en un distrito requiere la presencia de un número suficiente de alumnos con el mismo idioma nativo y el mismo nivel de inglés y que estén en el mismo curso académico), el estado (en algunos estados la educación bilingüe es obligatoria, mientras que otros la prohíben). Todas estas variables pueden afectar el rendimiento académico de los alumnos, por tanto, a la hora de evaluar estos programas si no se tienen en cuenta en las estimaciones generan un sesgo de variables

omitidas. Por ello, estimar el impacto causal de un programa bilingüe con estas características es un reto para la investigación académica (Chin, 2015).

### *Los países de América Latina*

Muchas escuelas en los países de América Latina hablantes de español también empezaron a ofrecer a sus alumnos educación bilingüe. Banfi y Day (2004) documentan este proceso para el caso de Argentina, donde la gran mayoría de los colegios bilingües pertenecen al sector privado. En estas escuelas, el contenido de algunas asignaturas se enseña en el idioma extranjero (en su gran mayoría es el inglés), pero a diferencia de los anteriores ejemplos de educación bilingüe, los colegios bilingües en Argentina representan un grupo muy diversificado en cuanto al tamaño, currículo, antigüedad y tradiciones, recursos o conexiones con el país del idioma extranjero. Por ello, la tarea de evaluar los programas bilingües es muy complicada y de hecho no existen estudios académicos que analicen o evalúen a estas escuelas en Argentina.

Un ejemplo parecido es el caso de Colombia. Ordoñez (2004) argumenta que la adquisición de un idioma extranjero, en particular el inglés, es una parte muy importante del sistema educativo en Colombia. Hay mucha presión política y social para un acceso desde una edad temprana y el aprendizaje del inglés está muy extendido a través de programas de inmersión en este idioma. Los colegios bilingües pertenecen al sector privado y ofrecen aproximadamente un 50% del tiempo de la enseñanza en inglés y el otro 50% en español. En estos programas, los alumnos usan el inglés desde la edad de cuatro o cinco años en asignaturas como ciencias, matemáticas o artes. Las actividades extraescolares se pueden desarrollar en ambos idiomas, en función de la disponibilidad de los profesores en inglés. En la gran mayoría de estos colegios, las competencias en lectura se enseñan primero en español y luego se extienden al inglés después del primer curso de Primaria. Ordoñez (2004) analiza los efectos de este tipo de educación bilingüe en la competencia narrativa oral de alumnos de 15 años, que llevan diez años estudiando en esos colegios en Colombia y encuentra que los alumnos bilingües muestran un rendimiento peor en estas competencias que los monolingües, tanto en español, como en inglés. Sin embargo, la autora de este estudio reconoce la necesidad de más investigaciones para apoyar esos resultados. Ella concluye que, aunque no se puede pretender que los bilingües tengan las competencias de los monolingües en ambos idiomas, es necesario un análisis cuidadoso de los costes que podría suponer para el primer idioma (el idioma nativo) la adquisición de

un segundo idioma en el colegio desde edades tempranas. Esto podría ayudar a los docentes a diseñar métodos de enseñanza para minimizar estos costes.

## 4. LA EVALUACIÓN DE LOS PROGRAMAS DE EDUCACIÓN BILINGÜE

Todos los programas descritos en la sección anterior involucran importantes recursos económicos y humanos, tantos públicos como privados, por ello es importante evaluar su impacto. En esta sección se presentan detenidamente algunos ejemplos de evaluaciones de programas de educación bilingüe. Los objetivos principales de estas evaluaciones suelen ser analizar, por un lado, las competencias lingüísticas de los alumnos en el idioma extranjero del programa, y, por otro lado, los conocimientos de los alumnos en las asignaturas que se estudian en el idioma extranjero.

Hay bastante escasez de estudios que evalúen programas de educación bilingüe y los que existen utilizan datos de alguna región o algún país donde se han implementado este tipo de programas. A la hora de evaluar el éxito de estos programas es importante tener en cuenta los entornos sociales, políticos y económicos que se analizan. Por ejemplo, en el caso de España, se trata de un entorno social en el que los alumnos casi no tienen que utilizar el inglés en su vida diaria fuera del colegio. En cambio, en otros países como los Países Bajos, los niños están expuestos al idioma inglés desde edades muy tempranas y en muchos entornos, al igual que los niños de inmigrantes en países como los EE. UU. Por tanto, los resultados que se obtienen al evaluar estos programas en un país o región determinada no se pueden generalizar fácilmente ni extrapolar a otros países.

### 4.1. Metodología y resultados de la evaluación de un programa de educación bilingüe en España

#### *4.1.1. Descripción de los datos y la metodología de estimación*

Una base de datos muy utilizada en varios trabajos de investigación en economía de la educación en España es la de los resultados en una prueba estandarizada para alumnos de 6º de Primaria (12-13 años) en la Comunidad de Madrid, llamada la prueba de Conocimientos y Destrezas Indispensables (CDI). Esta prueba evalúa las habilidades consideradas como imprescindibles para alumnos de 6º de Primaria. El examen se realizó desde el año académico 2004-2005 hasta

el año 2014-2015 y era obligatorio para todos los colegios de la Comunidad de Madrid (públicos, concertados y privados). Igual que la prueba PISA de la OCDE, el resultado en el examen no tiene ninguna consecuencia académica para los alumnos, su objetivo siendo el de proporcionar información adicional a los profesores, estudiantes y padres.

El examen consiste de 2 partes, cada una de 45 minutos: la primera parte incluye pruebas de dictado, lectura y conocimientos generales (ciencias naturales y sociales) y la segunda parte incluye ejercicios de matemáticas. Las pruebas se realizan en castellano para todos los alumnos, independientemente de si están o no atendiendo clases en un colegio con un programa bilingüe.

Antes de realizar el examen, los alumnos tienen que rellenar un cuestionario breve, donde se les pregunta sobre ellos mismos, sus padres y el entorno socioeconómico en el que viven. Las respuestas a estas preguntas proporcionan información sobre varias características de los alumnos: sexo, edad, país de nacimiento, nivel educativo de los padres, ocupación de los padres, composición del hogar dónde vive el alumno, edad a la que el alumno empezó a ir al colegio (educación infantil), si el alumno tiene alguna necesidad especial o alguna discapacidad.

Con la ayuda de estos datos, Anghel, Cabrales y Carro (2016) evalúan la implementación del programa bilingüe en la Comunidad de Madrid de la siguiente manera: se comparan los resultados de los alumnos de las escuelas donde se implement**ó** el programa bilingüe con los resultados de los alumnos de escuelas sin este tipo de programa. Además, como se sabe que el primer año en el que se implementó el programa fue 2004-2005, se pueden comparar las cohortes de alumnos que entraron justamente ese año en el colegio y que hicieron luego en sexto de Primaria la prueba CDI, con las cohortes que entraron en el colegio un año antes y que no están dentro del programa de educación bilingüe.

Anghel, Cabrales y Carro (2016) parten del siguiente modelo econométrico de la función de producción de la educación (véase también el capítulo 2 de este libro). Para un año académico, la nota en la prueba para un alumno *i* en la escuela *j* viene determinada por:

$$y_{ij} = \delta bil_j + \beta x_i + v_j + u_i + \varepsilon_{ij} \qquad (1)$$

donde $x_i$ incluye las características observables de los alumnos y de los padres, $bil_j$ es una variable *dummy* que indica si la escuela *j* participa en el programa bilingüe, $v_j$ son características no observables del colegio (por ejemplo, la calidad

de los profesores o del director del colegio), $u_i$ son características no observables de los alumnos (por ejemplo, el esfuerzo o la habilidad innata) y $\varepsilon_{ij}$ es un error aleatorio.

El parámetro de interés es el coeficiente $\delta$ que mide el impacto medio del programa bilingüe sobre $y_{ij}$. Sin embargo, al estimar la regresión de $y_{ij}$ sobre $bil_j$ y $x_i$ con los datos disponibles podría haber un sesgo de endogeneidad debido a dos problemas relacionados con la autoselección:

1) Los estudiantes no se distribuyen de forma aleatoria entre los colegios (véase también el capítulo 5a de este libro). Es decir, los alumnos, o, mejor dicho, los padres de los alumnos, eligen el colegio. En el caso de la Comunidad de Madrid, si no hay un exceso de demanda para las plazas en el colegio elegido, todos los niños están admitidos. En el caso en el que hay exceso de demanda, la admisión se basa en criterios como la proximidad de la vivienda del alumno al colegio y la renta de la familia. Ninguna de estas dos variables es aleatoria y están correlacionadas con el rendimiento académico de los alumnos.

2) Los colegios no se asignan de forma aleatoria para implementar el programa bilingüe. En la sección 3.2 se ha explicado que el programa se implementa solamente en aquellos colegios que lo solicitan, y, además, los colegios que solicitan el programa y que acceden a él tienen que cumplir ciertos requisitos. La solicitud para implementar un programa bilingüe podría ser una señal positiva de la calidad del director y de los profesores de la escuela, puesto que el programa implica un trabajo extra significativo por parte de ellos. Pero al mismo tiempo podría ser una señal de que el colegio tiene una baja demanda de admisiones, posiblemente debido a su baja calidad.

Para controlar la endogeneidad causada por los dos problemas de auto-selección, el artículo de Anghel, Cabrales y Carro (2016) proponen usar el método de diferencias en diferencias[6]. Para este contexto, este método consiste en comparar las notas de los alumnos en los colegios tratados (bilingües) antes y después del tratamiento –es decir, la implementación del programa bilingüe– con el grupo de colegios de control (no-tratados, es decir no bilingües) antes y

[6] Véase el capítulo 2 de métodos econométricos.

después del tratamiento. El grupo de tratamiento lo constituye la primera cohorte de alumnos en los colegios bilingües de la Comunidad de Madrid. Estos alumnos empezaron el curso de primero de Primaria en el año académico 2004-2005, por tanto, hicieron la prueba CDI seis años después, cuando estaban en sexto de Primaria, en el año 2009-2010. Esta cohorte de alumnos está formada por 25 colegios que fueron los primeros en implementar el programa bilingüe.

Se usan los datos de la prueba CDI de los años académicos 2008-2009 y 2009-2010, y los cuatro grupos de colegios que se analizan son los siguientes: el grupo de colegios bilingües en 2008-2009 (el grupo de tratamiento antes del tratamiento), el grupo de colegios no bilingües en 2008-2009 (el grupo de control antes del tratamiento), el grupo de colegios bilingües en 2009-2010 (el grupo de tratamiento después del tratamiento) y el grupo de colegios no bilingües en 2009-2010 (el grupo de control después del tratamiento).

Este método soluciona el problema del sesgo de autoselección de los colegios en el programa dado que se observa el rendimiento académico en la prueba CDI de los alumnos del mismo colegio en el primer año en el que el programa bilingüe se implementó y también en el año anterior. Otro aspecto importante para la estrategia de identificación del impacto es que se puede suponer que los únicos cambios significativos en los recursos materiales y en el personal de los colegios de un año al otro tuvieron que ver con el programa bilingüe.

En cuanto al problema del sesgo de autoselección de los alumnos en los colegios, el método de diferencias en diferencias también ayuda a solucionarlo, en la medida en la que se puede suponer que las diferencias entre la primera cohorte de alumnos tratados y las cohortes anteriores no están relacionadas con la introducción del programa. Para ello, Anghel, Cabrales y Carro (2016) realizan un análisis muy detallado de las características de los alumnos de los grupos de tratamiento y de control. Los autores observan un cierto cambio en las características de los alumnos en los colegios bilingües después de la introducción del programa, no muy grande, pero posiblemente endógeno. Este cambio consiste en una mejora de algunas características de los alumnos en los colegios tratados como el nivel educativo de los padres, la ocupación de los padres, o la nacionalidad, factores que la literatura académica ha identificado como determinantes del éxito de los alumnos (véase el capítulo 3b). Para solventar este problema y controlar por estos cambios, el método de diferencias en diferencias permite incorporar las características observables de los alumnos en la estimación. Por tanto, los autores estiman la siguiente regresión:

$$y_{ij} = \alpha_0 + \alpha_1 bil_j + \alpha_2 año\ 2010 + \delta(año\ 2010 * bil_j) + \beta x_i + \varepsilon_{ij} \qquad (2)$$

donde la variable es una variable *dummy* para el año académico 2009-2010, que es el primer año en el que se observan a los niños del programa bilingüe en la prueba CDI.

### *4.1.2. Resultados empíricos*

En la ecuación (2), el parámetro $\delta$ asociado al término de interacción *año2010*$bil_j$ indica el efecto del programa bilingüe sobre los resultados de los alumnos en las tres pruebas del examen CDI: matemáticas, lectura y conocimientos generales. Cuando se estima la regresión de la ecuación (2) sin ninguna variable de control $x_i$, los autores encuentran que el impacto del programa no es estadísticamente significativo (véase tabla 8 en Anghel, Cabrales y Carro, 2016). Cuando se incluyen las variables de control, que son características de los alumnos, el impacto del programa se reduce, lo que refleja que puede haber cierta autoselección de los alumnos en los colegios, una vez que se implementa el programa bilingüe. Para matemáticas y lectura, el coeficiente no es estadísticamente significativo, aunque pasa de ser positivo a ser negativo. Sin embargo, para los resultados en la prueba de conocimientos generales, el programa bilingüe tiene un impacto negativo y estadísticamente significativo. Tal y como se ha explicado en la sección 3.1 en la descripción del programa bilingüe en la Comunidad de Madrid, la asignatura de conocimientos generales es la única de las tres pruebas evaluadas que se enseña en inglés en los colegios bilingües, pero se evalúa en castellano en la prueba CDI. Por tanto, este resultado indicar**ía** por un lado que la enseñanza de una asignatura en inglés en Educación Primaria podría empeorar el rendimiento de los alumnos. Por otro lado, esto podría mostrar que posiblemente los alumnos tengan los conocimientos necesarios, pero que simplemente no sean capaces de expresarlos en castellano puesto que lo aprenden en inglés. No obstante, si este es el caso, se trataría también de un impacto negativo del programa, dado que los alumnos no tienen el nivel de inglés necesario para superar la barrera lingüística.

Los autores de este estudio realizan varios ejercicios de robustez de sus resultados, pero todos llevan a la misma conclusión: hay un efecto negativo claro y cuantitativamente importante del programa bilingüe sobre la asignatura que se enseña en inglés, y el efecto no es estadísticamente diferente de cero en las habilidades de matemáticas y lectura en castellano. Asimismo, los autores resaltan dos aspectos de este resultado que son muy relevantes para sus implicaciones de

políticas educativas: primero, el impacto negativo se concentra en el grupo de alumnos cuyos padres tienen un nivel educativo bajo y segundo, cuando realizan las mismas estimaciones para el grupo de colegios que implementaron el programa bilingüe en el segundo año (2005-2006), el impacto negativo se reduce y ya no es estadísticamente significativo, ni en promedio para toda la muestra de alumnos y tampoco para la muestra de alumnos con padres menos educados. Una posible explicación para este segundo resultado es que posiblemente los colegios seleccionados en el segundo año son mejores y mejor preparados para la implementación del programa, dado que por ejemplo se requiere un nivel más alto de conocimientos de inglés de los profesores en estos colegios. Sin embargo, esta es solamente una hipótesis que no se puede comprobar con los datos utilizados por estos autores.

## 4.2. Metodología y resultados de evaluaciones de programas de educación bilingüe en otros países

### *4.2.1. Los Países Bajos*

El estudio de Admiraal, Westhof y de Bot (2006) evalúa el programa de educación bilingüe en las escuelas de Educación Secundaria en los Países Bajos, descrito en la sección 3.2 del capítulo. A través de varios estudios piloto, los autores construyen un estudio de evaluación longitudinal. Su principal interés se centra en los efectos del uso de inglés como idioma de enseñanza sobre las competencias de los alumnos en inglés como idioma extranjero, aunque también analizan el impacto de la educación bilingüe sobre las asignaturas que se enseñan en inglés, así como sobre las competencias en holandés.

Las competencias de los alumnos en inglés se evalúan a través de las siguientes medidas: los conocimientos de vocabulario receptivo en inglés –un test de vocabulario denominado "English as a Foreign Language Vocabulary Test" (test de vocabulario EFL)–, la comprensión lectora –un examen final de inglés a nivel nacional para Educación Secundaria intermedia– y la competencia oral –un examen diseñado para el último año de la Educación Secundaria intermedia–.

En cuanto a los conocimientos en las asignaturas que se enseñan en inglés, se usan los resultados de los exámenes finales de la Educación Secundaria superior en las asignaturas de geografía e historia que se imparten en inglés en las escuelas bilingües.

Los autores también disponen de características de los alumnos que podrían estar correlacionadas con el nivel de competencias y de adquisición del inglés: sexo, los resultados de una prueba al final de la Educación Primaria, la lengua hablada en casa, contacto con la lengua fuera del colegio, y motivación de aprender el inglés.

Para estimar el impacto del programa, utilizan una muestra de alumnos que están matriculados en colegios con programas de educación bilingüe (el grupo de tratamiento) y otra muestra de alumnos que siguen los cursos regulares de Educación Secundaria en holandés (el grupo de control) –un total de 1.305 alumnos–. La información cubre un periodo de seis años académicos. Los datos tienen una estructura jerárquica: las notas de los alumnos en distintos exámenes, para cuatro cohortes de alumnos en nueve escuelas. Los datos se analizan a través del análisis de regresión multinivel que tiene en cuenta esta estructura. Además, dado que disponen de un número reducido de observaciones, los autores del estudio utilizan el método de *Restricted Iterative Generalised Least Squares* (RIGLS; véase Goldstein, 2011).

El análisis de este trabajo muestra que no hay ninguna diferencia significativa entre el grupo de alumnos del programa bilingüe y el grupo de control con respecto a la puntuación media en la prueba de vocabulario que es la medida de conocimientos de vocabulario receptivo. Sin embargo, tanto en comprensión lectora como en competencias orales en inglés, los alumnos del programa bilingüe obtienen resultados mejores que los del grupo de control que atienden el programa normal, incluso cuando se tienen en cuenta las características observables y relevantes de los alumnos. Por último, en cuanto a las asignaturas enseñadas en inglés (historia y geografía) y la lengua holandesa, no se encuentra ningún impacto del programa bilingüe.

Los autores de este estudio sugieren que es difícil generalizar las conclusiones obtenidas para la totalidad de escuelas que ofrecen programas bilingües en los Países Bajos, puesto que ellos analizan una muestra pequeña de alumnos y colegios que fueron los pioneros en implementar este tipo de educación. Al igual que el estudio para los colegios de la Comunidad de Madrid (Anghel, Cabrales y Carro, 2016), Admiraal, Westhoff y de Bot (2006) señalan que su investigación describe los efectos de la educación bilingüe en los colegios de Secundaria en los primeros años de implementación, antes de que se transformase en un programa educacional completamente desarrollado con alrededor de 120 colegios de

educación Secundaria. Por tanto, no está claro si los resultados serían los mismos para las escuelas que implementaron la educación bilingüe más tarde, dado que posiblemente estas se beneficiarían de la experiencia y de los conocimientos de las escuelas pioneras.

Por último, los autores también enfatizan que su análisis se centra en el caso de un país donde el inglés tiene una posición muy importante en la sociedad en general. Los niños están muy motivados en aprender el inglés, que es un idioma con un lugar destacado en la sociedad holandesa, en el sistema educativo y en el mundo científico, y, por tanto, están muy expuestos al idioma inglés también fuera de sus colegios.

### *4.2.2. Estados Unidos*

Un ejemplo de evaluación de un programa bilingüe en los EE. UU. es un estudio del estado de Texas (Chin, Daysal e Imberman, 2013), donde un distrito escolar[7] está obligado a ofrecer educación bilingüe ("*bilingual education*", BE) cuando 20 o más alumnos con competencias bajas en inglés ("*limited English proficiency students*", alumnos LEP) están matriculados en el mismo curso y hablan el mismo idioma nativo. Por debajo de este umbral de 20 alumnos, los distritos escolares tienen la libertad de ofrecer educación bilingüe o enseñar todas las asignaturas en inglés y ofrecer el inglés como un segundo idioma ("*English as a second language*", ESL). En esta situación, la mayoría eligen esta segunda opción dado que es la menos costosa.

La mayoría de los programas bilingües en los EE. UU. se ofrece a los inmigrantes nativos españoles. En el caso de Texas, el español es el idioma materno de aproximadamente un 90% de los alumnos inmigrantes matriculados en ESL.

Los programas BE y ESL tienen tanto ventajas como desventajas para los alumnos LEP. Por ejemplo, la enseñanza en el idioma nativo del alumno podría retrasar el aprendizaje del inglés, pero al mismo tiempo podría ayudar a los alumnos a estar al día con las matemáticas y otras asignaturas, mientras están aprendiendo el inglés. Por otro lado, un programa ESL podría mejorar el aprendizaje tanto del inglés como de otras materias a través de una redistribución del

[7] En los EE. UU. la mayoría de las escuelas funcionan como unidades de un distrito escolar que suele gestionar varias escuelas.

tiempo que se utilizaría para enseñar a los alumnos LEP a leer y escribir en su idioma nativo.

Además, para la evaluación de estos programas hay que tener en cuenta también el impacto de la presencia de estos alumnos LEP sobre sus compañeros que no son LEP (véase también el capítulo 3c sobre el efecto de los compañeros de clase). Este impacto se podría producir a través de varios mecanismos. Por ejemplo, los profesores con alumnos LEP tendrían que prestar más atención a estos en detrimento del tiempo destinado a los otros estudiantes. Además, en la medida en la que el rendimiento académico de los alumnos LEP es más bajo que el de los alumnos no-LEP, podría haber un efecto negativo de los compañeros LEP sobre los resultados de los alumnos no-LEP. Sin embargo, también podría haber efectos positivos de los compañeros LEP porque podrían estar más avanzados en otras dimensiones, como por ejemplo los conocimientos de otras asignaturas, habilidades no-cognitivas, etc. que son relevantes para el rendimiento académico o también podría haber ventajas de estar en una clase más diversa.

En este contexto, identificar el efecto causal de la educación bilingüe sobre el rendimiento de los alumnos es complicado, puesto que la provisión de la BE en un distrito escolar es una variable endógena. Los distritos que ofrecen BE no se asignan de forma aleatoria: podrían ser los que tienen más alumnos LEP, más aumento en el número de alumnos LEP, más riqueza, alumnos LEP con un nivel considerablemente bajo del inglés, etc.

Para corregir este sesgo de endogeneidad los autores utilizan el umbral de 20 alumnos que les ofrece la posibilidad de diseñar una regresión discontinua[8]. En concreto, el efecto de la BE (relativo a la enseñanza del inglés como ESL) sobre el rendimiento académico de los alumnos se puede obtener comparando los resultados de los alumnos en los distritos escolares justo por encima de este umbral (donde es más probable que se ofrezca BE) con los resultados en los distritos justo por debajo. Por poner un ejemplo, los distritos escolares con 19 alumnos LEP no tendrían que ser muy diferentes que los distritos escolares con justo 20 alumnos LEP, pero dada la existencia de la política, estos últimos están obligados a ofrecer BE. Por tanto, se puede usar la diferencia entre los resultados

[8] Véase el capítulo 2.

en los distritos con 19 y los distritos con 20 alumnos para identificar el efecto de la provisión de BE en un distrito.

Las estimaciones a través de esta metodología llevan a las siguientes conclusiones. Primero, la provisión de programas BE a nivel de distrito escolar en general tiene un efecto positivo pero muy pequeño y estadísticamente no significativo sobre los alumnos cuyo idioma materno es el español. Estos alumnos serían los que cumplen los requisitos para participar en un programa de BE, por tanto, estos resultados indicarían que los beneficiarios de los programas para los alumnos LEP obtienen resultados similares en los programas de BE y de ESL.

Segundo, la provisión de programas de BE a nivel de distrito escolar tiene un impacto estadísticamente positivo sobre el rendimiento en matemáticas y lectura de los alumnos que no son LEP y cuyo idioma nativo no es el español. Esto indica la presencia de efectos *spillover* de estos programas.

## 5. OTROS EJEMPLOS DE PROGRAMAS O REFORMAS QUE AFECTAN EL IDIOMA DE LA ENSEÑANZA EN LAS ESCUELAS

### *Reformas que cambian el idioma de la enseñanza en la escuela*

Varios estudios se centran en algunos países donde por razones históricas el idioma de enseñanza en todas (o algunas) las asignaturas es diferente al idioma nativo. Angrist y Lavy (1997) analizan el impacto de una reforma en Marruecos en el año 1983 que sustituyó el idioma de enseñanza en Educación Secundaria del francés al árabe. Utilizando el método de diferencias en diferencias, encuentran un impacto negativo en las habilidades de escritura en francés de los individuos expuestos a la reforma. Además, este impacto negativo se traslada en un impacto negativo sobre los salarios en el caso de los individuos afectados.

Al contrario, Angrist, Chin y Godoy (2008) en un estudio que analiza una reforma en Puerto Rico que cambió el idioma de la enseñanza del inglés al español en la educación post-Primaria, no encuentran ningún efecto de la reforma en las competencias en inglés, medidas a través del nivel de las habilidades de hablar el inglés. Los autores sugieren que este resultado demuestra que la enseñanza intensiva en otro idioma (inglés en ese caso) no es suficiente para mejorar las habilidades en inglés y no es necesaria tampoco para tener unas buenas habilidades del inglés.

*El caso de los países con varios idiomas oficiales*

Un caso particular es el de los países o regiones donde coexisten dos o más idiomas oficiales y el uso de idiomas múltiples en el sistema educativo es muy común. En este sentido, Sudáfrica es un ejemplo muy relevante, puesto que hay 11 idiomas oficiales que coexisten (nueve idiomas africanos y dos idiomas europeos –el inglés y el afrikáans–, este último derivado del neerlandés). La mayoría de las escuelas de Educación Primaria utilizan la enseñanza en el idioma nativo durante los primeros tres años y luego cambian al inglés en el cuarto curso, mientras que algunas escuelas usan el inglés como idioma de la enseñanza, aunque para la mayoría de los alumnos este no es el idioma nativo. Además, los exámenes de final de la educación Secundaria se realizan en inglés o en afrikáans, sin embargo, solamente el 23% de los surafricanos hablan alguno de estos dos idiomas (Taylor y von Fintel, 2016). Dado que en Sudáfrica hay mucha desigualdad dentro del sistema educativo público, el impacto del idioma sobre los resultados educativos de los niños podría ser muy relevante.

Para analizar este impacto, Taylor y von Fintel (2016) utilizan algunos cambios en las normas sobre los idiomas de la enseñanza en algunas escuelas, que ofrecen una variación exógena que se usa para identificar el impacto de estas normas. Los autores construyen una base de datos que cubre el periodo 2007-2012 para toda la población de escuelas de Sudáfrica. Los resultados estimados utilizando el método MCO indican que la enseñanza en inglés en edades tempranas se asocia con unos rendimientos mejores en inglés en los cursos escolares posteriores, Sin embargo, cuando usan un modelo con efectos fijos de escuelas (que incluyen características como la calidad de la escuelas), encuentran que estar en las escuelas donde se usa el idioma nativo como idioma de la enseñanza en los primeros tres años lleva a un mejor rendimiento de inglés en los siguientes cursos. Según los autores, este resultado está en consonancia con la teoría pedagógica que promueve el uso del primer idioma como idioma de la instrucción durante tres hasta aproximadamente seis años. En este periodo, los niños alcanzarían un cierto nivel de competencia académica en su primer idioma, que permitiría que posteriormente la enseñanza se cambiase a un segundo idioma (Hakuta, Butler y Witt, 2000).

Ivlevs y King (2014) explotan una reforma en Letonia en el año 2004, cuando el idioma de enseñanza en las escuelas de la minoría rusa de este país cambió desde un 100% ruso a un 60% letón y un 40% ruso. En concreto, analizan el rendimiento de los colegios de la minoría rusa y de los colegios de la mayoría letona, en un examen estandarizado al final de la Educación Secundaria que es idéntico para todas

las escuelas del país. La reforma y el examen permiten la construcción de un grupo de tratamiento (el grupo de colegios de la minoría rusa que están afectados por la reforma) y el grupo de control (el grupo de colegios letones no afectados por la reforma). El resultado principal es que el rendimiento de los colegios de minoría rusa sufrió un deterioro importante después de la reforma, relativo a los colegios del grupo de control, sobre todo en los años en los que las primeras cohortes tratadas se graduaron. Sin embargo, resultados preliminares de su análisis muestran que, en años posteriores, en algunas materias, el rendimiento de los colegios de la minoría rusa recupera los niveles de antes de la reforma, indicando que los colegios tratados posiblemente se hayan ajustado al cambio. No obstante, los autores señalan que se necesitaría información de un número más grande de años posteriores a la reforma para confirmar este resultado.

El idioma de la enseñanza es relevante también en la educación Terciaria, puesto que la movilidad internacional de los estudiantes está en constante crecimiento. Bernhofer y Tonin (2022) utilizan datos administrativos de la Universidad Libre de Bozen-Bolzano, que está ubicada en una región bilingüe de Italia, donde el 65% de la población usa el alemán como el idioma materno y el 27% usa el italiano como idioma materno. La Universidad es trilingüe –inglés, alemán e italiano–, lo que significa que los estudiantes pueden aprender y examinarse en los tres idiomas. Los autores encuentran una reducción importante en las notas de los estudiantes en exámenes que se realizan en un idioma diferente al idioma nativo, lo que podría indicar que estudiar una asignatura en un idioma que no es el nativo podría conllevar ciertos costes.

## 6. CONCLUSIONES

Este capítulo analiza la importancia de los idiomas extranjeros como un input relevante en la función de producción de educación de un individuo. Los conocimientos de idiomas son fundamentales para el progreso de un individuo en la sociedad en general, y en el mercado laboral en particular. La importancia de los idiomas extranjeros se ve reflejada también en el esfuerzo que hará la OCDE dentro de la prueba de PISA 2025 que evaluará las competencias en lenguas extranjeras de los alumnos a una escala global[9].

---

[9] PISA 2025 realizará una prueba opcional de inglés que evaluará tres habilidades en ese idioma: la lectura (*reading*), la escucha (*listening*) y el habla (*speaking*). Véase OCDE (2021).

El capítulo muestra que hay suficiente evidencia empírica a favor de los beneficios económicos de los idiomas extranjeros, en cuanto a mejores salarios y mejor situación en el mercado laboral. Dados estos beneficios, en la gran mayoría de los países se invierten recursos importantes tanto públicos como privados para la enseñanza de los idiomas, en algunos casos desde edades tempranas. La evaluación de algunos programas educativos que implementan el uso de un segundo idioma de la enseñanza en los colegios en diferentes niveles educativos (Primaria, Secundaria) muestra resultados distintos que dependen mucho del país o la región, del entorno social, del sistema educativo, de las características del programa. Por tanto, la generalización de los resultados de estos estudios es muy complicada y se tiene que hacer con mucha cautela. Para obtener resultados comparables entre los países se tendría que disponer de pruebas estandarizadas a nivel internacional, donde a los alumnos se les evalúen las competencias en un idioma extranjero estudiado en la escuela (el inglés, por ejemplo) de la misma manera. Además, se tendría que disponer de características socioeconómicas de los alumnos y sus familias, de características de las escuelas y de los sistemas educativos de los países.

Asimismo, para poder hacer una evaluación completa y rigurosa de estos programas se tendría que disponer de bases de datos donde se pueda seguir a los alumnos de los programas de educación bilingüe a lo largo de su vida académica y profesional y obtener información sobre, por ejemplo, sus carreras universitarias, si las hay, o sobre su situación en el mercado laboral (salarios, estado laboral, etc.). Disponer de esta información permitiría realmente evaluar si estos programas han conseguido lograr sus objetivos.

Una limitación de los estudios analizados es que no tienen en cuenta la calidad de los profesores que enseñan en los programas de educación bilingüe. Tal y como se explica en el capítulo 3d sobre profesores, estos son un factor determinante del rendimiento académico de los estudiantes. En el caso de los programas que implementan el uso de un idioma extranjero como idioma de la enseñanza, los profesores tienen que cumplir también unos requisitos extra relacionados con el nivel de competencia en el idioma extranjero, lo que exigiría mayores conocimientos y esfuerzo por su parte. Esto podría ser una señal de su motivación y calidad como docentes. Por tanto, información sobre los profesores y la calidad de la docencia sería fundamental para poder explicar mejor los resultados sobre los impactos de este tipo de programas educativos que se han encontrado en los estudios resumidos en este capítulo.

El capítulo abre varias puertas para futuros temas de investigación, como, por ejemplo: a) ¿A qué edad y cómo se deberían introducir los idiomas extranjeros en la educación de los alumnos en las escuelas; b) ¿Se debería usar un idioma extranjero como idioma de la enseñanza de algunas asignaturas en los colegios?; y c) Si la respuesta es sí, ¿a qué edades debería hacerse?

## EJERCICIOS

Ejercicio 1. *En la página web de PISA (https://www.oecd.org/pisa/data/2018 database/) descarga los microdatos de PISA 2018 (también puedes utilizar datos de PISA de otros años, o los datos del último año disponible), que están bajo el nombre "Student questionnaire data file". Utiliza solamente los datos de España (o de cualquier otro país).*

*a) Identifica en la base de datos:*

- *Las variables que corresponden a los conocimientos de idiomas extranjeros de los alumnos.*
- *Las variables que corresponden a las expectativas de los alumnos en cuanto a realizar una carrera universitaria y en cuanto a ser un manager o especialista.*

*b) Con la ayuda de estas variables, realiza una estimación por MCO donde la variable dependiente es: (a) las expectativas de los alumnos en cuanto a realizar una carrera universitaria; y (b) las expectativas de los alumnos en cuanto a ser un manager o especialista, y las variables independientes son: número de idiomas extranjeros, sexo del alumno, el nivel socioeconómico del alumno (medido a través del* PISA index of economic, social and cultural status*), el estatus de inmigrante. De esta forma, comprueba la afirmación que se hace en la Introducción del libro.*

*c) Con la ayuda de estas variables, comprueba si el número de idiomas extranjeros que conoce un alumno está relacionado con sus resultados en algunas de las pruebas de PISA (matemáticas, lectura, ciencias, o competencias globales), teniendo en cuenta su nivel socioeconómico, el estatus de inmigrante y el sexo.*

Ejercicio 2. *¿Cuáles son los problemas metodológicos de estimar los rendimientos salariales de los conocimientos de idiomas extranjeros? Utiliza la ecuación de Mincer explicada en el capítulo 4b para argumentar tus respuestas.*

Ejercicio 3. *El estudio de Anghel, Cabrales y Carro (2016) se detalla en la sección 4.1 de este capítulo. Ellos encuentran que cuando realizan las mismas estimaciones para el grupo de colegios que implementaron el programa bilingüe en el segundo año (2005-2006), el impacto negativo se reduce y ya no es estadísticamente significativo. Discute cuáles podrían ser las explicaciones de este resultado y qué datos se necesitarían para comprobar las hipótesis planteadas.*

## BIBLIOGRAFÍA


Admiraal, W., Westhoff G. y De Bot K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation, 12*(1), 75-93.

Aldashev, A., Gernandt J. y Thomsen S.L. (2009). Language usage, participation, employment and earnings: Evidence for foreigners in West Germany with multiple sources of selection. *Labour Economics, 16*(3), 330-341.

Anghel B., Cabrales A. y Carro J.M. (2016). Evaluating a Bilingual Education Program in Spain: The Impact Beyond Foreign Language Learning. *Economic Inquiry,* Vol. 54, No. 2, 1202-1223.

Angrist, J. D. y Lavy V. (1997). The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco. *Journal of Labor Economics, 15*(1), Part 2, S48-S76.

Araújo, L., Dinis da Costa, P., Flisi, S. y Soto Calvo, E. (2015). *Languages and Employability.* European Commission. https://crell.jrc.ec.europa.eu/sites/default/files/files/JRC97544.pdf

Banfi, C. y Day, R. (2007). The Evolution of Bilingual Schools in Argentina. International *Journal of Bilingual Education and Bilingualism, 7*, 398-411.

Bernhofer, J. y Tonin M. (2022). The effect of the language of instruction on academic performance. *Labour Economics, 78.*

Chin, A., Daysal N. M. e Imberman, S.A. (2013). Impact of bilingual education programs on limited English proficient students and their peers: Regression discontinuity evidence from Texas. *Journal of Public Economics, 107*, 63-78.

Chin, A. (2015). Impact of bilingual education on student achievement. *IZA World of Labor.*

Comisión Europea (2006). Effects on the European Union economy of shortages of foreign language skills in enterprise.

Comisión Europea (2012). Europeans and their languages. Special Eurobarometer 386. https://europa.eu/eurobarometer/api/deliverable/download/file?deliverableId=40204

Comisión Europea / EACEA / Eurydice (2023). Cifras clave de la enseñanza de lenguas en los centros escolares de Europa, edición 2023, Informe Eurydice, Oficina de Publicaciones de la Unión Europea, Luxemburgo.

Costa, A. y Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience, (15)*5, 336-345.

Covacevich, C. y Vargas, J. (2021). *How Language Learning Opens Doors*. OECD Publishing.

Dustmann, C. y Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal*, vol. 113, no.489, 695-717.

Eurydice (2023). Key data on teaching languages at school in Europe, 2023 edition. https://eurydice.eacea.ec.europa.eu/publications/key-data-teaching-languages-school-europe-2023-edition

Fidrmuc, J. y Fidrmuc, J. (2016). Foreign languages and trade: evidence from a natural experiment. *Empirical Economics, 50,* 31-49. https://doi.org/10.1007/s00181-015-0999-7.

Fry, R. y Lowell, L. B. (2003). The Value of Bilingualism in the U.S. Labor Market. *Industrial & Labor Relations Review, 57*, 128-40.

Ginsburgh, V.A. y Prieto-Rodriguez, J. (2011). Returns to foreign languages of native workers in the European Union. *Industrial and Labor Relations Review 64*(3), 599-618.

Goldstein, H. (2011). *Multilevel statistical models*. John Wiley & Sons.

Hakuta, K., Butler, Y. G. y Witt, D. (2001). *How long does it take English learners to attain proficiency?* Policy Report 2000-1. Santa Barbara, CA: The University of California Language Minority Research Institute.

Hutchinson, W. (2002). Does ease of communication increase trade? Commonality of language and bilateral trade. *Scott J Polit Econ, 49*, 544-556.

Hulstijn, J. H. (1997). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition, 19*(2), 131-143.

Isphording, I. E. (2013). Returns to Foreign Language Skills of Immigrants in Spain. *Labour, 27*(4), 443-461.

Ivlevs, A. y King, R. M. (2004). Minority Education Reform and pupil performance in Latvia. *Economics of Education Review, 38*, 151-166.

Melitz, J. (2008). Language and foreign trade. *European Economic Review,* Vol. 52/4, 667- 699.

Melitz J. y Toubal F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics, 93*, 351-363.

OECD (2021). PISA 2025 Foreign Language Assessment and Analytical Framework. OECD Publishing, Paris. Disponible en: https://www.oecd.org/pisa/foreign-language/PISA-2025-FLA-Framework-Learning-Another-Language.pdf

Ordóñez, C. L. (2004). EFL and Native Spanish in Elite Bilingual Schools in Colombia: A First Look at Bilingual Adolescent Frog Stories. *International Journal of Bilingual Education and Bilingualism, 7*, 449-74.

Saiz, A. y Zoido, E. (2005). Listening to What the World Says: Bilingualism and Earnings in the United States. *Review of Economics and Statistics, 87*, 523-38.

Taylor, S. y von Fintel, M. (2016). Estimating the impact of language of instruction in South African primary schools: A fixed effects approach. *Economics of Education Review, 50,* 75-89.

Williams, D. R. (2011). Multiple language usage and earnings in Western Europe. *International Journal of Manpower, 32*(4), 372-393.

# Capítulo 4a
# LA EDUCACIÓN INFANTIL

Álvaro Choi
Departament d'Economia
Universitat de Barcelona
ORCID: https://orcid.org/0000-0003-1286-4124/0000-0003-3394-0181

Gabriela Sicilia
Departamento de Economía, Contabilidad y Finanzas
Universidad de La Laguna
ORCID: https://orcid.org/0000-0003-1286-41240000-00021078-7762

**Resumen**

En este capítulo se reflexiona acerca de diversos elementos a considerar para abordar el análisis de los impactos de la Educación Infantil. No pretende realizar una revisión de la literatura, sino ser un punto de partida que facilite al lector profundizar en el análisis y la evaluación, especialmente de impacto, de la Educación Infantil. Para ello, se proporciona una breve revisión de la evolución reciente de la escolarización infantil en España y se presenta el marco teórico de la producción de Educación Infantil y la decisión de las familias al respecto, subyacente a la mayor parte de evaluaciones de impacto. Por último, se describen las distintas aproximaciones empíricas de la evaluación de impacto de la Educación Infantil –experimentales y cuasiexperimentales–, y destacando algunos elementos fundamentales a tener en cuenta a la hora de realizar este tipo de evaluaciones.

## 1. INTRODUCCIÓN

La Educación Infantil es el primer nivel de los sistemas educativos y abarca desde los primeros meses de edad hasta el momento de inicio de la Educación Primaria, que en la mayor parte de países suele establecerse entre los 5 y los 7 años. Tal y como sucede con la Educación Primaria, la Educación Infantil puede dividirse, a su vez, en distintos ciclos, en función del diseño institucional de cada

país. Así, por ejemplo, en el caso español se diferencia entre un primer ciclo, de los 0 hasta los 3 años, y un segundo ciclo de los 3 hasta los 6 años.

La Educación Infantil se expandió rápidamente en los países desarrollados a partir de la década de 1960, en gran medida como respuesta a la incorporación de la mujer al mercado laboral. Paulatinamente, fue abandonando la concepción asistencialista inicial para ir siendo considerada progresivamente como una etapa vital dentro del proceso de formación de capital humano. La Educación Infantil estimula el desarrollo de las habilidades físicas, sociales, cognitivas, emocionales y lingüísticas y su relevancia queda recogida en el Objetivo de Desarrollo Sostenible (ODS) 4, que establece como un objetivo para el año 2030 "garantizar que todos los niños y niñas tengan acceso a un desarrollo, atención y Educación pre-Primaria de calidad en la primera infancia para que estén preparados para la Educación Primaria" (ONU, 2015). De hecho, el ODS 4 recoge solo parcialmente la importancia de la Educación Infantil, ya que los beneficios asociados a la escolarización durante los años iniciales de vida exceden a los puramente educativos.

La escolarización en Educación Infantil tiene efectos positivos, tal y como ha demostrado la literatura centrada en evaluar el impacto de la asistencia (Duncan y Magnusson, 2013), la edad de escolarización (Duncan et al., 2022) y la calidad (estructural –ratios de profesorado, calidad del profesorado, duración de la jornada– así como de proceso –aproximaciones pedagógicas, tipos de interacción profesorado-alumnado–; véanse Ulferts et al., 2019; o Bowne et al., 2017), tanto en el corto como en el largo plazo (Heckman, 2012). Si bien la evidencia muestra que la Educación Infantil tiene un impacto positivo sobre el desarrollo del alumnado, los efectos son heterogéneos en función del perfil del alumno y el tipo de *outcome* evaluado, tanto para la ratio alumnado-profesorado (Duncan et al., 2022), como la edad óptima de inicio de la escolarización. La literatura ofrece resultados mixtos, explicada parcialmente por el tipo de cuidado alternativo, como los cuidados en el hogar, que suelen constituir el contrafactual (Lubotsky y Kaestner, 2016; Li et al., 2020), y la duración de la jornada –lógicamente hay diferencias importantes entre la escolarización durante algunas horas diarias, a media jornada, o a jornada completa– (Felfe y Zierow, 2018).

Entre sus efectos a corto plazo, la Educación Infantil tiene el potencial para incrementar tanto las competencias cognitivas (Blau, 2021) como las no cognitivas o socioemocionales (Joo et al., 2020), que facilitan los proce-

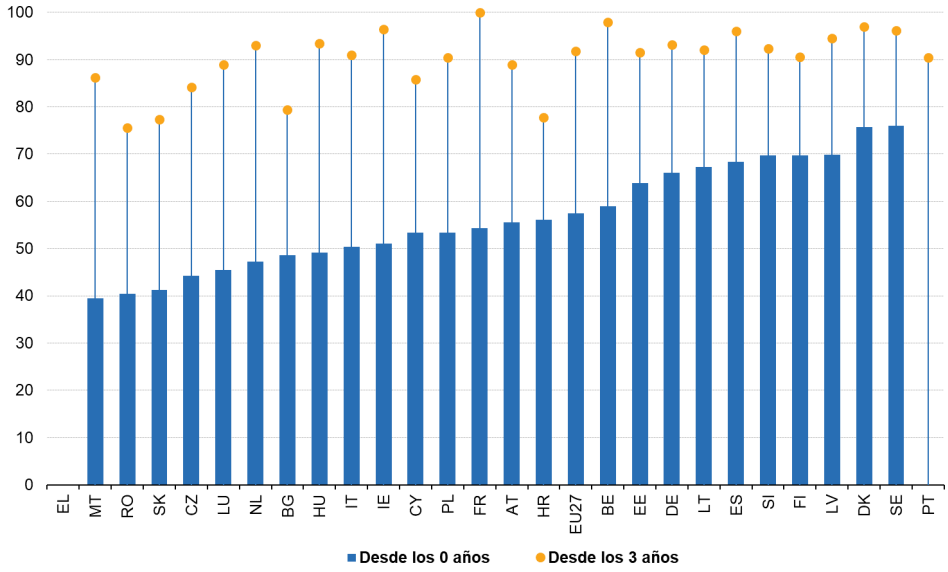
sos educativos posteriores. Por otro lado, estos beneficios no se distribuyen de forma homogénea entre la población, siendo mayores para el alumnado proveniente de hogares con un nivel socioeconómico bajo (Van Huizen y Plantenga, 2018).

Bien sea de forma directa o indirecta, en el largo plazo la participación en Educación Infantil reduce las probabilidades de repetición de curso e incrementa las tasas de matriculación y graduación en niveles educativos ulteriores (McCoy et al., 2017). Nuevamente, estos efectos positivos de la Educación Infantil no se distribuyen de forma homogénea entre la población, convirtiéndose en un instrumento de política educativa con la capacidad de alterar la igualdad de oportunidades educativas (Heckman, 2011). Ello resulta todavía más evidente cuando se consideran los beneficios adicionales vinculados a la Educación Infantil, entre los que se encuentran, entre otros, mayores tasas de ocupación y mejores salarios (De Haan y Leuven, 2020), menores tasas de embarazo juvenil (Barr y Gibbs, 2022), y una reducción en la probabilidad de comisión de delitos en edad adulta (García et al., 2019). Por tanto, los beneficios sociales de la Educación Infantil exceden los individuales, abriendo la puerta a la posible intervención pública (es un bien preferente con externalidades positivas al consumo) en dicho ámbito. De hecho, los análisis de Cascio (2023) y García et al. (2023) muestran que la inversión en Educación Infantil es coste-efectiva.

A pesar de esta creciente evidencia acerca de la rentabilidad individual y social de la inversión en Educación Infantil, su evolución a nivel internacional ha sido muy desigual, incluso entre países relativamente similares en su nivel de desarrollo económico (gráfico 1), debido a la confluencia de factores de oferta y de demanda. Por el lado de la oferta, resulta crucial el grado de intervención del sector público en la regulación y provisión del servicio de Educación Infantil, dado el elevado coste de su provisión –al exigir ratios profesor-alumno más altas que para los niveles educativos posteriores–. Por el lado de la demanda, resultan relevantes factores como el nivel de participación de la mujer al mercado laboral, los distintos modelos de familia, las diferencias culturales y la dispersión geográfica de la población. La existencia de políticas de conciliación laboral y familiar también afecta a los anteriores factores y a las preferencias familiares.

Como se puede observar en el gráfico 1, existen importantes diferencias entre países en sus grados de escolarización total y por ciclos de Educación Infantil. A

**Gráfico 1. Tasa de participación en Educación Infantil de la población desde 0 años y desde 3 años hasta la edad de escolarización en Educación Primaria (en %), año 2021.**




Fuente: Elaboración propia a partir de Eurostat. Datos para Grecia no disponibles. Datos para Portugal (desde los 0 años) no disponibles.

su vez, existen desigualdades en términos de escolarización en Educación Infantil dentro de cada país en función de variables geográficas y socioeconómicas, entre otras. Dados los efectos de la Educación Infantil sobre los resultados posteriores en distintos ámbitos vitales, las desigualdades en el acceso a la Educación Infantil de calidad pueden traducirse en un factor potenciador de la desigualdad de oportunidades. Así pues, la inversión en Educación Infantil puede analizarse –y no de forma excluyente– tanto desde el punto de vista de la efectividad y eficiencia, como del de la equidad.

En este capítulo se reflexiona acerca de diversos elementos a considerar para abordar el análisis de los impactos de la Educación Infantil. No pretende realizar una revisión de la literatura –en ese sentido, resultan excelentes, entre otros, los trabajos de Van Huizen y Plantenga (2018), Blau (2021), Cascio (2021) y Duncan et al. (2023)–, sino ser un punto de partida que facilite al lector profundizar en el análisis y la evaluación, especialmente de impacto, de la Educación Infantil. Para ello, en primer lugar (apartado 2) se procede a una breve revisión de la evolución reciente de la escolarización infantil en España. Seguidamente

(apartado 3), se presenta el marco teórico de la producción de Educación Infantil y la decisión de las familias al respecto, subyacente a la mayor parte de evaluaciones de impacto. En el apartado 4 se describen las distintas aproximaciones empíricas de la evaluación de impacto de la Educación Infantil, diferenciándose las experimentales de las cuasiexperimentales, y destacando algunos elementos fundamentales a tener en cuenta a la hora de realizar este tipo de evaluaciones. Finalmente, el apartado 5 presenta las principales conclusiones del capítulo.

## 2. LA EDUCACIÓN INFANTIL EN ESPAÑA

Tal y como se vio en el capítulo 1, la Educación Infantil en España, como en la mayoría de los países europeos, se ordena en dos ciclos; el primero, desde los 0 hasta los 3 años y, el segundo, desde los 3 a los 6. Si bien ambos ciclos son de carácter voluntario y no obligatorio, el segundo ciclo de Educación Infantil comenzó a ser gratuito a partir de la Ley Orgánica General del Sistema Educativo (LOGSE) de 1990. El segundo ciclo se imparte en los mismos centros que la Educación Primaria y se estableció un horario escolar, currículo educativo y regulación (tamaño del aula, formación docente, etc.) común con el fin de garantizar un nivel mínimo de calidad en los diferentes centros. Este cambio legislativo llevó a un incremento paulatino de las tasas de escolarización a los 3 años, pasando del 38% en el curso 1991-1992, al 84% en el curso 1999-2000 y a ser casi universal actualmente (97% en el curso 2021-2022). Como se desprende de la tabla 1, la evolución según Comunidad Autónoma (CC.AA. en adelante) ha sido muy distinta.

Por el contrario, el primer ciclo de Educación Infantil no es gratuito y no está regulado por el sector público. Tal y como sucede con el resto de etapas educativas, las competencias en materia de Educación Infantil corresponden, desde la LOGSE, a las CC.AA. Por consiguiente, el diseño de las políticas del primer ciclo de Educación Infantil depende de los gobiernos regionales. Ello incluye el sistema de financiación, el cual depende de los gobiernos regionales y locales. En términos generales, las CC.AA. financian una red de centros públicos con precios y becas para facilitar el acceso a las familias con menos recursos. Sin embargo, la oferta pública es aún insuficiente para cubrir toda la demanda. Por su parte, los diferentes ayuntamientos proporcionan cheques o bonos guardería a las familias con menores recursos para facilitar el acceso a los centros privados. Como consecuencia, la Educación Infantil de este ciclo continúa dependiendo

**Tabla 1. Evolución de las tasas netas de escolaridad a los 3 años por Comunidad Autónoma.**

| | 2021-2022 | 2014-2015 | 2009-2010 | 2004-2005 | 1999-2000 | 1996-1997 | 1991-1992 |
|---|---|---|---|---|---|---|---|
| España | 96,6 | 95,4 | 95,7 | 94,5 | 84,3 | 67,3 | 38.4 |
| Andalucía | 98,1 | 96,9 | 97,3 | 91,8 | 54,3 | 24,7 | 13,1 |
| Aragón | 96,2 | 95,6 | 95,4 | 97,1 | 95,7 | 88,2 | 59,1 |
| Asturias | 97,3 | 96,2 | 98,5 | 93,7 | 94,8 | 85,1 | 37,2 |
| Baleares | 94,4 | 94,4 | 92,6 | 94,9 | 89,5 | 71,0 | 35,4 |
| Canarias | 93,4 | 94,4 | 95,6 | 92,4 | 88,5 | 63,7 | 11,3 |
| Cantabria | 97,1 | 96,3 | 94,5 | 95,4 | 92,7 | 85,3 | 51,3 |
| Castilla y León | 97,5 | 96,9 | 97,5 | 99,0 | 97,7 | 94,8 | 58,2 |
| Castilla-La Mancha | 96,3 | 96,1 | 97,3 | 97,8 | 91,8 | 75,7 | 35,5 |
| Cataluña | 95,4 | 94,2 | 95,8 | 96,5 | 99,5 | 100,0 | 63,3 |
| Valencia | 96,3 | 92,9 | 91,0 | 94,0 | 82,4 | 47,1 | 32,8 |
| Extremadura | 99,5 | 97,9 | 98,6 | 94,2 | 91,4 | 76,1 | 37,8 |
| Galicia | 96,5 | 97,1 | 97,0 | 94,2 | 87,3 | 82,5 | 46,0 |
| Madrid | 96,4 | 94,0 | 94,5 | 92,6 | 91,3 | 75,6 | 34,7 |
| Murcia | 98,0 | 96,5 | 96,6 | 95,7 | 87,6 | 66,7 | 20,7 |
| Navarra | 96,6 | 95,7 | 95,8 | 96,8 | 96,5 | 97,1 | 94,2 |
| País Vasco | 97,7 | 98,8 | 98,0 | 98,2 | 100,0 | 100,0 | 97,2 |
| La Rioja | 96,1 | 97,1 | 94,2 | 95,0 | 97,9 | 94,0 | 36,5 |
| Ceuta | 95,8 | 94,7 | 97,3 | 84,9 | 74,9 | 50,8 | 13,5 |
| Melilla | 82,7 | 84,2 | 90,7 | 93,6 | 76,4 | 34,6 | 21,0 |

Fuente: Estadística de las Enseñanzas no universitarias. Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional.

en gran medida de los centros privados (46%), con variaciones importantes según la CC.AA. (tabla 2). En cuanto a las tasas netas de escolarización del primer ciclo de Educación Infantil, tal y como muestra la tabla 2, son aún bajas, sobre todo para niños menores de 2 años. Además, las tasas netas de escolarización en el primer ciclo en España están muy asociadas al nivel educativo de las madres. Mientras que la tasa neta de escolarización de Educación Infantil de primer ciclo en España es del 60% para niños con madres con educación universitaria, esta proporción se reduce al 20% para madres con Educación Primaria o inferior (MEFP, 2023).

**Tabla 2. Tasas netas de escolaridad en 0, 1 y 2 años por Comunidad Autónoma, según titularidad del centro. Curso 2020-2021.**

| | Tasa neta de escolaridad | | | | % Alumnado de Primer Ciclo | |
|---|---|---|---|---|---|---|
| | 0-2 años | 0 años | 1 año | 2 años | Centros Públicos | Centros Privados |
| España | 35,8 | 10,9 | 37,7 | 56,2 | 53,8 | 46,2 |
| Andalucía | 43,1 | 13,2 | 46,7 | 66,6 | 36,1 | 63,9 |
| Aragón | 31,0 | 8,5 | 32,5 | 49,4 | 64,0 | 36,0 |
| Asturias | 23,6 | 10,1 | 25,2 | 33,2 | 91,1 | 8,9 |
| Islas Baleares | 24,3 | 6,8 | 25,5 | 38,6 | 67,6 | 32,4 |
| Islas Canarias | 20,7 | 7,2 | 21,2 | 31,7 | 45,7 | 54,3 |
| Cantabria | 29,9 | 2,8 | 8,8 | 73,1 | 81,4 | 18,6 |
| Castilla y León | 18,5 | 4,6 | 20,0 | 29,3 | 72,4 | 27,6 |
| Castilla-La Mancha | 26,0 | 6,7 | 26,9 | 42,1 | 67,2 | 32,8 |
| Cataluña | 34,0 | 8,4 | 37,6 | 53,4 | 66,0 | 34,0 |
| Ceuta | 19,0 | 3,0 | 14,5 | 37,3 | 61,0 | 39,0 |
| Extremadura | 26,9 | 7,2 | 27,9 | 44,4 | 88,4 | 11,6 |
| Galicia | 42,9 | 17,1 | 46,6 | 62,5 | 60,5 | 39,5 |
| La Rioja | 41,2 | 10,6 | 47,0 | 64,5 | 46,9 | 53,1 |
| Madrid | 45,8 | 18,2 | 49,9 | 66,6 | 53,1 | 46,9 |
| Melilla | 24,6 | 6,4 | 19,2 | 43,8 | 60,7 | 39,3 |
| Murcia | 17,7 | 4,5 | 19,2 | 28,5 | 51,2 | 48,8 |
| Navarra | 26,2 | 6,5 | 29,0 | 41,1 | 82,6 | 17,4 |
| País Vasco | 50,5 | 14,8 | 42,8 | 89,7 | 53,8 | 46,2 |
| Valencia | 34,9 | 7,8 | 34,4 | 59,6 | 47,3 | 52,7 |

Fuente: MEFP (2023).

A pesar de que las tasas de escolarización de primer ciclo de Educación Infantil son aún bajas, su crecimiento en las últimas décadas ha sido notable. La tabla 3 recoge la evolución de la tasa de escolarización de 0 a 2 años desde inicios de la década de los 90 hasta la actualidad: del 3% del curso 1991-1992 se pasó al 42%, en 2021-2022. Nuevamente, las trayectorias según CC.AA. son muy heterogéneas. Ello viene explicado por una serie de factores entre los que sobresalen la diversidad de políticas públicas llevadas a cabo por las administraciones

regionales y locales, las diferencias socioeconómicas regionales, así como de la disparidad cultural y de preferencias de las familias.

**Tabla 3. Evolución de las tasas netas de escolaridad de 0 a 2 años por Comunidad Autónoma.**

| | 2021-2022 | 2016-2017 | 2011-2012 | 2006-2007 | 2001-2002 | 1996-1997 | 1991-1992 |
|---|---|---|---|---|---|---|---|
| España | 41,6 | 37,5 | 30,7 | 17,8 | 10,6 | 6,2 | 3,3 |
| Andalucía | 50,8 | 43,3 | 33,7 | 3,9 | 1,3 | 0,7 | 0,7 |
| Aragón | 35,8 | 35,7 | 31,7 | 28,8 | 4,3 | 1,8 | 2,6 |
| Asturias | 27,0 | 21,1 | 16,9 | 8,5 | 1,8 | 1,2 | 0,8 |
| Baleares | 28,3 | 25,4 | 21,5 | 10,9 | 6,6 | 4,1 | 1,7 |
| Canarias | 24,9 | 20,6 | 7,1 | ,, | 1,1 | 1,5 | 1,0 |
| Cantabria | 32,6 | 27,9 | 24,0 | 15,3 | 3,4 | 0,0 | 0,5 |
| Castilla y León | 23,8 | 21,7 | 16,9 | 11,9 | 8,2 | 1,7 | 1,6 |
| Castilla-La Mancha | 34,1 | 33,7 | 33,9 | 2,4 | 1,9 | 1,6 | 1,0 |
| Cataluña | 39,6 | 38,0 | 36,4 | 30,2 | 26,9 | 24,5 | 9,9 |
| Valencia | 39,6 | 33,3 | 25,3 | 11,9 | 6,8 | 3,1 | 7,7 |
| Extremadura | 34,6 | 28,6 | 3,3 | 2,9 | 1,3 | 1,0 | 0,7 |
| Galicia | 49,9 | 45,3 | 22,8 | 16,6 | 12,9 | 0,0 | 2,8 |
| Madrid | 51,7 | 48,2 | 43,3 | 32,9 | 16,3 | 6,7 | 2,6 |
| Murcia | 21,7 | 18,4 | 15,6 | 15,7 | 8,6 | 6,0 | 1,4 |
| Navarra | 31,5 | 27,9 | 10,5 | 25,4 | 13,2 | 0,0 | 0,0 |
| País Vasco | 54,2 | 52,1 | 51,9 | 46,6 | 23,1 | 14,0 | 2,1 |
| La Rioja | 50,1 | 36,7 | 15,7 | 4,5 | 2,5 | 2,5 | 1,5 |
| Ceuta | 20,3 | 12,6 | 3,8 | 4,1 | 3,4 | 1,2 | 0,9 |
| Melilla | 31,4 | 21,3 | 17,8 | 16,6 | 8,5 | 1,1 | 0,1 |

Fuente: Estadística de las Enseñanzas no universitarias. Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional.

Finalmente, un aspecto relevante a la hora de analizar el impacto de la Educación Infantil de primer ciclo es la presencia de una alta heterogeneidad en la calidad de los servicios prestados, como consecuencia de la falta de regulación. Inicialmente, las guarderías surgen en España principalmente para dar respuesta a la necesidad de las familias de disponer de un servicio de cuidado de los niños,

como consecuencia de la masiva incorporación de la mujer en el mercado laboral y la falta de políticas de conciliación familiar y laboral. Con el tiempo, esta etapa ha ido adquiriendo cada vez un carácter más educativo y menos asistencial, si bien aún conviven ambos modelos. La falta de regulación del servicio en el primer ciclo implica que cada centro tiene completa autonomía. Por ello es importante, a la hora de evaluar los efectos de la asistencia a la Educación Infantil, incorporar no solo la cantidad de años de asistencia (o la edad de ingreso), sino también controlar por las características del centro al que asisten (ratios, formación del personal, procesos pedagógicos, etc.) y, en definitiva, la calidad del servicio. En cualquier caso, medir la calidad de los centros supone un reto para la evaluación de la Educación Infantil. Lamentablemente, la mayoría de las bases de datos educativas disponibles para España (internacionales y nacionales) no incluyen esta información, lo cual resulta una limitante a la hora de interpretar el impacto de los servicios de Educación Infantil de primer ciclo sobre los resultados futuros de los niños.

## 3. MARCO CONCEPTUAL

El objetivo de esta sección es proporcionar un marco teórico que permita comprender los fundamentos subyacentes a los análisis empíricos que miden el impacto de la Educación Infantil sobre las habilidades y el bienestar de los niños. Para ello, primero se describe la función de producción educativa (FPE) y se detallan sus principales características e implicaciones para el nivel de Educación Infantil. Posteriormente, se presenta un modelo económico para explicar los determinantes de la decisión de las familias en materia de Educación Infantil, y que permite comprender los mecanismos a través de los cuales pueden operar algunas de las políticas públicas y programas comúnmente evaluados en la literatura. En este capítulo introducimos un modelo simple, pero en el capítulo 3b se discuten otras extensiones sobre el efecto de los modelos parentales y la transmisión de valores "culturales" en educación. Por último, se discuten brevemente los principales desafíos empíricos asociados a la estimación de los impactos de las políticas de Educación Infantil, teniendo en cuenta las limitaciones de información y, en particular, la dificultad de identificar efectos causales en la práctica.

### 3.1. La función de producción educativa

La función de producción educativa hunde sus raíces en la teoría del capital humano. Según esta teoría, el nivel de capital humano es el resultado de las in-

versiones realizadas por los individuos, dependiendo éstas de sus preferencias y de los beneficios esperados de la misma (Schultz, 1961, 1962; Becker, 1962). En este marco, la inversión en educación sería el principal canal de adquisición de capital humano. En términos generales, la FPE establece la relación productiva entre los recursos educativos (familiares y escolares) y los resultados educativos (Hanushek, 1979; Todd y Wolpin, 2003). El principal antecedente de la FPE –o, al menos, el más influyente– es el conocido "Informe Coleman" (1966), el cual a partir de una encuesta realizada a más de medio millón de niños pertenecientes a más de 3.000 escuelas en Estados Unidos, estudió la asociación estadística entre los resultados educativos y las características socioeconómicas familiares, introduciendo la discusión sobre el rol de los recursos escolares en el resultado académico. Los principales resultados del informe indicaban que las diferencias entre los centros escolares tenían poco que ver con los logros del alumnado, y que el determinante más importante eran los antecedentes familiares (véase el capítulo 3b donde se presentan diferentes modelos que podrían explicar los canales por los que operan dichos antecedentes) y las características del resto del alumnado del centro (véase el capítulo 3c para una discusión detallada del efecto de los compañeros o *peer effects*). Estas conclusiones y sus implicaciones de política educativa despertaron gran controversia y debates, así como un notable interés por el estudio de la relación *inputs-outputs* en el ámbito de la economía de la educación (Hanushek, 1986; 2003). Ello dio lugar al desarrollo de diferentes modelos económicos basados en la FPE para la estimación de relaciones causales entre los recursos y los resultados educativos (ver Todd y Wolpin, 2003, para una discusión detallada de los principales aspectos teóricos y empíricos de la estimación de la FPE).

La FPE aplicada al contexto de la Educación Infantil está estrechamente relacionada con la extensa e influyente investigación desarrollada en las dos últimas décadas por James Heckman y sus colaboradores, focalizada en el estudio de la formación de habilidades cognitivas y no cognitivas, y combinando evidencia multidisciplinar proveniente de la Economía, la Neurociencia y la Psicología[1]. En sus trabajos seminales, Cunha y Heckman (2007, 2008) proponen un modelo económico para explicar la "tecnología de formación de las habilidades" (TFH) cognitivas y no cognitivas para múltiples etapas de la infancia, con el

[1] Para una revisión de los trabajos llevados a cabo en los últimos 20 años por esta línea de investigación véase https://heckmanequation.org/.

objetivo de sistematizar la evidencia disponible en materia de desarrollo infantil. Posteriormente, se han llevado a cabo diferentes extensiones para proponer especificaciones alternativas o imponer supuestos de identificación en el modelo económico con el objetivo de abordar los desafíos más relevantes de la estimación de la TFH (ver, por ejemplo, los trabajos de Cunha y Heckman, 2008; Cunha et al., 2010; Attanasio, 2015; Attanasio et al., 2022a).

Si bien en términos conceptuales ambas aproximaciones (la FPE y la TFH) comparten el mismo objetivo (estudiar el proceso de formación de las habilidades o resultados educativos de los niños), existen algunos matices en cuanto al enfoque y supuestos de cada una. La TFH está principalmente centrada en el estudio del rol familiar y en cómo el entorno, las características y las inversiones familiares afectan al desarrollo de las habilidades de los niños. En este caso, el foco del análisis está en las decisiones familiares, y los recursos escolares –y la inversión pública– están implícitos en el modelo, incluidos en el componente de 'inversión parental'. Es decir, las familias deciden invertir en uno u otro modelo de escolarización (por ejemplo, enviar a sus hijos a centros privados o públicos), en base a sus preferencias, su restricción presupuestaria y la información disponible, lo cual, bajo el supuesto de información perfecta, determinaría el nivel de recursos escolares (volveremos sobre este punto más adelante). Por su parte, la FPE está centrada en la relación productiva entre los recursos y los resultados educativos (*input-output*), donde los recursos escolares tienen un rol explícito en el modelo y no dependen exclusivamente de las decisiones familiares, sino también de las decisiones de los propios centros y/o de las administraciones educativas (ver la sección 6 del capítulo 3b para una discusión detallada sobre como la interacción de las decisiones de múltiples actores puede afectar al desempeño de los estudiantes y a la asignación de recursos y esfuerzos en la educación). Dado que este capítulo se centra en la medición del impacto de la Educación Infantil sobre las habilidades de los niños, se ha optado por utilizar la FPE como marco teórico principal, incorporando algunas de las características de la tecnología de formación de habilidades propuesta por Cunha y Heckman (2007).

En términos generales, la FPE de la Educación Infantil se puede expresar de la siguiente forma:

$$A_{it} = f_t\,(X_i, F_i, I_{i,1\ldots t}, S_{i,1\ldots t}) \qquad (1)$$

donde $A_{it}$ representa el resultado educativo de interés del niño al final del periodo $t$, $X_i$, recoge las características individuales (incluida la habilidad innata); el vec-

tor $F_i$ representa las características familiares exógenas; el vector $I_{i,1...t}$ captura las inversiones familiares en recursos educativos y tiempo acumuladas hasta el periodo $t$; y, finalmente, el componente $S_{i,1...t}$ corresponde a los recursos del sistema educativo acumulados hasta el periodo $t$. La función $f_t$ indica que los parámetros de la FPE pueden variar a lo largo del tiempo, es decir, el impacto de los recursos educativos sobre las habilidades de los niños puede cambiar según su edad.

Tal y como se desprende de la ecuación 1, el proceso de formación de habilidades es acumulativo. Así, las habilidades producidas en una determinada etapa se acumulan a las obtenidas en etapas posteriores, es decir, $(\partial f_t (X, F_{1...t}, S_{1...t}) / \partial A_{t-1} > 0$. Este efecto es lo que Heckman y Cunha (2007) denominan "autoproductividad" (*self productivity*). La "autoproductividad" refleja la idea de que las habilidades se refuerzan a sí mismas y se potencian mutuamente. Al estudiarse múltiples habilidades cognitivas y no cognitivas, esta característica incluye efectos propios y cruzados. Por ejemplo, la adquisición de habilidades no cognitivas como la autodisciplina en edades tempranas permite alcanzar mayores niveles de habilidades cognitivas durante las etapas escolares posteriores (Duckworth y Seligman, 2005). Otra de las características de la formación de habilidades es lo que en la literatura se conoce como la "complementariedad dinámica" (*dynamic complementary*), que se refiere a que el nivel de habilidades adquiridas en el periodo, $t-1$, $A_{it-1}$ (como consecuencia de las inversiones en $t-1$), hace que las inversiones en el periodo $t$ sean más productivas en la generación de $A_{it}$ (ver por ejemplo Heckman y Cunha, 2007; Cunha et al., 2010). En términos de la FPE de la ecuación 1, la complementariedad dinámica de las inversiones familiares, $F$, implicaría que $(\partial^2 f_t (X, F_{1...t}, S_{1...t}) / \partial A_{t-1} \partial F'_t > 0$. Es decir, las inversiones en recursos educativos en edades tempranas tienen efectos no solo en el periodo en que se realizan, sino también en edades posteriores. Asimismo, la complementariedad dinámica implica también que, para que las inversiones en edades tempranas sean productivas en periodos posteriores, éstas deben ser seguidas de inversiones en edades más tardías. Ambas características –autoproductividad y complementariedad dinámica– producen efectos multiplicadores que son los mecanismos a través de los cuales se produce el efecto "*skills beget skills*" (las habilidades engendran habilidades), popularizado por Heckman y colaboradores. Estas características son particularmente relevantes para comprender los mecanismos a través de los cuales la Educación Infantil impacta en el nivel de desarrollo de habilidades de los niños y, sobre todo, para interpretar los efectos de medio y largo plazo. Resulta a su vez de gran utilidad para entender por qué las

inversiones en edades tempranas en Educación Infantil son altamente rentables para los alumnos más desaventajados económicamente y, por tanto, para promover la igualdad de oportunidades.

Ejercicio 1. *Asumiendo la siguiente FPE*

$$lnA_{it} = \beta_1 \, lnA_{it-1} + \beta_2 \, lnF_i + \beta_3 \, lnI_{it} + \beta_4 \, lnS_{it} + \varepsilon_{it}$$

*derivar cuál sería el efecto de una política de reducción del tamaño del aula en Educación Infantil en el período t–1, sobre el resultado educativo en t–1 y en t. ¿Cómo se relaciona el resultado obtenido con los conceptos de autoproductividad y complementariedad dinámica vistos anteriormente?*

En la ecuación 1 se asume que las características individuales $X_i$ y las características familiares $F_i$ son componentes exógenos (en éstas últimas se encuentran las preferencias y creencias de las familias, el nivel educativo de los padres, etc.), mientras que las inversiones familiares en recursos educativos $I_{i1...t}$ son endógenas. En cuanto a las inversiones escolares $S_{i1...t}$, una parte viene determinada por la decisión de las familias (por ejemplo, la edad de inicio de la escolarización, las horas de asistencia al centro o el tipo de centro educativo), y por tanto es endógena, y otra parte es exógena, ya que depende de decisiones de los propios centros o de las autoridades educativas –y que las familias desconocen al tomar sus decisiones de inversión–.

Es importante destacar que las familias pueden reaccionar (tanto en términos de recursos educativos en el hogar $I_{i1...t}$ como de escolarización de los hijos $S_{i1...t}$) a la evolución del resto de los componentes de la FPE. Por lo tanto, la modelización del proceso de toma de decisiones familiares permite comprender los mecanismos a través de los cuales los programas o las políticas públicas de Educación Infantil pueden impactar sobre el desarrollo de las habilidades. Por ejemplo, si las familias anticipan una política pública o programa (como puede ser la introducción de un subsidio en el precio de los servicios de Educación Infantil) y modifican su nivel de inversión en recursos educativos en el hogar, el efecto de la política puede verse afectado por dicho cambio (compensando o potenciando el efecto directo). A continuación, se presenta un modelo económico que permite comprender qué factores intervienen en la decisión de las familias sobre cuánto invertir en educación temprana para sus hijos. Tal y como mencionamos anteriormente, en este capítulo presentamos un modelo simple adaptado al caso de la Educación Infantil. El lector interesado puede consultar el capítulo

3b para modelos más específicos sobre el efecto de los modelos parentales y la transmisión de valores culturales en educación.

### 3.2. Determinantes de la inversión de las familias en Educación Infantil

Una de las primeras decisiones a las que deben enfrentarse las familias en relación con la educación de sus hijos es cuánto invertir en ella desde edades muy tempranas. En concreto, las familias deben decidir a qué edad comenzar la escolarización, cuántas horas diarias y, sobre todo, a qué centro educativo enviar a sus hijos, considerando varios factores como la posibilidad de conciliar la vida laboral y familiar (restricción temporal), qué tan efectiva creen que es la Educación Infantil para el desarrollo de sus hijos, el coste de los servicios y la restricción presupuestaria familiar, entre otros factores objetivos y emocionales. Además, las familias deben decidir cuánto tiempo y dinero invertir en sus hijos dentro del hogar, también considerando los factores previamente mencionados. Por otra parte, las autoridades educativas deciden, también condicionadas por sus propias restricciones presupuestarias, acerca de la edad a partir de la cual la Educación Infantil debe ser gratuita y universal, además de regular los contenidos y la formación de los docentes necesarios para impartir este nivel, entre otras cuestiones. Ello también puede afectar a las decisiones familiares. La disposición de un marco conceptual de referencia permite comprender los mecanismos a través de los cuales operan los efectos de los programas y las políticas públicas (facilitando, idealmente, su diseño adecuadamente). Además, la modelización del comportamiento de las familias puede ser útil también como estrategia de identificación de los efectos causales de los recursos escolares, como veremos más adelante.

A continuación, siguiendo a Attanasio (2015), se presenta un modelo donde la decisión de las familias sobre la inversión en recursos educativos está basada en la optimización de la utilidad familiar sujeta a una serie de restricciones. Por simplicidad, llamaremos $E_{i1...t}$ al conjunto de inversiones familiares en recursos educativos, que incluye las inversiones en recursos educativos (materiales y tiempo) en el hogar $I_{i1...t}$, y la inversión en la escolarización de los hijos (años de educación, tiempo de asistencia, tipo de centro, etc.), recogida en parte en el componente $S_{i1...t}$ de la ecuación 1. Nótese que, en el caso de asumir información perfecta por parte de las familias respecto de la calidad y características de los centros educativos, así como de la importancia de la Educación Infantil, el com-

ponente $S_{i1...t}$ dependería exclusivamente de la decisión familiar y, por tanto, sería totalmente endógeno y podría incluirse dentro del componente , $I_{i1...t}$ (ver Cunha y Heckman, 2007). En la realidad, las familias disponen de cierto grado de información acerca de las características de los centros educativos (algunas más que otras), pero ésta dista de ser perfecta. El modelo que se presenta a continuación es la versión más simple de la literatura, ya que el objetivo del apartado es discutir los factores más relevantes que pueden afectar a las decisiones familiares. Para especificaciones más complejas y la relajación de diversos supuestos pueden consultarse los trabajos de Cuhna et al. (2010) o Attanasio et al. (2020).

En términos generales, el problema de maximización al que se enfrentan las familias a la hora de decidir cuántos recursos educativos invertir en edades tempranas puede expresarse de la siguiente forma:

$$\max_{\{C_t, E_{i,t}\}} U(C_t, A_{i,t}) \tag{2}$$

$$s.a. \quad C_t + P_t^E E_{i,t} = Y_t \tag{3}$$

$$A_{i,t} = f_t(X_i, F_i, E_{i,t}) \tag{4}$$

donde $C_t$ corresponde al consumo familiar, $P_t^E$ refleja el vector de precios de las inversiones educativas $E_{i,t}$ (cuotas escolares, materiales escolares, etc.) e $Y_t$ son los recursos económicos disponibles de la familia. Por tanto, las decisiones de inversión dependerán de las preferencias y gustos familiares ($U$), los parámetros de la FPE de la ecuación 1 ($f_t$), de los precios de las inversiones, $P_t$ y de la restricción presupuestaria de las familias $Y_t$. La función de inversión educativa puede expresarse de la siguiente forma:

$$E_{i,t} = z_t(X_i, F_i, P_t, Y_t, f_t, U) \tag{5}$$

En otras palabras, los hogares maximizan su utilidad en base a sus preferencias y gustos, a su restricción presupuestaria, y al nivel de información que poseen sobre los parámetros de la FPE.

Ejercicio 2. *Asumiendo la misma FPE que en el ejercicio 1 y teniendo en cuenta la siguiente función de inversión educativa* $E_{i,t} = \delta_1 F_i + \delta_2 P_t + \delta_3 Y_t + \delta_4 S_{it}$, *derivar el efecto de la misma política de reducción del tamaño del aula implementada en el periodo t sobre el resultado educativo en t, si las familias reaccionaran con una reducción del nivel de inversión* $E_{i,t}$. *¿Y si lo aumentaran?*

Es interesante destacar, que en el problema de maximización (2) se asume que no existe ahorro. Por tanto, las preferencias y gustos familiares dependen de la utilidad que les reporta el consumo y el nivel de habilidades que desarrollen sus hijos. Factores como el nivel de altruismo o paternalismo explican que algunas familias valoren en mayor medida el desarrollo de habilidades y, por lo tanto, inviertan más en recursos educativos para sus hijos. Asimismo, algunas familias darán mayor importancia al desarrollo de cierto tipo de habilidades respecto de otras y, dependiendo de cómo repercutan las inversiones en cada tipo de habilidades, invertirán más o menos. Por tanto, los parámetros de la FPE son un factor determinante de las inversiones familiares.

Ejercicio 3. *Suponga que las familias desconocen la FPE y consideran que el tamaño del aula no afecta al resultado educativo de su hijo , ¿cómo cambiarían los resultados del ejercicio 2?*

El modelo que se presenta asume que solo existe un hijo en la familia pero, de haber varios, los padres deberán decidir cuánto invertir en cada uno teniendo en cuenta sus características (edad, género, habilidad percibida, etc.). Las preferencias respecto de la asignación de inversiones dentro del hogar pueden ser muy heterogéneas. Por ejemplo, algunas familias pueden optar por compensar las diferencias iniciales en habilidades o, por el contrario, apostar por invertir en mayor medida en quienes tienen mayores habilidades (Almond, y Mazumder, 2013). La literatura sobre la asignación de inversiones educativas dentro del hogar es extensa y excede el alcance de este capítulo, pero se puede consultar Attanasio et al. (2022b) para una discusión más amplia sobre esta cuestión. Por último, los cambios en la sociedad y las preferencias sociales también pueden influir sobre los gustos y preferencias familiares, con lo cual hay que tenerlos en cuenta a la hora de interpretar los resultados en los análisis empíricos.

Otro factor determinante para las familias a la hora de decidir el nivel óptimo de inversión en Educación Infantil es la restricción presupuestaria, que viene determinada por el vector de precios de las inversiones (dado para las familias) y los recursos económicos familiares. A mayor precio de las inversiones o menor nivel de recursos, menores inversiones educativas. En este sentido, políticas focalizadas a familias de baja renta, que reduzcan los precios de las inversiones (gratuidad de los servicios de Educación Infantil, bonos de guardería, etc.) suelen ser mecanismos efectivos para incrementar la asistencia a la Educación Infantil en esta franja de población.

Por último, otro elemento sumamente relevante en la toma de decisiones de inversiones educativas es el conocimiento y creencias que los padres tienen respecto al proceso de formación de habilidades de sus hijos y de la importancia que tienen las inversiones en recursos educativos. Es decir, de los parámetros de la FPE de la ecuación 1. Si bien inicialmente los modelos económicos estándar asumían que las familias eran racionales y tenían información perfecta respecto de la FPE, la evidencia de las últimas dos décadas ha puesto de manifiesto la importancia de flexibilizar este supuesto para considerar el papel de las imperfecciones de información en la determinación de las inversiones educativas. La falta de información de las familias, tanto en términos de la calidad de los servicios de Educación Infantil disponibles, como del impacto que tienen los recursos educativos en el desarrollo de las habilidades de sus hijos, influye sobre las preferencias parentales y, por tanto, sobre el nivel de inversiones que realizan (Cunha et al., 2020; Attanasio et al., 2022a). Esto implica que, a pesar de que las familias maximizan su función de utilidad (ecuación 2), las decisiones que tomen en términos de inversión educativa pueden ser sub-óptimas, ya que son tomadas en base a lo que los padres perciben sobre el proceso educativo y el nivel de desarrollo de sus hijos.

Las restricciones de conocimiento y de información son particularmente relevantes en las familias de rentas bajas, quienes tienden a subestimar los retornos de las inversiones educativas (Cunha et al., 2013; Attanasio et al., 2019). Este hecho, sumado a las creencias sobre la crianza de los hijos, puede ser un factor explicativo clave de los déficits en inversión en Educación Infantil en familias vulnerables. En estos contextos, políticas públicas destinadas a proporcionar más y mejor información a los padres pueden ser mecanismos efectivos para promover una mayor inversión en la Educación Infantil y así impulsar el desarrollo de las habilidades de los niños.

Ejercicio 4. *Discutir los pros y contras de los siguientes mecanismos empleados por los hogares para la obtención de información acerca de la importancia y calidad de la Educación Infantil:*

1) *De boca en boca entre familiares y amistades.*
2) *Rankings publicados en diversos medios de comunicación.*
3) *Campañas públicas de información en publicidad y medios.*
4) *Evaluaciones públicas de valor añadido de los centros de Educación Infantil.*

### 3.3. Retos de la estimación empírica del impacto de la Educación Infantil

La realización de evaluaciones de impacto para conocer la efectividad de la Educación Infantil implica tener que afrontar una serie de retos. Entre ellos, destacan: 1) los problemas de endogeneidad y autoselección; 2) la dificultad en la definición y medición del resultado de interés (*outcome*); y 3) la falta de datos suficientes para acometer las evaluaciones.

Empecemos por la endogeneidad y la autoselección. En Educación Infantil, a la asignación no aleatoria del alumnado entre los distintos centros –con características y calidades diversas, y un alumnado heterogéneo–, se le añade la complejidad, propia de los niveles educativos no obligatorios, de que el alumnado escolarizado tiene unas características personales y familiares distintas a quienes deciden no participar. Así, a modo de ejemplo, pongamos que estuviéramos interesados en comparar directamente el resultado medio entre participantes y no participantes en un programa de Educación Infantil. Si los participantes en educación infantil provinieran principalmente de entornos socioeconómicos favorables, la estimación estaría sesgada al alza, incluso en el caso de que pudiéramos controlar por una serie de variables socioeconómicas en nuestro análisis. En otras palabras: al alumnado participante le hubiera ido mejor que al no participante, incluso si no hubiera participado en ese programa de educación infantil ya que tiene unos padres con unas características más favorables que el grupo no participante. Lo contrario sucedería si, el programa en cuestión hubiera ido dirigido, principalmente, a alumnado con circunstancias personales y/o familiares desfavorables.

Vale la pena destacar aquí también que resulta crítica la elección del grupo contrafactual, ya que al evaluar la efectividad de la educación infantil puede compararse la participación respecto a la no participación (quedarse en casa con algún familiar o cuidador), o bien la participación en un programa alternativo respecto a quienes participan en el programa de educación infantil a evaluar. En este caso, la participación por parte del grupo de control en algún tipo de educación infantil alternativo al programa evaluado puede minorar el efecto de dicho programa o ser una de las explicaciones al hallazgo de efectos decrecientes en el tiempo. Naturalmente, esta dinámica dependerá de la calidad de las alternativas al programa evaluado y de las características del alumnado participante –en el programa evaluado, y en el contrafactual–.

En base a lo anterior, no es de extrañar que el estudio de las relaciones causales a nivel medio poblacional (*average treatment effects* – ATE) resulte especial-

mente difícil para este nivel educativo, siendo más habitual el análisis de los efectos de la participación sobre subgrupos poblacionales concretos o de políticas específicas sobre los participantes. Se trata frecuentemente, por consiguiente, de evaluaciones de impacto con elevados niveles de validez interna, pero de validez externa limitada. En cualquier caso, la obtención de dicha validez interna pasa por la corrección de los sesgos asociados a la endogeneidad y a la autoselección muestral. Dicha corrección puede realizarse, tal y como se verá en el subapartado 3.2, mediante técnicas experimentales y cuasiexperimentales.

Un segundo reto al que se enfrentan las evaluaciones de impacto en Educación Infantil es la dificultad en la definición y medición del *outcome* o resultado de interés. Como hemos mencionado anteriormente, se trata de un nivel educativo en el que los contenidos y objetivos del currículo se encuentran definidos de forma menos estructurada que en los niveles obligatorios (sobre todo en el primer ciclo de Educación Infantil), y en el que la adquisición de competencias no cognitivas, como la interacción personal, el autocontrol o la capacidad de atención resulta central. Incluso en el caso de estar definidas, la medición de estas competencias resulta más complicada que, por ejemplo, los resultados en matemáticas o lengua, cuantificadas habitualmente a partir del nivel de Primaria a través de pruebas estandarizadas, ya sea a nivel nacional o internacional. Así pues, la evaluación en el corto y medio plazo del impacto de los resultados de la Educación Infantil hace oportunos los abordajes multidisciplinares que incorporen elementos psicométricos –como los tests Peabody o SDQ, procedentes de la psicología– para la medición de las competencias no cognitivas para la cuantificación precisa de los *outcomes*.

Ahora bien, buena parte de la literatura ha centrado su interés en el impacto de la Educación Infantil sobre otro tipo de resultados en la edad adulta como la trayectoria académica, laboral o criminal del individuo. En tal caso, resulta imprescindible el seguimiento de los participantes en Educación Infantil a lo largo de décadas y, a menudo, la vinculación de la información de participación en Educación Infantil con otros registros administrativos.

Ello enlaza con el tercero de los retos planteados para la realización de evaluaciones de impacto en Educación Infantil: la escasez de información relativa a la trayectoria de escolarización y resultados de los participantes. Esta escasez, unida a la necesidad de disponer de perspectiva temporal para el análisis de *outcomes* a lo largo del ciclo vital explican el menor número relativo de evaluaciones de

impacto para Educación Infantil y la utilización de información de programas implementados con años o incluso décadas de antelación, con las limitaciones que ello implica en países con una cultura de la evaluación menos desarrollada como es el caso de España.

## 4. MEDICIÓN DEL IMPACTO DE LA EDUCACIÓN INFANTIL

Para solventar los retos anteriormente mencionados, las estrategias seguidas para la realización de evaluaciones de impacto en Educación Infantil pueden dividirse entre las aproximaciones experimentales y las cuasiexperimentales.

### 4.1. Evaluaciones experimentales

Los estudios experimentales permiten la estimación precisa del impacto de la Educación Infantil al asignar de forma aleatoria la participación en el programa en cuestión. Esta aleatorización al producirse de forma intencional, controlando quien participa y quien no en el programa, se denominan experimentos aleatorios controlados o RCT, por sus siglas en inglés. De esta forma, los RCT garantizan la comparabilidad entre las características observables y no observables de los miembros del grupo tratado y del grupo de control (ver el capítulo 2 para mayor detalle sobre RCTs). Por tanto, si la aleatorización se ha realizado de forma adecuada, las diferencias entre ambos subgrupos pueden atribuirse a la participación en el programa. Es de destacar que, en algunas situaciones, los niños no reciben el tratamiento que se les asignó, y existe lo que se conoce en la literatura como un problema de *non-compliance.* Por ejemplo, algunos niños asignados al programa de Educación Infantil pueden no participar finalmente porque sus familias así lo deciden, lo que se conoce como *Never-Takers*. Por el contrario, algunos niños del grupo de control pueden finalmente participar en el programa, los que se conocen como *Allways-Takers*. En este caso, la comparación entre el grupo de tratamiento y de control no permitiría estimar el ATE para la población, sino que identifica el efecto medio de la intención de tratar (*intention-to-treat*; ITT). Bajo este escenario, es necesario corregir las estimaciones para estimar el *Local Average Treatment Effect* (LATE) –véase el capítulo 2, para una discusión al respecto– únicamente sobre aquellos niños que efectivamente cumplen con la asignación original al programa (*compliers*).

A pesar de tratarse de evaluaciones con una elevada validez interna, el grueso de la evidencia basada en RCT se ha concentrado principalmente en el análisis

de una serie de programas implementados a nivel local en Estados Unidos y destinados a subgrupos poblacionales con bajo nivel socioeconómico. Los ejemplos más famosos son el Perry Preschool Program (1962) de Michigan, y el Carolina Abecedarian Project (1972), de Carolina del Norte, en los que se evaluaba el impacto de la participación en Educación Infantil para el alumnado desaventajado; y el Proyecto STAR (Student-Teacher Achievement Ratio, 1985) de Tennessee, centrado en analizar el impacto del tamaño del aula[2]. Elango (2016) proporciona un buen resumen de los aprendizajes derivados de los dos primeros experimentos, mientras que Chetty et al. (2011) y Dynarski et al. (2013) son dos ejemplos destacados de las evaluaciones de impacto realizadas sobre el Proyecto STAR.

El impacto del Tennessee Voluntary Prekindergarten Program (TN-VPK), un programa destinado a alumnado de cuatro años de edad y proveniente de hogares con un nivel socioeconómico bajo, con alguna discapacidad o que no hablan la lengua inglesa, también ha sido evaluado a través de RCT, aprovechando el exceso de demanda del programa (Lipsey et al., 2013). Los resultados ofrecen una visión menos optimista acerca de la participación en Educación Infantil que la obtenida a partir de la evaluación de otros programas, ya que los efectos positivos –sobre competencias cognitivas y no cognitivas– que se identifican al año de haber participado en TN-VPK, se diluyen al cabo de pocos cursos.

Por último, otro programa ampliamente evaluado en la literatura ha sido el programa Head Start Impact Study de EE. UU., la evaluación aleatorizada del programa federal Head Start, que proporciona de forma gratuita, entre otros servicios, Educación Infantil a niños provenientes de hogares con bajos niveles de ingresos. Para ello, se aleatorizaron las ofertas para participar en el programa en aquellos lugares en los que existía un exceso de demanda y se siguió a este alumnado hasta tercer año de Primaria. Lipsey et al. (2015a) y Kline y Walters (2016) constituyen ejemplos ilustrativos de la evaluación de la efectividad de Head Start sobre una serie de resultados educativos y sanitarios.

A nivel europeo, destaca el empleo de RCT para la evaluación de programas de Educación Infantil en Reino Unido y Dinamarca. En el primer caso, el Hackney Day Care Study evaluó los efectos de participar en Educación Infantil pública de

[2] El Proyecto STAR se analiza detenidamente en el capítulo 3a.

alta calidad sobre el comportamiento y competencias no cognitivas del alumnado de bajo nivel socioeconómico, de entre 6 meses y tres años, de dicho barrio de Londres (Toroyan et al., 2004). En el caso danés, Jensen et al. (2013) analizaron el impacto de Action Competences in Social Pedagogical Work with Socially Endangered Children and Youth (ASP), un programa consistente en mejorar la formación del profesorado de los centros de Educación Infantil, sobre las competencias cognitivas y los niveles de bienestar de alumnado, también de bajo nivel socioeconómico, de entre 3 y 6 años de edad de dos municipios daneses.

En resumen, a pesar de su *a priori* mayor validez interna, el número de RCT en Educación Infantil resulta reducido, dado su elevado coste –que condiciona el tamaño muestral–, los posibles problemas de desgaste de la muestra (abandono del programa o del seguimiento de algunos participantes), especialmente agudos cuando se trata de analizar resultados a largo plazo, y la ya comentada falta de validez externa.

En el caso de España, no se dispone hasta la fecha de evidencia experimental en Educación Infantil. Con una filosofía similar, Santín y Sicilia (2018), aprovechando el experimento natural generado por aquellos centros de Educación Primaria con más de un grupo por curso, explotaron las diferencias entre clases donde el alumnado fue asignado de forma aleatoria para evaluar el impacto de la participación en Educación Infantil en España. Sin embargo, sí existen, situaciones análogas a la descrita, por ejemplo, en el Head Start Impact Study. Así, en las grandes ciudades españolas suele darse un exceso de demanda –o, más bien, defecto de oferta– de plazas públicas de Educación Infantil. La realización de este tipo de asignación experimental para el caso español requeriría que, de forma previa, las administraciones públicas reconocieran el acceso gratuito o subvencionado a la Educación Infantil como un derecho a partir de la superación de un determinado umbral y posteriormente, se realizara una lotería para asignar las plazas ofertadas dentro de los demandantes que no superan el umbral. En cualquier caso, valga como ejemplo para ilustrar que, con imaginación y la complicidad de las administraciones, la realización de un experimento social resultaría factible en España en el contexto actual.

Ejercicio 5. *Proponer el método más adecuado para la evaluación del impacto de las siguientes políticas públicas. Pensar en el impacto de interés y el tipo de información necesaria para la realización de la evaluación, y discutir a su vez el alcance y limitaciones de los resultados.*

*1) Implementación de un programa consistente en la exoneración del pago de la matrícula en los centros de Educación Infantil a partir de un umbral de renta familiar.*

*2) Reducción de la ratio de alumnado por docente en los centros de una Comunidad Autónoma, disponiendo los centros de un período de diez años para la adaptación a dicha normativa.*

*3) Extensión de la obligatoriedad de la Educación Infantil desde los dos años de edad en toda España a partir del 1 de enero de dentro de cinco años.*

### 4.2. Evaluaciones cuasiexperimentales

El hecho de que la participación no se asigne habitualmente de forma aleatoria dificulta la obtención de evidencia causal sobre los efectos de la Educación Infantil. No obstante, el empleo de técnicas cuasiexperimentales como diferencias en diferencias (DD), modelos de efectos fijos, variables instrumentales (VI) y regresiones en discontinuidad (RD) han permitido proporcionar, bajo las condiciones que se discutirán a continuación, dicha evidencia. La aplicación de estas técnicas parte de la premisa de disponer de bases de microdatos en diversos momentos temporales que permitan vincular la participación en Educación Infantil con momentos vitales posteriores.

Los modelos de DD explotan diferencias temporales y/o geográficas en la implementación de programas y han sido utilizados ampliamente para la evaluación de la Educación Infantil. Una característica que hace popular el uso de este tipo de aproximación es que no requiere disponer de bases de datos longitudinales –relativamente más escasas– sino que puede aplicarse empleando microdatos que siguen a una misma cohorte a lo largo del tiempo. La idea principal consiste en construir grupos tratados y de control, para diversos años, en función de los distintos ritmos de implementación de los programas de Educación Infantil. Por tanto, se explotan dos diferencias: el cambio en el tiempo o entre cohortes en las zonas (ciudades, distritos, regiones, etc.) tratadas, y el cambio en el tiempo o entre cohortes en las zonas inicialmente no tratadas. Así, por ejemplo, Felfe et al. (2014) y Bosque-Mercader (2022) aprovecharon las disparidades en la velocidad de adaptación de los sistemas educativos de las Comunidades Autónomas en España a la Ley de Ordenación General del Sistema Educativo (LOGSE), que establecía la gratuidad de la Educación Infantil entre los 3 y 5 años, para evaluar respectivamente los efectos de la participación en Educación Infantil sobre

la repetición de curso y competencias cognitivas a los 15 años de edad, y sobre la probabilidad de desarrollar enfermedades en edad adulta. Callaway y Sant'Anna (2021) y Goodman-Bacon (2021) desarrollan soluciones metodológicas para los desafíos que plantea la estimación de DD con programas de implementación gradual.

Tal y como sucede en otras áreas de conocimiento, los efectos heterogéneos de la Educación Infantil suelen evaluarse a través de análisis de diferencias en diferencias en diferencias (DDD) y los esfuerzos para convencer acerca de la validez interna de los modelos de DD se centran en demostrar las tendencias paralelas previas a la implementación del programa. Una cuestión relevante por considerar es la posible endogeneidad de los distintos ritmos de implementación del programa, proponiéndose en tal caso la realización de tests de sensibilidad utilizando información del año previo a la implementación del programa (Cascio, 2021). Debe también considerarse la posibilidad de que las familias modifiquen su conducta para conseguir que su hijo consiga acceder al programa (mediante movilidad geográfica o mediante la insistencia a los centros para conseguir la admisión). En tales casos, el empleo de tests de sensibilidad excluyendo del análisis a los años cercanos al momento de la implementación pueden resultar un complemento interesante a los clásicos tests placebo en los que se simula la implementación del programa en años o territorios distintos a los verdaderos –con la esperanza de obtener resultados estadísticamente no significativos–.

Los modelos de efectos fijos explotan la diversidad de tratamiento (asistencia a Educación Infantil) dentro de una unidad de agrupamiento. Típicamente, estos modelos toman el hogar como unidad de agrupamiento y comparan los resultados entre hermanos. Pages et al. (2020) y Berlinski et al. (2008) son ejemplos representativos. La clave para la estimación de estos modelos consiste en disponer de bases de datos longitudinales –razón que explica la concentración de este tipo de análisis para, principalmente, los EE. UU.– y, naturalmente, la identificación de hogares en los que solo algunos de los hijos hayan sido afectados por una reforma de Educación Infantil sobre la que los hogares no tengan capacidad de decisión. Bajo este último supuesto, los resultados obtenidos tendrán una elevada validez interna, si bien restringida a los hogares que tienen más de un hijo.

Una parte de los estudios que han evaluado el impacto de la Educación Infantil, especialmente en el ámbito europeo y asiático, ha utilizado el método de

VI. Para ello, parten de un modelo principal basado en la FPE detallada en la sección 2, donde un determinado resultado depende de una batería de variables entre las que se encuentra la participación en un programa de Educación Infantil. Estas aproximaciones plantean superar los problemas de endogeneidad y autoselección asociados a la participación en dichos programas mediante el cálculo de un modelo previo (primera etapa) en el que la variable endógena –la participación en el programa de Educación Infantil– depende de una serie de variables de control –las empleadas en el modelo de la segunda etapa– y al menos una variable –instrumento– exógena que a) debe estar fuertemente correlacionada con la variable endógena (validez del instrumento); y b) solo debe afectar a la variable dependiente del modelo de la segunda etapa –el *outcome* de interés– a través de su incidencia en la variable endógena (exogeneidad del instrumento) –el capítulo 2 proporciona una descripción más detallada de las técnicas de VI–. En Educación Infantil, las fuentes de variación exógena que se han utilizado con mayor frecuencia como instrumentos son la fecha de nacimiento del niño y la variación temporal a nivel local en la cobertura de la Educación Infantil –Nguyen (2022) y Zachrisson et al. (2023)–, los cuales constituyen buenos ejemplos, respectivamente.

Los resultados obtenidos a partir de VI en Educación Infantil se enfrentan a los mismos retos que en otros ámbitos. Así, debe considerarse que los estimadores presentados miden LATE, por lo que solo estarán midiendo el efecto de aquellas personas que vieron efectivamente modificada su participación como consecuencia del instrumento seleccionado. Por otro lado, la validez y exogeneidad de los instrumentos debe ser probada. En este sentido, resulta pertinente comprobar que las familias no han modificado su conducta. Para los dos instrumentos mencionados se trataría de garantizar, por un lado, que los centros de Educación Infantil no pudieran saltarse la normativa relativa a la edad para la escolarización y, por otro lado, que resulta creíble el supuesto de no movilidad geográfica.

Finalmente, los modelos de RD se aplican frecuentemente para la evaluación de programas cuyo diseño establece la superación de algún umbral para la admisión del participante. En el caso de la Educación Infantil, dos de los criterios utilizados con mayor frecuencia para determinar la admisión son la edad del alumno (Berlinski et al., 2011) y el nivel familiar de ingresos (Carneiro y Ginja, 2014). Las evaluaciones que emplean RD proporcionan evidencia muy similar a la que se obtiene mediante aproximaciones experimentales, partiendo

del supuesto de que las personas que quedan cerca del umbral en cuestión son, en términos medios, idénticas en todas sus características (observables y no observables), difiriendo solamente en haber sido admitidas o no en el programa de Educación Infantil. Cabe realizar, no obstante, varias advertencias. En primer lugar, los efectos que se estimarán no serán los ATE, sino LATE. Es decir: los efectos estimados no serán extrapolables al conjunto de la población, sino solamente al de las personas que quieren participar en Educación Infantil y que se encuentran cercanas al umbral. En segundo lugar, tal y como exponen Lipsey et al. (2015), el empleo de la edad como variable de decisión resulta especialmente delicada. A modo de ejemplo, teniendo en cuenta la reducida edad de los niños, pocos meses de diferencia pueden dar lugar a diferencias importantes en los niveles de maduración, por lo que los resultados pueden ser especialmente sensibles, cuando se emplea la edad como discontinuidad, al intervalo de días o meses que se seleccione. Por otro lado, el contrafactual, al emplearse la edad como variable de decisión (*running variable*), serán todas las experiencias y actividades del niño durante el año previo a que resulten elegibles (superen el umbral) para dicho programa. Finalmente, la literatura muestra la existencia de fuertes efectos heterogéneos de la Educación Infantil (Lee et al., 2022), por lo que el empleo de RD para su estimación, al centrarse en una parte muy concreta de la población (la muy cercana al umbral), puede no ser la más aconsejable. En definitiva, tal y como sucedía con los RCT, se trata de un tipo de evaluaciones con una elevada validez interna, pero con una reducida validez externa.

## 5. CONCLUSIONES

En las últimas décadas la Educación Infantil ha experimentado una notable expansión, abarcando cada vez edades más tempranas. Esta evolución viene explicada por múltiples factores como la incorporación de la mujer al mercado laboral, cambios en los modelos familiares y el creciente reconocimiento de los efectos positivos de la escolarización temprana. La evidencia ha demostrado la existencia de altos beneficios, tanto a nivel individual como social, de intervenciones tempranas como la escolarización en Educación Infantil a jornada completa, la mejora de la formación del profesorado o la universalización del acceso a partir de los 3 años. Además, estos retornos son mayores para los niños provenientes de contextos más desfavorecidos. Aunque buena parte de esta evidencia proviene de la implementación de programas focalizados en las familias con menores recursos económicos en EE. UU., también se ha demostrado que

las políticas de carácter universal, más propias de países europeos, son efectivas –al generar, por ejemplo, mayores efectos compañero y homogeneizar el tipo de educación recibida por todo el alumnado al acceder al primer curso de Primaria–. Sin embargo, la evidencia disponible no es concluyente respecto a otras cuestiones también relevantes en el diseño de las políticas de Educación Infantil como la reducción del tamaño del aula, la edad de inicio de la escolarización temprana (antes de los 3 años) o los determinantes de la persistencia en el medio plazo de las habilidades adquiridas durante la Educación Infantil.

Por otro lado, la validez externa de los resultados obtenidos a través de estrategias como RCT, VI o RD resulta limitada, al centrarse en la estimación de LATE o, en el caso de los RCT, en subgrupos poblacionales muy específicos. La evidencia ha dejado clara la existencia de una elevada heterogeneidad en los efectos de las políticas en función de los diversos perfiles poblacionales. Estos estudios resultan útiles para identificar programas que funcionan para colectivos concretos, pero no tanto para identificar los motivos por los que dichos programas funcionan ni para elaborar políticas a nivel sistémico. Por ello, una línea interesante a explorar es la combinación de evaluaciones de impacto con la estimación de los parámetros de los modelos comentados en la sección 3 de este capítulo. Por ejemplo, aprovechando la implementación de experimentos sociales para estimar los parámetros de la FPE o de la función de inversión familiar en Educación Infantil. Este tipo de ejercicio permitiría además de estimar los impactos de las intervenciones, identificar los mecanismos a través de los cuales operan.

Adicionalmente, debe señalarse que, en el nivel infantil, resulta especialmente importante el acompañamiento de las evaluaciones de impacto con estimaciones de su coste. Ello es así al tratarse de un nivel educativo para el que existen alternativas fuera de la escuela (permanecer en el hogar, por ejemplo) y donde se pueden aplicar intervenciones alternativas, como los programas consistentes en el fortalecimiento del vínculo familia-niño que pueden ser igualmente efectivas y menos costosas. Así mismo, es importante tener en cuenta que, a edades tempranas, el simple hecho de conversar con el niño, leerle un libro o involucrarlo en simples juegos puede tener grandes efectos en el desarrollo cognitivo del niño, y en ocasiones son intervenciones más coste-efectivas que invertir más recursos en la Educación Infantil.

Finalmente, quedan abiertos múltiples frentes para avanzar en el conocimiento de los efectos de la Educación Infantil en España. El impacto de la edu-

cación infantil sobre el desarrollo de competencias distintas a las medidas por evaluaciones internacionales y de diagnóstico, la identificación de los efectos más allá de la adolescencia o la medición de la efectividad de distintas prácticas educativas son, entre otras, cuestiones todavía por dilucidar. El avance queda condicionado por la dificultad para el acceso a bases de datos administrativas, la falta de seguimiento de la trayectoria de los niños participantes en el largo plazo y la posibilidad de evaluar programas piloto o implementar RCTs. Tampoco facilita la labor el hecho de que se trate de un nivel educativo cuantitativa y cualitativamente muy heterogéneo a nivel territorial si bien, tal y como se ha expuesto, precisamente esta heterogeneidad ha sido explotada para la identificación de la –escasa– evidencia disponible para España. Adicionalmente, el lector interesado en investigar la Educación Infantil en España deberá tomar en consideración también la falta de un currículo común que establezca unos contenidos u objetivos mínimos a alcanzar, así como la escasez de evaluaciones educativas en los niveles educativos tempranos; lo cual hace complejo el conocimiento del proceso de formación de las habilidades y la posible complementariedad entre disciplinas en dicho proceso. En definitiva, la mayor parte de las dificultades señaladas están vinculadas, en mayor o menor grado, a la falta de un "sistema" nacional de Educación Infantil.

## BIBLIOGRAFÍA


Almond, D. y Mazumder, B. (2013). Fetal origins and parental responses. *Annu. Rev. Econ.*, *5*(1), 37-56.

Attanasio, O. P. (2015). The determinants of human capital formation during the early years of life: Theory, measurement, and policies. *Journal of the European Economic Association*, *13*(6), 949-997.

Attanasio, O., Cunha, F. y Jervis, P. (2019). Subjective parental beliefs. their measurement and role (No. w26516). *National Bureau of Economic Research*.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C. y Rubio-Codina, M. (2020). Estimating the production function for human capital: results from a randomized controlled trial in Colombia. *American Economic Review*, *110*(1), 48-85.

Attanasio, O., Boneva, T. y Rauh, C. (2022a). Parental beliefs about returns to different types of investments in school children. *Journal of Human Resources*, *57*(6), 1789-1825.

Attanasio, O., Cattan, S. y Meghir, C. (2022b). Early childhood development, human capital, and poverty. *Annual Review of Economics*, *14*, 853-892.

Barr, A. y Gibbs, C. R. (2022). Breaking the cycle? intergenerational effects of an antipoverty program in early childhood. *Journal of Political Economy*, *130*(12), 3253-3285.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, *70*(5, Part 2), 9-49.

Berlinski, Samuel, Sebastian Galiani y Marco Manacorda (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics, 92*, 1416-1440. https://doi.org/10.1016/j.jpubeco.2007.10.00

Berlinski, S., Galiani, S. y Mc Ewan, P. J. (2011). Preschool and maternal labor market outcomes: Evidence from a regression discontinuity design. *Economic Development and Cultural Change*, *59*(2), 313-344.

Blau, D. M. (2021). The effects of universal preschool on child and adult outcomes: A review of recent evidence from Europe with implications for the United States. *Early Childhood Research Quarterly*, *55*, 52-63.

Bosque-Mercader, L. (2022). *The effect of a universal preschool programme on long-term health outcomes: Evidence from Spain* (No. 2022-07). FEDEA.

Bowne, J. B., Magnuson, K. A., Schindler, H. S., Duncan, G. J. y Yoshikawa, H. (2017). A meta-analysis of class sizes and ratios in early childhood education programs: Are thresholds of quality associated with greater impacts on cognitive, achievement, and socioemotional outcomes? *Educational Evaluation and Policy Analysis*, *39*(3), 407-428.

Callaway, B. y Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200-230.

Carneiro, P. y Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, *6*(4), 135-173.

Cascio, E. U. (2021). *Early childhood education in the United States: What, when, where, who, how, and why* (No. w28722). National Bureau of Economic Research.

Cascio, E. U. (2023). Does universal preschool hit the target? Program access and preschool impacts. *Journal of Human Resources*, *58*(1), 1-42.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. y Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, *126*(4), 1593-1660.

Coleman, J. S. (1966). *Equality of educational opportunity* [summary report (Vol. 1). US Department of Health, Education and Welfare, Office of Education.

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

Cunha, F., Elo, I. y Culhane, J. (2013). Eliciting maternal expectations about the technology of cognitive skill formation (No. w19144). *National Bureau of Economic Research*.

De Haan, M. y Leuven, E. (2020). Head start and the distribution of long-term education and labor market outcomes. *Journal of Labor Economics*, *38*(3), 727-765.

Duncan, G. J. y Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, *27*(2), 109-132.

Duncan, G., Kalil, A., Mogstad, M. y Rege, M. (2023). Investing in Early Childhood Development in Preschool and at Home, en Machin, S.J., Woessmann, L. Hanushek, E.A. (Eds.), *Handbook of the Economics of Education.* Volume 7. North Holland, Amsterdam (en prensa).

Dynarski, S., Hyman, J. y Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management*, *32*(4), 692-717.

Duckworth, A. L. y Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939-944.

Elango, Sneha, Jorge Luis García, James J. Heckman y Andres Hojman (2016). Early Childhood Education, en *Economics of Means-Testing Transfer Programs in the United States,* Volume 2 (ed. Robert Moffitt), pp. 235-297.

Felfe, C., Nollenberger, N. y Rodríguez-Planas, N. (2015). Can't buy mommy's love? Universal childcare and children's long-term cognitive development. *Journal of Population Economics*, *28*, 393-422.

Felfe, C. y Zierow, L. (2018). From dawn till dusk: Implications of full-day care for children's development. *Labour Economics*, *55*, 259-281

García, J. L., Heckman, J. J. y Ziff, A. L. (2019). Early childhood education and crime. *Infant Mental Health Journal*, *40*(1), 141-151.

García, J. L., J. J. Heckman, D. E. Leaf y M. J. Prados (2020). Quantifying the Life-Cycle Benefits of a Prototypical Early Childhood Program. *Journal of Political Economy*, *128*(7), 2502-2541.

García, J. L., Heckman, J. J. y Ronda, V. (2023). The Lasting Effects of Early-Childhood Education on Promoting the Skills and Social Mobility of Disadvantaged African Americans and Their Children. *Journal of Political Economy*, *131*(6).

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254-277.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, *14*(3), 351-388.

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

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

Heckman, J. J. (2011). The economics of inequality: The value of early childhood education. *American Educator*, *35*(1), 31.

Heckman, J. J. (2012). Invest in early childhood development: Reduce deficits, strengthen the economy. *The Heckman Equation*, *7*(1-2).

Heckman, J. J. y Mosso, S. (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, *6*(1), 689-733.

Hidalgo-Hidalgo, M. y García-Pérez, J. I. (2013). Capítulo 4: Impacto de la asistencia a Educación Infantil sobre los resultados académicos del estudiante en Primaria, en *PIRLS-TIMS 2011. Estudio Internacional de progreso en comprensión lectora, matemáticas y ciencias*. IEA. Volumen II, 83.

Jensen, B., Holm, A. y Bremberg, S. (2013). Effectiveness of a Danish early year preschool program: A randomized trial. *International Journal of Educational Research*, *62*, 115-128.

Joo, Y. S., Magnuson, K., Duncan, G. J., Schindler, H. S., Yoshikawa, H. y Ziol-Guest, K. M. (2020). What works in early childhood education programs?: A meta–analysis of preschool enhancement programs. *Early Education and Development*, *31*(1), 1-26.

Kline, P. y Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of Head Start. *The Quarterly Journal of Economics*, *131*(4), 1795-1848.

Lee, S. Y., Rodgers, J., Kim, R. y Subramanian, S. V. (2022). Distributional effects on children's cognitive and social-emotional outcomes in the Head Start Impact Study: A quantile regression approach. *SSM-Population Health*, *18*, 101108.

Li, W., Duncan, G. J., Magnuson, K., Schindler, H. S., Yoshikawa, H. y Leak, J. (2020). Timing in Early Childhood Education: How Cognitive and Achievement Program Impacts Vary by Starting Age, Program Duration, and Time since the End of the Program. *Working Paper, 20,* 201. Annenberg Institute for School Reform at Brown University.

Lipsey, M. W., Farran, D. C. y Hofer, K. G. (2015a*). A Randomized Control Trial of a Statewide Voluntary Prekindergarten Program on Children's Skills and Behaviors through Third Grade.* Research Report. Peabody Research Institute.

Lipsey, M. W., Weiland, C., Yoshikawa, H., Wilson, S. J. y Hofer, K. G. (2015b). The prekindergarten age-cutoff regression-discontinuity design: Methodological issues and implications for application. *Educational Evaluation and Policy Analysis*, *37*(3), 296-313.

Lubotsky, D. y Kaestner, R. (2016). DoSkills Beget Skills'? Evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood. *Economics of Education Review*, *53*, 194-206.

MEFP (2023). *Las cifras de la educación en España. Curso 2020-2021* (edición 2023). Ministerio de Educación y Formación Profesional. Madrid.

McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., ... y Shonkoff, J. P. (2017). Impacts of early childhood education on medium-and long-term educational outcomes. *Educational Researcher*, *46*(8), 474-487.

Nguyen, C. V. (2022). The effect of preschool attendance on Children's health: Evidence from a lower middle-income country. *Health Economics*, *31*(8), 1558-1589.

ONU (2015). Transforming our world: the 2030 Agenda for Sustainable Development. *United Nations,* NY, USA.

Pages, R., Lukes, D. J., Bailey, D. H. y Duncan, G. J. (2020). Elusive longer-run impacts of head start: Replications within and across cohorts. *Educational Evaluation and Policy Analysis*, *42*(4), 471-492.

Santín, D. y Sicilia, G. (2018). Does preschool education attendance matter? Evidence from a natural experiment in Spain. *Applied Economics*, *50*(47), 5050-5063.

Schultz, T. W. (1960). Capital formation by education. *Journal of Political Economy*, *68*(6), 571-583.

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, *51*(1), 1-17.

Todd, P. E. y Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485), F3-F33.

Toroyan, T., Oakley, A., Laing, G., Roberts, I., Mugford, M. y Turner, J. (2004). The impact of day care on socially disadvantaged families: an example of the use of process evaluation within a randomized controlled trial. *Child: Care, Health and Development*, *30*(6), 691-698.

Ulferts, H., Wolf, K. M. y Anders, Y. (2019). Impact of process quality in early childhood education and care on academic outcomes: Longitudinal meta-analysis. *Child Development*, *90*(5), 1474-1489.

Van Huizen, T. y Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review*, *66*, 206-222.

Zachrisson, H. D., Dearing, E., Borgen, N. T., Sandsør, A. M. J. y Karoly, L. A. (2023). Universal early childhood education and care for toddlers and achievement outcomes in middle childhood. *Journal of Research on Educational Effectiveness*, 1-29.

# Capítulo 4b
# LA EDUCACIÓN POST-OBLIGATORIA

Marcel Jansen
Departamento de Análisis Económico: Teoría Económica e Historia Económica
Universidad Autónoma de Madrid
FEDEA
ORCID: https://orcid.org/0000-0003-4337-3790

Aitor Lacuesta
Banco de España
ORCID: https://orcid.org/0000-0002-7050-8462

**Resumen**

El capítulo trata la elección de educación post-obligatoria describiendo algunos artículos que estiman la interrelación de esta decisión con su impacto en los ingresos laborales a lo largo de la vida. Para contextualizar el capítulo se documenta como la mejora educativa en España tras la crisis financiera favoreció a los estudios universitarios y, entre los hombres, a los ciclos formativos de Grado superior. Se presenta un marco teórico y evidencia descriptiva que racionaliza este patrón a partir de la evolución de los rendimientos de la Universidad y de la formación profesional (FP) Superior. A continuación, se presentan algunos artículos que han enfrentado a la formación académica y profesional. En concreto, si bien la FP puede facilitar la entrada en el mercado laboral, la formación académica podría aportar competencias que pudieran ser más útiles en un contexto de cambio tecnológico acelerado. Finalmente, dado el auge reciente de la FP en España y dada la apuesta del Gobierno español por la FP dual, se presenta evidencia para España sobre el impacto en los primeros años de carrera laboral de haber cursado un Grado de FP tradicional o uno de FP dual.

**Nota y agradecimientos**

Las opiniones de este capítulo corresponden a los autores y no necesariamente coinciden con la visión del Banco de España o el Eurosistema. Queremos

agradecer el trabajo empírico de Ana Regil y la revisión de Ismael Sanz y Lucas Gortázar.

## 1. INTRODUCCIÓN

En la actualidad, la educación obligatoria en España dura diez cursos. Empieza el año que el niño cumple 6 y finaliza cuando cumple 16. Al acabar esta etapa, el alumno puede decidir ponerse a trabajar, seguir estudios de FP o plantearse una educación académica. Esta decisión tiene unas consecuencias muy importantes para el bienestar de la persona ya que cambiará de por vida su capacidad de alcanzar mayores conocimientos y otros títulos académicos superiores que reportarán otras corrientes de ingresos futuras. En realidad, esta decisión se va tomando día a día y de una forma secuencial desde que se cumplen los 16 años hasta que se decide poner fin a la etapa de formación, al menos por un tiempo[1].

En España, tradicionalmente, mucha gente dejaba de estudiar con anterioridad o justo en el momento de conseguir el título de Educación Secundaria obligatoria. En el caso de seguir estudiando, muchos abandonaban antes de alcanzar un título superior. Como consecuencia, España se convirtió en uno de los países con peores estadísticas de abandono escolar en el área del euro. En segundo lugar, de los que seguían en la formación post-obligatoria, la gran mayoría escogía una vía académica para posteriormente iniciar los estudios universitarios. En este sentido, la FP Media y la FP Superior han quedado tradicionalmente relegadas a ser opciones marginales y, en consecuencia, España ha presentado un déficit de perfiles técnicos y de nivel educativo medio.

Tras la crisis financiera de 2008, estos patrones tradicionales han ido cambiando lo que ha permitido a España converger, en parte, con el área del euro. La segunda sección del capítulo documenta los hechos más relevantes en el cambio reciente de patrón de las decisiones educativas tras la etapa de enseñanza obligatoria. En particular, más allá de documentar cambios en los stocks, se utiliza información de la Encuesta de Población Activa para apuntar la dirección y magnitud de los cambios en las transiciones entre niveles educativos. La tercera

---

[1] Este capítulo no trata la formación profesional continua al centrarse en la elección de los estudiantes a los 17 años. Sin embargo, este elemento es de gran importancia para la actualización de las capacidades de los trabajadores. En concreto, en la última sección se propone un ejercicio que abre la puerta a analizar con datos este tema.

sección presenta un modelo económico que ayuda a pensar en las motivaciones detrás de las decisiones educativas. Con este modelo en mente, se presenta en esta sección un marco empírico y evidencia descriptiva sobre la evolución de los rendimientos de diferentes grados educativos en España entre 2001 y 2019. Para tal fin se utilizan las secciones cruzadas de la Encuesta Financiera de las Familias.

El reciente auge de la FP hace necesario profundizar algo más en algunas de las cuestiones más relevantes relativas a este tipo de enseñanza y sus diferencias con la formación académica que ya han sido analizadas en otros países. En particular, algunos autores abogan por la ventaja de la FP en facilitar la inserción laboral de sus graduados respecto a las ramas más académicas que, por el contrario, tendrían ventaja en el largo plazo al facilitar la inserción laboral en entornos cambiantes. La evidencia empírica de esta hipótesis tiene interés en la comparativa de la FP Superior con la Universidad, pero también dentro de la propia FP, cuando se comparan sistemas con diferente grado de clases prácticas, como la FP tradicional y la dual. La cuarta sección presenta evidencia relativa a la comparativa de los ingresos a lo largo de la vida de FP Superior y Universidad utilizando la Encuesta Financiera de las Familias. Finalmente, la quinta sección presenta evidencia relativa a los primeros años de inserción laboral tras graduarse en FP tradicional o en FP dual. Para tal fin se utilizan datos de bases de cotización de la Seguridad Social de graduados en diferentes sistemas de FP. Esta última cuestión es especialmente relevante dado que la ley que organiza el sistema de FP aprobada en 2022 por el Gobierno español apuesta decididamente por la modalidad dual introducida en 2012.

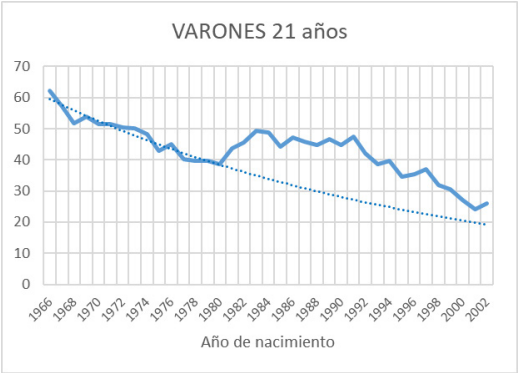
## 2. LA ELECCIÓN DE ESTUDIOS POST-OBLIGATORIOS EN ESPAÑA EN LOS ÚLTIMOS 60 AÑOS

Hace 60 años destacaba el elevado porcentaje de estudiantes que no seguían los estudios más allá de lo estrictamente obligatorio. De acuerdo con la Encuesta de Población Activa (EPA), a los 21 años, alrededor de un 62% de los hombres y un 57% de las mujeres nacidos a mediados de los años 60 no disponían de ningún título por encima de la formación estrictamente obligatoria (gráfico 1)[2].

[2] Según las diferentes legislaciones en España, la formación obligatoria permitiría obtener el título de Bachiller elemental, Educación General Básica –EGB–, Educación Secundaria Obligatoria –ESO–, iniciación profesional o Formación Profesional –FP– básica.

Estas cifras mejoraron de forma continua hasta 1995, momento en el que tocó a la generación nacida en 1978 decidir si continuar sus estudios. El abandono escolar de esa generación se situaba en el 39% de los hombres y el 26% de las mujeres. A partir de entonces, la mejora educativa se frenó durante unos años. La generación nacida en 1990 aún presentaba un 40% de hombres y un 26% de mujeres con la Educación Secundaria Obligatoria (ESO) como mayor título formativo a los 21 años. Posteriormente, a partir de 2009, se inició una nueva senda de mejora que ha continuado hasta 2023. Así, los nacidos en 2002 que no disponían de un título superior a lo estrictamente obligatorio representaban, a los 21 años, un porcentaje del 21% de los hombres y del 14% de las mujeres.

**Gráfico 1. Porcentaje de la población de 21 años de los que tienen ESO como máximo (por año de nacimiento).**



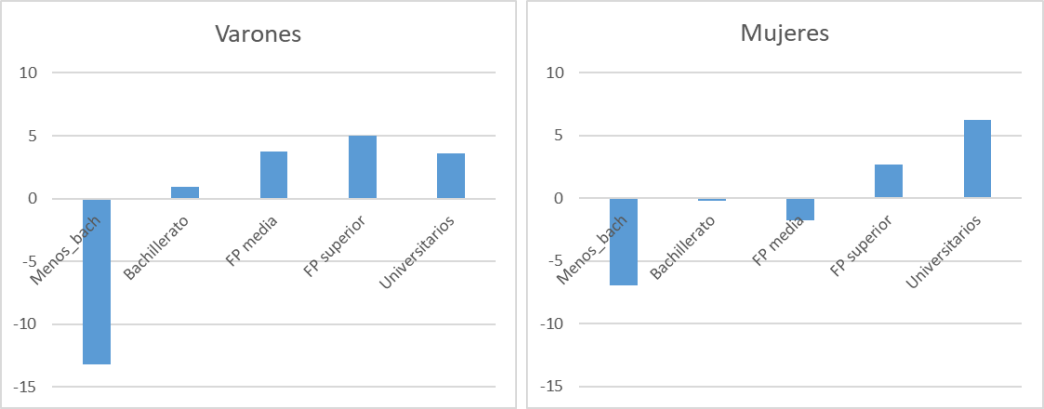
Fuente: EPA 1987-2023 (INE).
Nota: No tienen un nivel adquirido superior a los estudios obligatorios, ni ningún tipo de FP, salvo la Básica. Clas. 1

A pesar de la mejora, España sigue presentando una tasa de abandono escolar muy superior a la de otros países europeos de su entorno, principalmente entre los hombres. Para más detalles sobre la comparativa internacional de abandono escolar véase el capítulo 1.

Centrándose en las generaciones nacidas a partir de 1990 (gráfico 2), se observa que, a los 25 años, el descenso del abandono escolar supuso principalmente un aumento de la formación superior. En el caso de los hombres, el mayor incremento fue para la FP Superior. Así, la generación nacida en el 1998, a los 25 años, presentó 5pp más de FP Superior que aquellos nacidos 8 años antes. El aumento en este periodo de graduados con título universitario fue de 3,6pp,

similar al de FP Media que experimentó un aumento de 3,7pp. Es esperable que a futuro la distribución de máxima formación alcanzada incremente aún más la incidencia de FP Superior ya que, como se verá más adelante, las generaciones nacidas a partir del año 2000 han incrementado la entrada en FP Media tras la ESO y la entrada en FP Superior tras el Bachillerato. En el caso de las mujeres, el cambio de formación máxima alcanzada se concentró en educación universitaria (+6,2pp). También se observó un aumento de la FP Superior, aunque fue de menor tamaño (+2,6pp). En cualquier caso, en este colectivo, también es esperable que se incremente el porcentaje de mujeres con FP como en el caso de los hombres, aunque el cambio de transiciones es algo más leve y tardío.

**Gráfico 2. Cambio en el % de la máxima formación alcanzada a los 25 años entre nacidos 1997/1998 y nacidos 1988/1992.**



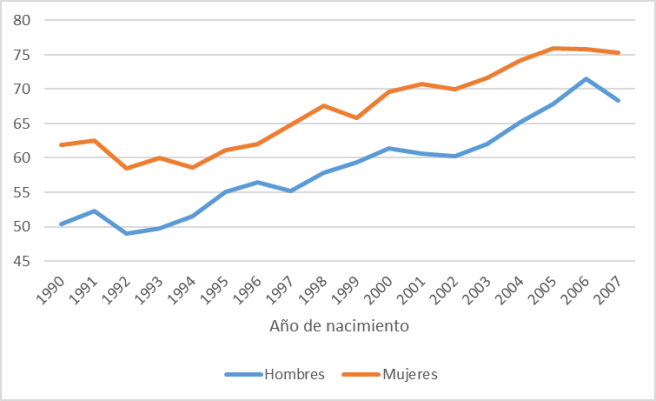

Fuente: EPA 2015-2023 (INE).

Los indicadores anteriores sobre el nivel de formación alcanzado no ayudan lo suficiente para entender cómo se han modificado las transiciones educativas en España al finalizar la etapa obligatoria en los últimos años. Para analizar estas transiciones podemos utilizar de nuevo los datos de la EPA. Esta encuesta permite describir la transición educativo-formativa de las personas entre 16 y 17 años. Concretamente, además de la edad, se dispone de las respuestas a dos preguntas: el nivel formativo más elevado alcanzado por la persona y el Grado que se está cursando, en el caso de seguir estudiando, durante las cuatro semanas anteriores a la encuesta. Con esta información se puede observar tanto el éxito de conseguir un Grado a una determinada edad como las transiciones educativas tras la obtención de un Grado.

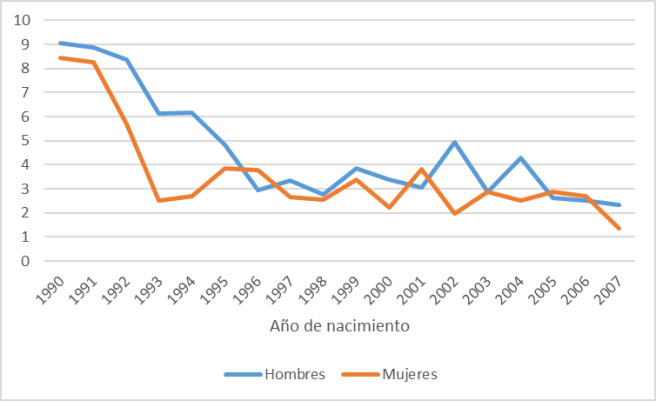
Analizando, el porcentaje de estudiantes que poseía el título de la ESO a los 16 y 17 años (gráfico 3.1.), se observa un claro aumento de 18pp para hombres y 13pp para mujeres entre las generaciones de 1990 y las nacidas a principios de siglo XXI. Este aumento considerable de graduados no derivó en un incremento del porcentaje de estudiantes que decidió ponerse a trabajar tras la obtención del título, algo en parte normal desde 2008 dadas las escasas oportunidades laborales para todo el mundo y especialmente para las personas con poca formación a lo largo de la crisis financiera. Al contrario, el porcentaje de estudiantes que no siguió estudiando a pesar de graduarse en la ESO se redujo considerablemente en ese mismo periodo desde el 9% y el 8% hasta el 2% y 1% para hombres y mujeres (gráfico 3.2.). Este nuevo patrón de continuidad formativa post-obligatoria, que se ha mantenido incluso en los años de recuperación tras 2013, ha sido clave para la reducción del abandono escolar.

**Gráfico 3. Transiciones educativas desde la ESO (por año de nacimiento).**

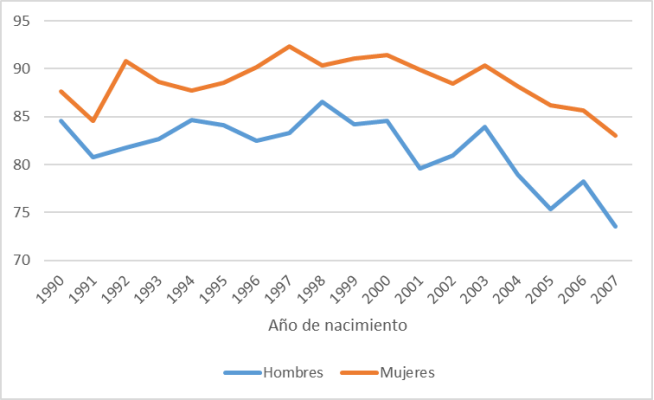
**3.1. Graduados ESO con 16-17 años (en % total).**



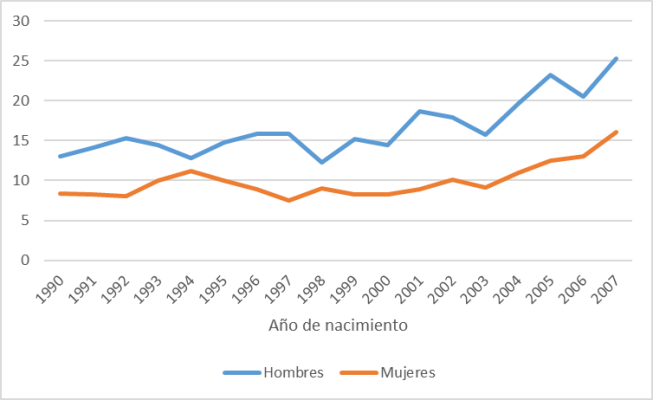

**3.2. Graduados ESO con 16-17 años que no siguen estudiando (en % de graduados).**



**3.3. Graduados ESO con 16-17 años que estudian Bachillerato (en % de graduados que siguen estudiando).**




**3.4. Graduados ESO con 16-17 años que estudian FP Media (en % de graduados que siguen estudiando).**




Fuente: EPA 2006-2023 (INE).

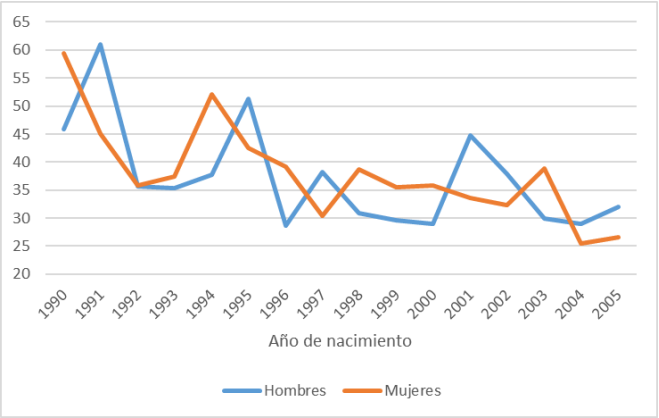
También se observaron cambios en las transiciones de la ESO a otros estudios medios (gráficos 3.3. y 3.4.). Analizando los datos año a año, hasta 2015, no se observó prácticamente ningún cambio en esta transición, pero a partir de entonces, la transición de ESO a ciclos formativos de Grado medio aumentó de forma notable, especialmente entre los hombres. El porcentaje de titulados de la ESO que decidió seguir estudiando FP Media se mantuvo estable desde un 13% para los nacidos en 1990 hasta un 14% para los nacidos en 2000 y posteriormente aumento hasta un 25% para los nacidos en 2007. El crecimiento de la transición a FP Media para las mujeres fue algo menos intenso y tardío. Se mantuvo en

el 8% para las nacidas entre 1990 y el 2000, para posteriormente aumentar hasta el 16% para las nacidas en 2007.

Centrándonos en la FP Media, no existen series largas oficiales sobre la tasa de graduación de diferentes cohortes de nuevo ingreso. Los datos proporcionados por el Ministerio de Educación y Formación Profesional indican que tan sólo el 57% de los hombres y el 67% de las mujeres que ingresaron en 2016-2017 habían finalizado los estudios en 4 o menos años[3]. En cualquier caso, la FP Media, que para muchos era una vía suave de entrada al mercado laboral, parece estarse convirtiendo poco a poco en una etapa intermedia más del sistema educativo. Así, el porcentaje de alumnos que tras graduarse de FP Media decide no seguir con los estudios se ha reducido considerablemente (gráfico 4.1.). Si bien la serie es algo volátil, de los nacidos entre 1990 y 1992 el 47% de los graduados de 18-19 años no siguieron estudiando. Sin embargo, para las generaciones de 2003 y 2005 este porcentaje se redujo al 30%. Por otro lado, la transición de FP Media a FP Superior ha ido aumentando de forma continua (gráfico 4.2.). En particular, si para los nacidos entre 1990 y 1992 el 67% de los hombres y el 62% de las mujeres proseguían su formación con estudios superiores de FP, para las generaciones entre 2003 y 2005 los porcentajes habían aumentado hasta el 82% y el 72% respectivamente.

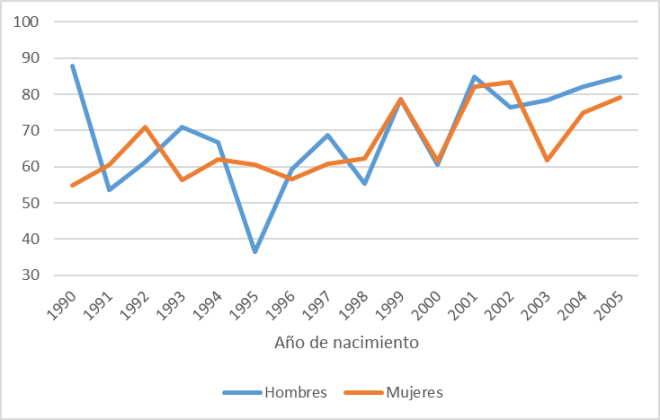
**Gráfico 4. Transiciones educativas desde FP Media (por año de nacimiento).**

**4.1. Graduados FP Media con 18-19 años que no siguen estudiando (en % de graduados).**




[3] Información de indicadores de rendimiento académico del alumnado que accede a FP del Ministerio de Educación y Formación Profesional.

**4.2 Graduados FP Media con 18-19 años que estudian FP Superior (en % de graduados que siguen estudiando).**



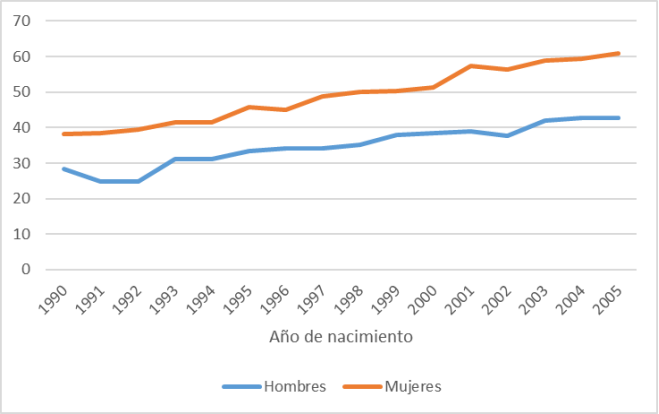

Fuente: EPA 2008-2023 (INE).

Por su parte, la tasa de éxito de los estudios de Bachillerato ha ido aumentando en los últimos años. Del curso 2007-2008 al curso 2020-2021 ha habido un aumento de 8pp en el porcentaje de alumnos de primero que promocionan a segundo hasta alcanzar el 86,4% y un aumento de 10pp de alumnos de segundo que finalizan el curso con éxito hasta el 85,2%[4]. Esto ha provocado un aumento notable de graduados de Bachillerato a los 18-19 años (gráfico 5.1.) y, en consecuencia, una disminución de los que se quedaban en el camino tan sólo con el título de ESO. La gran mayoría de los graduados de Bachillerato siguen sus estudios (gráfico 5.2.) y van a la Universidad (gráfico 5.3.) si bien también ha habido un aumento de la población que escoge seguir los estudios de FP Superior especialmente a partir de 2015 para hombres y 2019 para mujeres. Concretamente, como se ve en el gráfico 5.4., el porcentaje de hombres que siguen estudiando tras el Bachillerato y escogen FP Superior ha aumentado del 13% para la generación del 1990 hasta el 17% para los nacidos en el 1997 y 21% para los nacidos en 2005. Estos porcentajes para mujeres fueron del 14% para la generación del 1990 y 1997 y del 18% para la del 2005.

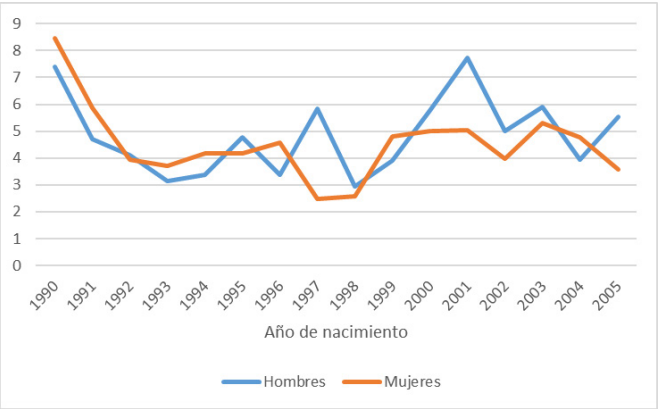
[4] Estadísticas de porcentaje de alumnado que promociona de curso por enseñanza del Ministerio de Educación y Formación Profesional.

**Gráfico 5. Transiciones educativas desde Bachillerato (por año de nacimiento).**

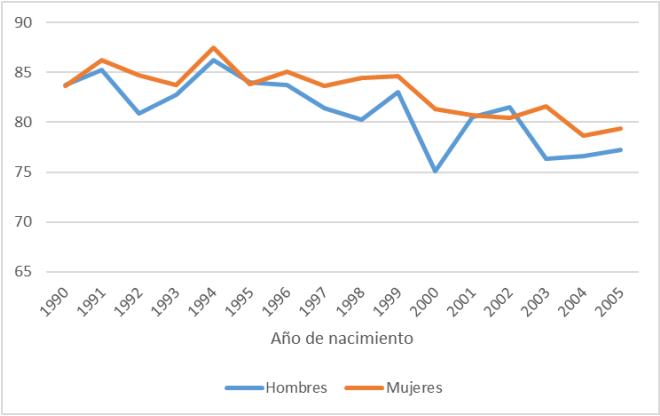
**5.1. Graduados Bachillerato con 18-19 años (en % total).**



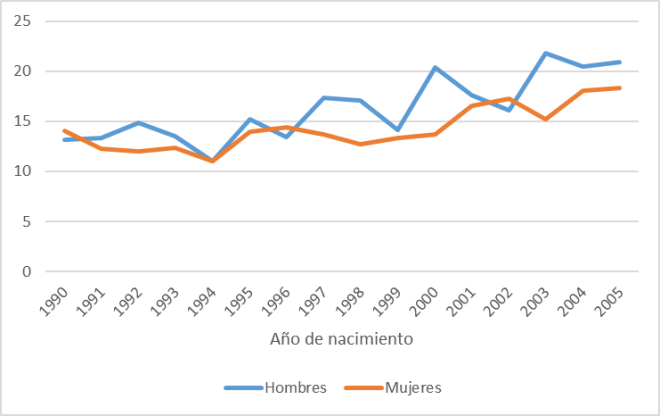

**5.2. Graduados Bachillerato con 18-19 años que no siguen estudiando (en % de graduados).**



**5.3. Graduados Bachillerato con 18-19 años que estudian Universidad (en % de graduados que siguen estudiando).**




**5.4 Graduados Bachillerato con 18-19 años que estudian FP Superior (en % de graduados que siguen estudiando).**




Fuente: EPA 2008-2023 (INE).

Respecto a la FP Superior, de acuerdo a los datos del Ministerio de Educación y FP el porcentaje de alumnado de nuevo ingreso en FP Superior que finalizó el Grado en 4 o menos años fue del 70% para hombres y del 78% para mujeres en el curso de nuevo ingreso de 2016-2017[5]. Asimismo, el 21% de alumnos varo-

[5] Estadísticas de seguimiento educativo y rendimiento académico del alumnado que accede a FP del Ministerio de Educación y Formación Profesional.

nes y el 30% de mujeres que se graduaron de FP Superior en el curso académico 2018-2019 se matricularon en la Universidad posteriormente[6]. Estas estadísticas son relativamente nuevas por lo que no existe una serie suficientemente larga como para ver la evolución a lo largo del tiempo.

Finalmente, relativo a estudios universitarios y de acuerdo a los datos del Ministerio de Universidades, la tasa de graduación de la Universidad definida como el porcentaje de estudiantes que finaliza la titulación en el tiempo teórico previsto o un año más se situó en el 48% para los nuevos ingresos masculinos y el 61% para los femeninos en el curso 2012-2013 y aumentó unas décimas hasta el curso 2015-2016[7]. La tasa de transición de Grado a Máster (dentro del sistema español) en el curso inmediatamente posterior fue del 27% para los graduados en el año 2014-2015 y del 20% para las graduadas. Cinco años más tarde, si bien el porcentaje de los hombres no varió el de las mujeres había aumentado 3pp hasta el 23%.

En resumen, las generaciones nacidas tras 1978 experimentaron un estancamiento de la mejora formativa que había caracterizado a las cohortes anteriores, perpetuando un nivel de abandono escolar muy elevado para los estándares europeos. La reducción del abandono escolar se retomó con la generación de 1990 por un incremento de las personas que deciden seguir estudiando tras la ESO, una mayor entrada en la FP Media y un mayor éxito en la consecución del Bachillerato. Además de reducir el abandono escolar, las generaciones nacidas a partir del 2000 también han incrementado la transición de estudios obligatorios y de Bachillerato hacia estudios profesionales, especialmente entre los hombres. La sección 3 presenta un marco teórico y empírico para analizar algunas de las razones que pueden explicar estos movimientos.

Ejercicio 1. *El Instituto Nacional de Estadística (INE) realiza la encuesta de transición educativa-formativa e inserción laboral. Esta encuesta, de periodicidad cuatrienal, que se realiza a estudiantes de diferentes niveles educativos no universitarios aporta información muy relevante sobre sus transiciones educativas y laborales. Utilizando los resultados de 2019, analiza en qué medida las transiciones de la encuesta son coherentes con las publicadas en este capítulo a partir de la encuesta*

---

[6] Estadísticas de seguimiento educativo posterior de los graduados de FP del Ministerio de Educación y Formación Profesional.

[7] Indicadores de rendimiento académico del Ministerio de Universidades.

*de población activa. Razona las ventajas y desventajas de utilizar la información de una u otra fuente.*

## 3. ¿MERECE LA PENA SEGUIR ESTUDIANDO?

### 3.1. Marco teórico

Para analizar qué factores se encuentran detrás de la decisión de estudiar tras la educación obligatoria se presenta brevemente el modelo de Altonji (1993). En un primer momento, periodo 1, el estudiante finaliza la ESO y debe decidir qué hacer ($s_1$): puede ponerse a trabajar ($s_1=l$), dirigirse a estudios de FP ($s_1=p$) o seguir la vía académica ($s_1=a$). Seguir estudiando tras la ESO corresponde a cursar FP Media en la vía profesional y Bachillerato en la vía académica. Si el individuo se pone a trabajar dando por finalizado el periodo formativo, obtiene el valor presente descontado de su expectativa de ingresos a lo largo de la vida $Y_0$. Seguir estudiando dota a la persona de conocimientos que posteriormente serán remunerados con una rentabilidad por encima de la corriente de ingresos anterior. En las próximas subsecciones se tratará la estimación de diferentes rentabilidades según el nivel de estudio. Por un lado, la formación profesional ofrece conocimientos profesionales ($k_p$) distintos a los académicos ($k_a$), normalmente más aplicados y destinados a aprender el funcionamiento de un oficio en el nivel tecnológico actual. Los conocimientos académicos suelen ser más teóricos y útiles para adaptarse a un mundo cambiante o para desarrollar nuevas tecnologías. Esta diferencia entre ambos conocimientos puede dar resultado a rendimientos distintos a lo largo de la vida.

Algo importante a definir es la forma en que cada persona acumula conocimientos mientras está estudiando. Esta acumulación de conocimientos servirá para conseguir aprobar grados que darán acceso a estudios superiores. De manera simplificada, especificamos un nivel inicial $k_0$ de conocimientos que pueden depender de muchos ingredientes, algunos de ellos tratados en el capítulo 3b de este libro. La acumulación de conocimientos en el periodo 1 dependerá de este nivel inicial de conocimiento, del conjunto de decisiones de formación y de un componente aleatorio que puede relacionarse con, entre otras cosas, la posibilidad de tener un mal profesor o una mala clase, elementos que se discuten en los capítulos 3c y 3d. Como más adelante se definirá un segundo periodo en el que se podrá decidir si seguir con estudios superiores, de manera general, se puede definir la acumulación de conocimientos en cada periodo como:

$$k_{t+1,p} = f(k_{t,p}, s_t, \varepsilon_t)$$
$$k_{t+1,a} = f(k_{t,a}, s_t, \varepsilon_t)$$

Si se toma la decisión de seguir estudiando en FP Media o en Bachillerato, el estudiante descubrirá en ese momento si es un estudiante del tipo $\theta_0$ (persona a la que no le gusta estudiar ni la vía académica ni la profesional, de tal manera que a pesar de recibir un salario mayor con FP Superior o Universidad prefiere no seguir formándose), tipo $\theta_1$ (persona a la que no le gusta estudiar, pero está dispuesta a aprender algún oficio, con lo que sus preferencias serían tal que el salario universitario no compensaría nunca la utilidad de ponerse a trabajar o estudiar FP) o tipo $\theta_2$ (que no tiene restricción de preferencias).

Para facilitar la exposición, y seguir lo más posible el modelo de Altonji, se asume que la FP Media y el Bachillerato duran 1 periodo y que todo el mundo aprueba ese Grado entrando en lo que denominaremos el periodo 2. En realidad, dado lo observado en el caso español, se podría añadir algún parámetro que permitiera diferentes probabilidades de éxito de la FP Media y el Bachillerato. En el periodo 2, de nuevo, se debe decidir si abandonar el sistema educativo o seguir estudiando. Realizar Bachillerato o FP Media difiere en términos de rendimiento económico de haberse quedado con la ESO en una proporción $r_1$ pero no hay diferencia entre ellas. Como los ingresos se cobran tras haber pasado el primer periodo su valor presente se debe de descontar con el tipo de interés de un año: $Y_0(1+r_1)/(1+R)$.

Si el estudiante continúa en el sistema, puede decidir proseguir por la rama de estudios que había escogido anteriormente o cambiar a la otra si lo prefiere. Concretamente, al revelarse la información de la tipología en este periodo, es posible que haya alumnos que habían decidido cursar Bachillerato que deciden cambiarse a FP Superior al ser del tipo $\theta_1$ Recibir ingresos superiores por los conocimientos aprendidos requiere de la obtención de un título superior por lo que el éxito de la segunda etapa vendrá determinado por la probabilidad de finalizar con título los estudios. Las probabilidades de éxito $g_p$ y $g_a$ dependerán de los conocimientos que se tengan al finalizar los estudios por lo que en definitiva dependerán, de los conocimientos iniciales, las decisiones de formación anteriores y los diferentes factores aleatorios fuera del control del estudiante $F_p(g_p \mid k_0, s_1)$, $F_a(g_p \mid k_0, s_1)$.

Finalizar los estudios de FP superior supone un incremento de $r_p$ lo que supone unos ingresos a lo largo de la vida de: $Y_0(1+r_p)/(1+R)^2$ y los estudios universitarios de $r_a$: $Y_0(1+r_a)/(1+R)^2$.

Con este planteamiento ya se puede resolver la decisión de continuar estudiando tras la ESO. En este sentido, en el periodo 2, aquellas personas que son $\theta_0$ no proseguirán con sus estudios y ganarán:

$$V_2(g_2, 0) = \frac{Y_0(1+r_1)}{(1+R)}$$

Por otro lado, los del tipo $\theta_1$ tendrán que decidir si ponerse a trabajar o realizar FP Superior según la opción que les reporte más ingresos:

$$V_2(g_2, 1) = max\left[g_{2p}\frac{Y_0(1+r_p)}{(1+R)^2} + (1-g_{2p})\frac{Y_0(1+r_1)}{(1+R)}, \frac{Y_0(1+r_1)}{(1+R)}\right]$$

Y los de tipo $\theta_2$, tendrán las tres opciones para escoger:

$$V_2(g_2, 2) = max\left[g_{2a}\frac{Y_0(1+r_a)}{(1+R)^2} + (1-g_{2a})\frac{Y_0(1+r_1)}{(1+R)}, g_{2p}\frac{Y_0(1+r_p)}{(1+R)^2} + (1-g_{2p})\frac{Y_0(1+r_1)}{(1+R)}, \frac{Y_0(1+r_1)}{(1+R)}\right]$$

En consecuencia, tras la ESO, cuando uno no sabe sus preferencias, pero sí las probabilidades incondicionales en la población ($\theta_0$, $\theta_1$, $\theta_2$), escoger la vía formativa $s_1$ reportará unos ingresos esperados:

$$\begin{aligned} V_1(F, s_1) = {} & \theta_0 E\{V_2(g_2, 0) \mid k_0, s_1\} \\ & + \theta_1 E\{V_2(g_2, 1) \mid k_0, s_1\} \\ & + \theta_2 E\{V_2(g_2, 2) \mid k_0, s_1\} \end{aligned}$$

Por tanto, una persona decide empezar el Bachillerato o la FP Media en la medida que:

$$max[V_1(k_0, p), V_1(k_0, a)] > Y_0$$

Y decidirá por la FP Media en la medida que:

$$V_1(k_0, p) > V_1(k_0, a)$$

Altonji (1993) muestra diferentes implicaciones de este modelo que son de interés para el análisis empírico que se va a realizar posteriormente.

La probabilidad de seguir con estudios más allá de la ESO se incrementa con $r_1$, $r_a$, $r_p$. Por otro lado, cambios en alguna de las anteriores variables que au-

menten el porcentaje de alumnos que continúan los estudios post-obligatorios no necesariamente tienen que aumentar la probabilidad de graduarse. Concretamente, la probabilidad de graduarse en estudios superiores no tiene porqué subir con aumentos de $r_a$, $r_p$. Esto es así porque en el caso de que los rendimientos sean muy elevados, se atraerá a estudiantes que deciden probar esa opción, aunque tengan pocas probabilidades de éxito. En este sentido, un aumento de la proporción de alumnos que prosiguen estudios post-obligatorios no se traslada automáticamente a un aumento del porcentaje de graduados superiores.

Asimismo, la decisión de cursar formación profesional en relación a la vía académica aumenta con $r_p - r_a$.

Finalmente, el modelo incorpora el carácter secuencial de la decisión de educación con lo que se debe distinguir entre rendimiento exante y expost, algo que va a tener consecuencias para la estimación de los rendimientos a la educación de la sección 3.2.

Ejercicio 2. *Enuncia al menos dos mecanismos dentro del modelo de Altonji por los que un estudiante podría decidir abandonar los estudios a pesar de que la elección escogida presente un rendimiento expost elevado. ¿Se te ocurren otros mecanismos por los que alguien pudiera querer cambiar su elección y que no estén en el modelo presentado? [Pista: Una vez hayas pensado algún mecanismo adicional puedes encontrar algunas ideas relacionadas en Wiswall y Zafar (2015) o Coelli (2011)].*

## 3.2. Marco empírico de la decisión de educarse y los rendimientos de la educación

En la sección 3.2.1. se presentan los resultados de una pequeña muestra de estudios que estiman de forma causal la elasticidad de las oportunidades económicas de los trabajadores con poca cualificación sobre el abandono escolar. El marco empírico común correlaciona la decisión de seguir estudiando o la probabilidad de disponer de un Grado superior con cambios en las oportunidades económicas relativas de ambos colectivos. Si $S$ es el resultado educativo que depende de los ingresos relativos de las opciones educativas $\left[\frac{W(S)}{W(0)}\right]$, otras variables observables ($z$) e inobservables ($u$) para el económetra, la regresión utilizada en los estudios es la siguiente:

$$S(W(S)/W(0), z) = \beta_0 + \gamma \left[\frac{W(S)}{W(0)}\right] + \beta_1 z + u$$

Como se ha visto en la anterior sección, dentro de $u$ puede haber variables como los conocimientos innatos o la familia que afecten tanto a la decisión de educarse como al rendimiento laboral de los más formados. En consecuencia, la estimación anterior por Mínimos Cuadrados Ordinarios (MCO) no sería adecuada. Para evitar este sesgo se suelen buscar situaciones en las que una perturbación económica hubiera aumentado la demanda de oportunidades laborales de un colectivo en particular. Por ejemplo, un boom en la construcción, en la agricultura o algunas industrias como la minería o la exportación del textil debería aumentar en mayor medida las oportunidades laborales de trabajadores poco cualificados.

Muchos otros estudios no se han interesado en calcular la elasticidad de seguir estudiando respecto a las oportunidades laborales, sino que pretenden estimar directamente la tasa interna de rentabilidad de realizar un año más de educación ( $\rho$). Al pensar la educación como una inversión, este parámetro es suficiente para determinar si es necesario formar más a la gente o si existe una sobreinversión en educación. El rendimiento de realizar $s$ años de educación se define como la tasa de descuento que iguala la suma del valor presente de las ganancias a lo largo de los años de experiencia $x$ (hasta la edad de jubilación X que puede depender de los años de formación) con los costes tanto de oportunidad (el valor presente de los ingresos de no educarse) como los costes monetarios y no monetarios de formarse $C(s)$. Matemáticamente, en tiempo discreto, se definiría así:

$$\sum_{x=0}^{X(s)} \frac{W(s,x)}{(1+\rho)^{x+s}} - \sum_{x=0}^{X} \frac{W(0,x)}{(1+\rho)^{x}} - C(S) = 0$$

Ese rendimiento se puede comparar con el tipo de interés al que una persona se puede endeudar, o con el rendimiento de inversiones alternativas para decidir si es una "buena" o "mala" inversión seguir estudiando.

Desde Mincer (1974) el término rendimiento de la educación también se aplica al coeficiente de la variable "años de educación" en la siguiente ecuación. Si $W(s,x)$ es el nivel salarial dado $s$ años de educación y $x$ los años de experiencia, el logaritmo del salario vendría representado por:

$$\ln(W(S,x)) = \alpha_0 + \rho S + \alpha_1 x + \alpha_2 x^2 + \varepsilon$$

La justificación a este modelo proviene de la teoría de capital humano que indica que las inversiones en capital humano se hacen principalmente al prin-

cipio de la vida ya sea en forma de educación o como consecuencia de tiempo gastado en formación durante los primeros años en el trabajo (Becker, 1962). Para permitir un menor aprendizaje a lo largo de la vida se incorpora un término cuadrático en la experiencia.

Heckman et al. (2006) detalla las condiciones que se deben de dar para que el coeficiente de la anterior regresión estimada por MCO sea igual a la tasa interna de rendimiento de la anterior expresión:

1) No deben existir costes a educarse.
2) Los años de vida activa deben de aumentar uno a uno con los años de educación.
3) Los salarios deben aumentar linealmente con los años de educación sin existir no linealidades en diferentes niveles formativos.
4) Debe existir separabilidad entre la educación y los años de experiencia, o sea las pendientes a lo largo del ciclo vital de diferentes niveles educativos deben ser iguales y no tienen por qué serlo.
5) No debe de haber correlación entre $\varepsilon$ y $S$, o lo que es lo mismo no debe de existir ninguna variable omitida en la regresión que pueda afectar tanto al salario como a la decisión de educarse.
6) Los rendimientos deben de ser homogéneos en la población.

Ejercicio 3. *Si se asume que el salario dado un nivel educativo es constante a lo largo de la vida y que la edad de jubilación aumenta 1 a 1 con los años de educación, el valor presente de los ingresos, expresado en tiempo continuo, dado un nivel educativo s es:*

$$V(s) = \int_{s}^{T+s} \frac{W(s)}{e^{rt}}\, dt$$

*Demostrar, como decía Heckman, que en ausencia de costes de educarse la tasa interna de rentabilidad de la educación* $(V(s)=V(0))$ *es igual al coeficiente de una regresión de Mincer.*

Una revisión de varios artículos realizada por Card (2018) sugiere que el coeficiente por año de educación de esta regresión en Europa se situaría entre un 5% para Suecia y un 10% para Reino Unido. En EE. UU. el coeficiente se situaría cerca del 10%. Este rendimiento es muy elevado comparado con las rentabilidades de muchas inversiones y sugeriría un importante incentivo a estudiar. En

este sentido, es posible que varios de los factores anteriores estén operando y reduzcan los rendimientos a la educación para al menos un colectivo importante de la población. Heckman et al. (2006) muestra en el caso de Estados Unidos que los costes monetarios de educarse reducen el rendimiento estimado tan sólo en uno o dos puntos porcentuales, por lo que seguramente haya costes no monetarios importantes. Por otro lado, tener en cuenta una forma funcional diferente a la tradicional en la estimación sí que suele tener un mayor impacto en la estimación de los rendimientos. Concretamente, los anteriores autores utilizan formas funcionales que no obligan a una linealidad de los rendimientos por años de educación y a una independencia de la educación con la experiencia. Así, obtienen rendimientos diferentes a los estimados por la ecuación de Mincer principalmente en la obtención del Grado de Bachillerato (entre 10 y 12 años) y en el Grado Universitario (entre 12 y 16). Finalmente, respecto a los anteriores puntos (4) y (5), Card (2001) describe la forma en la que el sesgo de selección y la heterogeneidad de rendimientos en la población pueden afectar las estimaciones. En particular, el sesgo de selección debería sesgar al alza las estimaciones. Sin embargo, en la mayoría de los estudios que utilizan cambios en el sistema educativo para instrumentar la decisión de educarse se encuentra un coeficiente mayor que el de MCO. Según Card, esto podría ser debido a que, si bien las variables instrumentales solucionan el problema de endogeneidad, tienen una difícil interpretación si existe una elevada heterogeneidad de rendimientos en la población. En particular, la estimación por variables instrumentales identifica desproporcionadamente los rendimientos de la población[8] más afectada por los cambios en la variable instrumental. De esta manera, si se utiliza como instrumento una reforma del sistema educativo, es fácil pensar que esta puede tener como principal objetivo mejorar el rendimiento académico de los grupos más desfavorecidos. En la medida en que gran parte de estos colectivos no estudiara porque no podían costearse la educación en vez de porque no esperaba buenos rendimientos del estudio, el coeficiente de variables instrumentales sobre-ponderaría rendimientos más elevados. En definitiva, la estimación de la tasa interna de rendimiento en una economía y momento determinado no es sencilla. Los estudios que se aventuren a este propósito requieren, al menos, de una discusión de los supuestos detrás de ella, entre los que destacan la forma funcional escogi-

[8] Otros autores han abogado porque el coeficiente de MCO podría estar sesgado a la baja por la existencia de error de medida, algo que se corregiría con variables instrumentales (Griliches, 1977).

da en la estimación, así como de la heterogeneidad que pudiera existir de estos rendimientos en la población.

Existen otras consideraciones que se tienen que tener en cuenta cuando se estiman regresiones de Mincer. En primer lugar, nótese que los ingresos pueden ser brutos o netos de impuestos. En el segundo caso, la progresividad de los esquemas de imposición de la renta llevará en general a un menor rendimiento de la educación. Asimismo, la regresión se puede hacer para los ingresos de una hora de trabajo, de un mes o de un año. Dado que, en general, el número de horas y días que una persona trabaja a lo largo de un mes o año es mayor cuanta más educación tenga, los diferenciales tenderán a subir con la unidad de medida. Esto es especialmente relevante para España, país donde la tasa de paro y la precariedad es elevada para las personas con menor formación. Adicionalmente, algunos estudios calculan los diferenciales salariales de un Grado superior en relación a un Grado inferior sin usar la variable continua de años de educación. En este caso, no se debe interpretar directamente el coeficiente de la regresión como tasa interna de rentabilidad ya que el rendimiento debe de descontar los ingresos que se cobran tras pasar varios años estudiando. En consecuencia, el rendimiento del Grado será menor que el coeficiente estimado en la regresión.

Finalmente, hay que tener en cuenta que el rendimiento estimado a partir de una ecuación de Mincer es el rendimiento expost de un año de educación. Pongámonos en el caso del modelo especificado en 3.1. y el cálculo de los rendimientos de la FP Media. La tasa interna de rentabilidad de la FP Media $\rho_{FPm}^{expost}$ coincide con $r_1$

$$\frac{Y_0(1+r_1)}{(1+\rho_{FPm}^{expost})} - Y_0 = 0$$

$$\rho_{FPm}^{expost} = r_1$$

Sin embargo, la tasa interna de rentabilidad exante $\rho_{FPm}^{exante}$, que es la variable relevante para la decisión de seguir estudiando, es más compleja. En este sentido, el cálculo exante no sólo debe incorporar el rendimiento expost de la FP Media que se materializará en el caso de no seguir estudiando tras la graduación o en el caso de no haber sido capaz de finalizar estudios superiores a pesar de haberlos empezado ($p_{Fpm,a2}$, $p_{Fpm,p2}$) sino también los rendimientos de disponer de estudios universitarios o de FP superior en el caso de que los acabe ($p_{Fpm,a2}$, $p_{Fpm,p2}$):

$$(p_{Fpm,a1} + p_{Fpm,a2} + p_{Fpm,p2})\ \frac{Y_0(1+r_1)}{(1+\rho_{FPm}^{exante})} + p_{Fpm,a2}\frac{Y_0(1+r_a)}{(1+\rho_{FPm}^{exante})^2} + p_{Fpm,p2}\frac{Y_0(1+r_p)}{(1+\rho_{FPm}^{exante})^2} - Y_0 = 0$$

Evidentemente, la estimación de estos rendimientos exante requiere de un mayor número de datos, en particular, datos apropiados de transiciones educativas y probabilidades de éxito por lo que no hay muchos estudios que los calculen (véase Cunha y Heckman, 2007, para más detalles y comparaciones entre rendimientos exante y expost).

### *3.2.1. Seguir estudiando durante el boom inmobiliario 1995-2008*

A diferencia de otras economías, España redujo la prima salarial de la educación entre 1995 y 2008. Esta caída del diferencial salarial por nivel educativo ha sido corroborada a partir de diferentes bases de datos (Pijoan-Mas y Sánchez-Marcos, 2010; Izquierdo y Lacuesta, 2012; Bonhomme y Hospido, 2017) y contrasta con lo que se ha documentado para otros países como EE. UU. (Autor et al.. 2008), Reino Unido (Gosling et al., 2000), Canadá (Fortin et al., 2012) y Alemania (Dustmann et al., 2009). En esos países se ha argumentado que, en la parte alta de la distribución, hubo un incremento de la demanda de trabajadores con educación superior provocado por cambios en la tecnología. En particular, algunos autores han sugerido la importancia de la complementariedad entre la educación y la tecnología de la información. Sin embargo, el crecimiento de España a finales de los 90 y principios del 2000 fue provocado por un boom inmobiliario que incrementó la demanda relativa de trabajadores de la construcción. Es esperable que este boom de oportunidades laborales para los menos formados cambiara los incentivos a seguir en el sistema educativo. Sustrayendo 17 años al inicio del boom inmobiliario y la crisis financiera, tenemos que las generaciones nacidas entre 1978 y 1991 se deberían ver afectadas.

Usando este periodo de caída del rendimiento a la educación derivada del boom inmobiliario, varios autores han relacionado de forma causal las oportunidades económicas de los poco formados con el freno en la mejora educativa de España. Aparicio-Fenoll (2016) utiliza información de la Encuesta de Población Activa 2001-2006 para la población entre 16 y 18 años, información de salarios a nivel provincial para el mismo periodo de la MCVL 2007 y el valor añadido provincial de construcción de las cuentas regionales del INE. Su hipótesis es que el peso de la construcción a nivel provincial correlaciona con las oportunidades laborales de los hombres con poca formación en relación con las mujeres. Por otro lado, argumenta que el crecimiento de la actividad constructora a nivel provincial fue motivado por causas exógenas a la decisión de educarse, lo que lo convierte en un buen instrumento para las diferencias por género del diferencial

salarial entre formados y no formados. De esta forma, analiza diferencias en la decisión de seguir educándose de hombres y mujeres según variaciones en los diferenciales salariales relativos motivados por la actividad constructora a nivel provincial. El artículo encuentra que una disminución del 10% en el diferencial salarial de educados y no educados disminuye la probabilidad de matriculación en la educación posobligatoria en por lo menos un 2% de la tasa media de transición (de un 80%) y disminuye la proporción de quienes completan el Grado en al menos 0,2% de la tasa promedio de finalización de Grado (87%). Lacuesta et al. (2020) confirma el resultado con otra estrategia empírica. Los autores utilizan información de la Encuesta de Población Activa 1993-2007 para la población de 17 y 25 años, información de salarios a nivel provincial para el mismo periodo de las diferentes olas de la MCVL y el registro de convenios colectivos entre 1990 y 2009. La hipótesis de partida es que los convenios colectivos para diferentes categorías profesionales correlacionan bien con los diferenciales salariales por grupo de cotización. Asimismo, los autores argumentan que quien negocia los convenios no tiene en cuenta las oportunidades laborales de los jóvenes por lo que son un buen instrumento para los diferenciales salariales por nivel formativo. Así, analizan a nivel provincial como cambios en los salarios relativos derivados de lo negociado en convenio afectan las decisiones educativas. Encuentran que un aumento del 10% en los salarios relativos de trabajadores no cualificados, para las cohortes que cumplen 17 años genera un aumento de entre 2 y 6 puntos porcentuales en la probabilidad de quedarse sólo con la educación obligatoria. Su especificación empírica intenta ir más allá al introducir una distinción entre trabajadores con formación universitaria o técnica. Gracias a esta distinción pueden argumentar que el aumento en la proporción de jóvenes varones con escolaridad básica se da a expensas de la proporción de jóvenes varones con trayectoria académica Secundaria y no de formación profesional.

Elasticidades similares han sido encontradas en otros contextos geográficos y temporales. En particular, Black et al. (2005) analiza la decisión de matricularse en cursos de Secundaria post-obligatoria en Kentucky y Pensilvania en los 70's y 80's. Esta decisión fue afectada tras el embargo a los envíos petrolíferos de la OPEP a EE. UU. a principios de los 70's provocando un boom en las minas de carbón de la zona rural de Appalachia. Así, aquellos municipios con mayores reservas de carbón experimentaron una caída de los salarios relativos durante los 70's y una subida durante los 80's. Las estimaciones indican que un aumento del 10% en los ingresos de los trabajadores poco cualificados podría disminuir las tasas de matriculación

en la escuela post-obligatoria Secundaria hasta en un 5%-7%. Asimismo, Atkin (2016), documenta los cambios en la formación en México derivados del crecimiento de oportunidades laborales para las personas con menos nivel educativo durante un período de reformas comerciales que tuvo lugar entre los años 1986 y 2000. En concreto, aquellas generaciones que a los 16 años experimentaron en sus comunidades un mayor incremento del empleo de las industrias exportadoras alcanzaron un número de años de educación menor. Por cada 25 nuevos puestos de trabajo, habría un abandono de 1 estudiante de noveno Grado.

Ejercicio 4. *¿Podrías indicar qué fuentes de datos se han utilizado en Pijoan-Mas y Sánchez-Marcos (2010), Izquierdo y Lacuesta (2012), y Bonhomme y Hospido (2017) para estimar la evolución de los rendimientos de la educación en España a lo largo del tiempo? Fíjate en las diferentes definiciones del concepto de renta y en las diferentes muestras para describir las ventajas e inconvenientes de cada una de estas bases de datos.*

Ejercicio 5. *Además de cambiar el rendimiento de la educación, el boom inmobiliario provocó un cambio en el precio de los activos, especialmente en las viviendas. Argumenta en qué medida este canal podría afectar las decisiones educativas de las familias distinguiendo por el nivel de renta del hogar y el país de residencia [Pista: Una vez hayas pensado algún argumento puedes encontrar algunas ideas relacionadas en Lovenheim (2011)].*

### *3.2.2. Evidencia descriptiva de los rendimientos tras la crisis financiera*

Para disponer de un periodo temporal amplio antes y después de la crisis financiera y un nivel suficiente de desagregación de los niveles educativos, en esta subsección, se va a ilustrar cómo calcular la tasa interna de rentabilidad de diferentes niveles educativos a partir de datos de la encuesta financiera de las familias. En primer lugar, uno tiene que estimar los coeficientes de una regresión de Mincer para el ingreso laboral anual bruto de los asalariados de la Encuesta Financiera de las Familias entre 2001 y 2019[9]. Las regresiones incluyen experiencia, y experiencia al cuadrado calculadas como edad-16 (por los años a los que se empezaría a

[9] La variable de ingresos utilizada es el ingreso laboral del año anterior a la encuesta en términos brutos. Para convertir los ingresos brutos en netos y obtener una TIR que tenga en cuenta el IRPF español se podría aplicar a la renta bruta una función de impuestos con los parámetros de la tabla 10 en García-Miralles et al. (2019).

trabajar) para las personas sin formación post-obligatoria, edad-18 para FP Media y Bachillerato, edad-20 para FP Superior y edad-22 para universitarios.

A partir de estas regresiones se estiman ingresos laborales de 0 a 50 años de experiencia y se calculan tasas internas de rentabilidad (según la ecuación en tiempo discreto de esta sección) asumiendo que quien se pone a trabajar tras la ESO puede empezar a hacerlo 2 años antes que quien estudia FP Media o Bachillerato. Igualmente, quien estudia FP Superior empezará dos años más tarde que el que ha cursado estudios post-obligatorios y el universitario cuatro. La tasa se estima asumiendo que más allá de empezar a trabajar más tarde el coste de formarse es 0. Como se observa en la tabla 1, entre 2001 y 2007, el rendimiento de estudiar FP Media o Bachillerato se situó alrededor del 10% para hombres y entre el 13% y el 25% para mujeres. Realizar estudios de FP Superior en comparación con FP

**Tabla 1. Tasa Interna de Rentabilidad (TIR) de ingresos anuales entre diferentes niveles educativos.**

| | 2001 | 2007 | 2013 | 2019 |
|---|---|---|---|---|
| | **Hombres** | | | |
| Bachillerato o FP Media *vs* ESO | 10% | 10% | 20% | 18% |
| FP Superior *vs* Bachillerato o FP Media | 11% | 9% | 5% | 16% |
| Universidad *vs* Bachillerato o FP Media | 13% | 13% | 19% | 20% |
| | **Mujeres** | | | |
| Bachillerato o FP Media *vs* ESO | 25% | 13% | 20% | 26% |
| FP Superior *vs* Bachillerato o FP Media | 8% | 14% | 13% | 4% |
| Universidad *vs* Bachillerato o FP Media | 15% | 19% | 25% | 23% |

Fuente: Elaboración propia a partir de datos de la Encuesta Financiera de las Familias.
Nota: En el cuadro se presentan la TIR de realizar un estudio superior a otro. La TIR de realizar un nivel educativo se define como la tasa de descuento que iguala la suma del valor presente de las ganancias a lo largo de 50 años de experiencia con los costes de oportunidad (el valor presente de los ingresos de no educarse). Se asumen costes monetarios cero. Para obtener las ganancias laborales brutas por año de experiencia se realizan unas ecuaciones de Mincer restringidas a una muestra con los dos niveles educativos analizados añadiendo una constante, una variable ficticia del nivel educativo superior, una variable de experiencia potencial y su cuadrado definida como años-16 para ESO, años-18 para estudios medios, años-20 para FP Superior y años-22 para universitarios. Se tienen en cuenta pesos y las cinco imputaciones de la EFF en la estimación. A partir de las regresiones de la tabla 1 se estiman ingresos laborales de 0 a 50 años de experiencia y se calculan tasas internas de rentabilidad asumiendo que quien se pone a trabajar tras la ESO puede empezar a hacerlo 2 años antes de quien estudia FP Media o Bachillerato. Igualmente, quien estudia FP Superior empezará dos años más tarde que el que ha cursado estudios post-obligatorios y el universitario cuatro.

Media o Bachillerato reportó un 9% y un 11% para hombres y entre un 8% y un 14% para mujeres. Finalmente, disponer de un título universitario tuvo una rentabilidad de un 13% para hombres y entre 15% y 19% para mujeres. Con la crisis, entre 2007 y 2013, aumentó mucho la rentabilidad de realizar estudios medios, así como realizar estudios universitarios tanto para hombres como para mujeres. Sin embargo, el rendimiento de realizar estudios de FP Superior cayó 4pp y se mantuvo prácticamente estable para mujeres. Finalmente, con la recuperación económica tras 2013, el rendimiento de la FP Superior para hombres aumentó notablemente hasta el 16% para acercarse al rendimiento de los universitarios que se mantuvo estable. Sin embargo, el rendimiento de la FP Superior para mujeres no ha mejorado. Al contrario, incluso ha retrocedido hasta el 4%.

En resumen, la evolución de las tasas internas de rentabilidad entre 2007 y 2019 sería coherente con un aumento de los rendimientos de la formación postobligatoria entre la crisis financiera y la recuperación que se inició en 2013 para mantenerse constante posteriormente. Si bien entre 2007 y 2013 la formación universitaria amplió su rendimiento relativo respecto a la FP Superior, esta última opción ha ganado terreno tras 2013 para los hombres. Finalmente, la rentabilidad de la FP Superior para las mujeres no ha mejorado en esos años y sigue presentando un diferencial importante respecto a la Universidad.

Ejercicio 6. *Calcula la TIR de los estudios medios de las mujeres en 2001 si la regresión de Mincer de estudios medios respecto a ESO presenta los siguientes coeficientes: una constante de 8,5; un coeficiente de los estudios de 0,45; un coeficiente de la experiencia de 0,03 y un coeficiente de –0,0004 para su cuadrado. ¿Cómo varía esa TIR si se tiene en cuenta el coste monetario de estudiar y el precio de acceder a Bachillerato es de 1.000 euros por* año?

Ejercicio 7. *Discute cómo podría afectar el cambio en la composición de la población que tiene formación al nivel y a la evolución de la TIR según lo que se asuma sobre la heterogeneidad de los rendimientos en la población (Pista: Te puede ayudar pensar la regresión de Mincer como un modelo de coeficientes aleatorios).*

$$\ln\big(W(S,x)\big) = \alpha_0 + \rho_i S + \alpha_1 x + \alpha_2 x^2 + \varepsilon$$

Ejercicio 8. *Este capítulo trata la decisión individual de educarse para la que se tiene en cuenta su impacto en el ingreso potencial. Sin embargo, otra pregunta distinta sería la decisión del sector público en financiar una parte de esa educación para lo que nos deberíamos preguntar por diferencias entre el rendimiento social e*

*individual de la educación. A partir de lo aprendido en otros capítulos del libro, sugiere mecanismos que justificarían rendimientos sociales de la educación más allá de los individuales. ¿Podrías cuantificarlos en relación a los individuales?*

El siguiente recuadro documenta la evidencia internacional y los argumentos esgrimidos sobre el debilitamiento de los rendimientos universitarios tras el 2000 y en especial tras la crisis financiera. No conocemos ningún estudio que haya documentado un reciente auge de los rendimientos expost de FP Superior, si bien sí que existe evidencia de la ausencia de perfiles técnicos en un contexto de cambio tecnológico acelerado a nivel europeo (Arpaia y Halasz, 2023). En el caso de España, este crecimiento puede haberse visto reforzado tras el Covid gracias, por ejemplo, al despliegue de los proyectos de inversión vinculados al programa Next Generation EU (NGEU). En este sentido, el porcentaje de empresas que declaran que la escasez de mano de obra impacta de forma negativa o muy negativa en su actividad se ha elevado de forma considerable. De acuerdo con la Encuesta del Banco de España sobre la Actividad Empresarial (EBAE), en el primer trimestre de 2023 el 35% de las compañías compartían esta percepción, 25 puntos porcentuales (pp) más que a principios de 2021 (Cerezo e Izquierdo, 2023). Esta escasez es especialmente elevada en agricultura, construcción, hostelería, información y comunicación y transporte (Cuadrado y Montero, 2022). En la siguiente sección se presentarán algunos estudios que analizan en qué medida una mejor entrada en el mercado laboral de perfiles técnicos en relación a perfiles académicos se puede perpetuar a lo largo de la vida laboral.

**Debilidad de los rendimientos universitarios a nivel internacional tras el año 2000**

En comparativa internacional alrededor del año 2000, se empezó a ver un cambio de patrón tras la subida continua de la prima salarial de estudiar Universidad tras los años 80's y 90's en EE. UU. Valleta (2018) documentó con una muestra de trabajadores americanos entre 25 y 64 que en EE. UU. hubo un estancamiento de la prima salarial entre las personas con título universitario y las que poseen un Grado de Bachillerato entre 2000 y 2015. En su artículo y la posterior discusión de Autor publicada en el propio artículo Valleta (2018) se perfilan dos explicaciones:

1) La sustitución de graduados universitarios que realizaban tareas cognitivas rutinarias por máquinas sin tener opción de recolocarse en otro tipo de ocupaciones más abstractas al haberse reducido la demanda de estas tareas.

2) Un fuerte incremento del número de graduados en las universidades americanas a partir de 2004.

Para entender el argumento tecnológico, hay que partir de Acemoglu y Autor (2011) que cambia el foco del análisis de los diferenciales salariales entre colectivos al incorporar al marco conceptual tradicional no sólo diferencias educativas sino una heterogeneidad de ocupaciones que requieren tareas rutinarias y no-rutinarias. Este marco sirvió para explicar un cambio en el empleo que empezó a verse en los 90's y que no era monótono a lo largo de toda la distribución de habilidades. Concretamente, durante los años anteriores a los 90's EE. UU. presentó un crecimiento elevado de los salarios de la parte alta de la distribución. Desde entonces se observó que las ocupaciones de baja cualificación como la asistencia o el cuidado personal iban aumentando su incidencia en el empleo total. Esto se unió al crecimiento de las ocupaciones de elevada cualificación como los responsables y profesionales. Sin embargo, el peso en el empleo de las ocupaciones de educación media como vendedores, administrativos, operarios y peones empezó a disminuir. Autor y Dorn (2013) relaciona la polarización con la interacción entre las preferencias de los consumidores y un coste decreciente de automatizar ciertas tareas laborales rutinarias y codificables. Esta polarización y rutinización se fue encontrando en otros contextos y lugares. Goos et al. (2014) identifica el mismo patrón para 16 economías europeas, entre las que se encuentra España. Igualmente, muestran la importancia de la deslocalización de las tareas de cualificación intermedia tras un proceso de globalización influido también por el cambio tecnológico.

La crisis financiera parece ahondar en esta tendencia. En particular, varios autores concluyeron que esta pérdida de empleo rutinario se incrementó tras la crisis financiera, momento en el que el coste de oportunidad de invertir en nuevas tecnologías y el coste de despedir era menor. Por ejemplo, Hershbein y Kahn (2018) muestra que en las áreas metropolitanas que fueron más duramente golpeadas por la crisis financiera existió una mayor sustitución de empleo por tecnología en las ocupaciones rutinarias manuales. También en estas zonas habría habido un aumento de los requisitos educativos en las ofertas de empleo de ocupaciones rutinarias cognitivas. Sin embargo, los universitarios que perdieron el trabajo rutinario cognitivo por la automatización de sus tareas no pudieron acceder a puestos de mayor abstracción a partir del

año 2000 por una caída de la demanda de este tipo de ocupaciones. En este sentido, Beaudry et al. (2016) sugiere que, a partir del año 2000, la demanda de conocimientos abstractos se redujo al haber perdido fuerza la inversión tecnológica en EE. UU. Su única opción, en consecuencia, fue optar a una recolocación hacia trabajos manuales poco cualificados y de menor salario.

El fuerte aumento del número de graduados a partir del 2000 también tendría como consecuencia una reducción de los diferenciales salariales entre grados educativos incluso en el caso de no existir heterogeneidad en los rendimientos. Este proceso pudo deberse a varias causas. En primer lugar, puede ser que la opción de realizar estudios de postgrado en la Universidad haya aumentado los incentivos exante de acabar un Grado universitario. En este sentido, Lindley y Machin (2016) muestran que el estancamiento de la prima salarial se circunscribió a los universitarios que tan sólo acabaron la carrera y no a los graduados de un título post-obligatorio. Asimismo, Denning et al. (2022) documentan que la graduación universitaria creció de forma continua entre 1990 y 2010 dado un proceso de inflación de notas. Otras explicaciones tradicionales como la preparación de los alumnos, el trabajo durante la carrera, el coste de entrada o los recursos de la Universidad no tuvieron ningún poder explicativo. Esta mayor graduación por inflación de notas haría que la calidad media del colectivo de graduados empeorara y de esa forma el diferencial con otros grados educativos bajaría.

## 4. ¿FP O UNIVERSIDAD?

### 4.1. Marco teórico y evidencia para otros países

Como se deriva del modelo de la sección 3.1., las habilidades que se adquieren en la línea profesional no tienen por qué ser iguales a las que se aprenden en la línea académica, por lo que los ingresos laborales de un mismo número de años de educación pueden ser distintos según la vía escogida. Krueger y Kumar (2004) sugieren que la educación académica facilita a los trabajadores operar con tecnologías de producción nuevas, mientras que los estudios profesionales están enfocados a enseñar a operar tecnologías establecidas. Los autores utilizan esta idea para subrayar que las políticas de formación profesional en Europa facilitaron el crecimiento relativo a EE. UU. durante los años 60's y 70's, cuando

la tecnología era estable. Sin embargo, en la actualidad, con un contexto más cambiante, lo podrían estar perjudicando.

A partir de esta idea, Hanushek et al. (2015) sugiere que puede existir un *trade off* para la formación profesional entre el corto y el largo plazo. Así, las habilidades adquiridas en los cursos de formación profesional podrían facilitar la transición al mercado laboral en el corto plazo, pero podrían quedar obsoletas a un ritmo más rápido que las académicas en el largo plazo.

Es evidente que esta cuestión empírica requiere un marco similar al de la sección 3.2 pero 1) teniendo en cuenta explícitamente diferencias en las ganancias salariales a lo largo de la vida por tipo de formación, y 2) distinguiendo más allá de los años de educación, el tipo de formación que uno realiza. Un buen punto de partida para entender cómo la inclusión del tipo de currículum afecta a las estimaciones de las ecuaciones salariales es Altonji et al. (2012). Estos autores plantean que no sólo es importante la forma en que los conocimientos cambian los ingresos laborales sino también los incentivos a estudiar más años. De esta forma, si $C$ es un tipo de currículum (académico o profesional), los años de educación $S$ se verán afectados por la vía que finalmente se elija y por otras variables observables ($z$) y no observables ($v$):

$$S(C,z) = \beta_0 + \gamma_0 C + \beta_1 z + v$$

Además, los ingresos laborales expost también se verán afectados por el tipo de habilidades adquiridas en cada currículum y los años de educación que finalmente se hayan realizado en el sistema:

$$\ln\big(W(C,S,x)\big) = \alpha_0 + \gamma_1 C + \rho S + \alpha_1 (C,S)x + \alpha_2 (C,S)x^2 + \varepsilon$$

No es evidente si la ecuación salarial debe tener en cuenta de forma independiente a la vía escogida el número de años de educación. Tener en cuenta esta variable en la regresión implica aislar el impacto en los salarios de haber realizado un tipo de currículo educativo, más allá del número de años de educación que se han cursado. Si no se tiene en cuenta, el impacto incorpora la capacidad del currículum para afectar los años de educación. En cualquier caso, el problema principal de este análisis proviene de la dificultad para tratar los sesgos que pudieran derivarse de la auto-selección de estudiantes entre formación profesional y académica.

Hanushek et al. (2015) intenta evitar el problema de auto-selección proponiendo un análisis de diferencia en diferencias. El trabajo utiliza una sección cruzada de la Encuesta Internacional de Alfabetización de Adultos (IALS) con información para 11 países comparando los resultados laborales de personas de diferentes edades con educación académica o profesional. Bajo el supuesto de que el sesgo de auto-selección no varía en el tiempo ni a lo largo de la vida del individuo, el enfoque de diferencia en diferencias lo elimina. Por ejemplo, al restar la diferencia salarial entre formación profesional y académica a los 50 años y la diferencia a los 25 años se eliminaría el sesgo de selección, que se asume igual por generaciones, quedando tan sólo las diferencias en el perfil a lo largo de la vida de cada currículum. De esta forma, encuentran que las personas con formación académica tienen inicialmente peores resultados laborales que las personas con formación profesional si bien experimentan una mejora a medida que envejecen. Estos mismos resultados son confirmados por Hampf y Woessmann (2017) utilizando una base de datos diferente (PIAAC) que les permite incluir más controles de conocimientos cognitivos individuales. Los autores también utilizan la misma estrategia de identificación de diferencias en diferencias por edad. Los resultados parecen más robustos en países con un fuerte sistema de prácticas profesionales (Austria, Dinamarca y Alemania).

Estos resultados han sido corroborados con datos longitudinales que permiten seguir a lo largo del tiempo a la misma generación. Brunello y Rocco (2017) sigue a estudiantes nacidos en Reino Unido en 1958 entre los 23 y 55 años y a los nacidos en 1970 entre los 26 y 42 años. En términos de empleo, para los que tienen formación profesional superior encuentran una ventaja al principio de la carrera laboral que se va reduciendo, si bien esa diferencia nunca se torna negativa. Para los que tienen formación profesional media, no encuentran diferencias con respecto a la formación académica. En términos de salarios, encuentran una pequeña ventaja inicial que se torna desventaja a partir de los 30 años para los que tienen formación profesional media y nacieron en 1958 y para los que tienen formación profesional superior y nacieron en 1970. De forma similar, Golsteyn y Stenberg (2017) identifica las decisiones educativas de las cohortes nacidas en Suecia durante el periodo 1955-1963, así como sus ganancias anuales para cada año desde 1978 hasta 2011. Con estos datos encuentran un diferencial negativo a favor de los estudiantes de formación profesional para los primeros años de carrera que más allá de los 28 años se torna negativo.

El supuesto de que el sesgo de selección entre cohortes no cambia parece bastante restrictivo. Una manera más convincente de eliminar el sesgo de selección proviene de aquellos trabajos que evalúan reformas educativas que cambiaban el porcentaje de educación académica frente a la profesional. En general, los resultados son diferentes a los anteriores y la ventaja a favor de la formación académica en edades más avanzadas es menos clara. Por ejemplo, Malamud y Pop-Eleches (2011) analiza el impacto de una reforma educativa en Rumanía en 1973. Con anterioridad a ese año, el sistema iniciaba la formación profesional tras 8 años de escolaridad. Sin embargo, la reforma exigió recibir 2 años adicionales de estudios académicos. Usando datos del censo y técnicas de discontinuidad, se observa que los estudiantes nacidos a partir de 1959 tuvieron una mayor probabilidad de completar el currículo académico de escuela Secundaria en comparación con sus contrapartes que nacieron inmediatamente antes. Esto es especialmente importante para estudiantes de regiones pobres y rurales y de familias con menos educación, que tenían menos probabilidades de completar la Universidad. Sin embargo, aunque estos estudiantes desfavorecidos se volvieron significativamente más propensos a ser elegibles para la Universidad después de la reforma, esto no se tradujo en un aumento relativo en la graduación universitaria.

De manera similar, Bertrand et al. (2021) estudia una reforma de la formación profesional en Noruega en 1994. Anteriormente, el sistema profesional se caracterizaba por una especialización muy temprana de los alumnos cerrando la puerta a la Universidad. Tras la reforma, se simplificó el número de cursos básicos en los primeros años de formación profesional alcanzando una mayor especialización en los últimos 2 años. Se incrementó la formación académica en esos dos primeros años teniendo que completar un núcleo común en Lengua, Matemáticas, Ciencias Naturales y Ciencias Sociales, que permitía acumular créditos para poder acceder a la Universidad. Los graduados de formación profesional podían volver a la línea académica tras los dos primeros años e incluso podían ser elegibles para la Universidad tras los cuatro años de formación profesional a condición de completar un curso general de 6 meses suplementarios. Esta reforma se aplicó a los nacidos en 1978 con lo que los autores realizan un análisis de discontinuidad. Encuentran que la reforma incrementó la matriculación en formación profesional y en formación académica. En términos de formación alcanzada existieron diferencias por género. La reforma no indujo a cambios en el porcentaje de hombres que acabaron la Universidad, pero sí consiguió que aquellos que no lograron finalizar la formación universitaria, dejaran la rama

académica anticipadamente para finalizar los estudios profesionales. Estos alumnos incrementaron sus ingresos anuales en un 5%. Por otro lado, la reforma incrementó el porcentaje de mujeres que se graduó tanto en formación profesional como en formación académica. Concretamente, hubo más mujeres que hicieron ambos Grados. A pesar del aumento de mujeres elegibles para la Universidad, no se aumentó el porcentaje de mujeres que terminó el Grado universitario ni hubo cambios en sus salarios.

Finalmente, Silliman y Virtanen (2022) utiliza un análisis de discontinuidad motivado por el proceso de admisión en Finlandia. Concretamente se centra en los solicitantes que se postulan tanto para carreras profesionales como académicas y cuya admisión se determina por el exceso de solicitudes en alguna escuela. La ventaja de esta identificación a partir del exceso de demanda es que no se utiliza una reforma que puede cambiar el contexto. La estimación sugiere que la matriculación en formación profesional aumenta los ingresos anuales respecto a la formación académica y el diferencial persiste hasta los 33 años (un 6%). Los autores incorporan una idea interesante que es la de rendimientos heterogéneos según el tipo de estudiante. Esto ayudaría a explicar los resultados diferentes en distintos contextos y reformas. En particular, observan que sus resultados aplican a estudiantes en el rango medio de habilidad, mientras que matricularse en formación profesional para los estudiantes de más habilidad supondría una pérdida de ingresos a largo plazo.

### 4.2. Evidencia descriptiva para España

No tenemos conocimiento de ningún trabajo para España que haya comparado de forma causal las diferencias entre los ingresos a lo largo de la vida laboral de realizar estudios superiores de formación profesional con estudios universitarios. Sin embargo, se pueden utilizar los datos de la EFF para realizar un análisis descriptivo de esta comparativa. En la sección 3.2.2. se describían los diferenciales del ingreso anual entre asalariados con ESO, FP Media y Bachillerato, FP Superior y carrera universitaria asumiendo que cada nivel educativo producía un incremento del salario paralelo a lo largo de la vida laboral. Una forma sencilla de evitar este supuesto es incorporar en la especificación empírica tradicional de Mincer dos interacciones entre la educación, la experiencia y la experiencia al cuadrado.

La tabla 2 muestra que, en general, se rechaza el paralelismo de los ingresos laborales a lo largo de la vida para los diferentes niveles educativos. En primer

lugar, tanto antes como después de la crisis financiera, los ingresos de los asalariados con FP Media y Bachillerato aumentaron a lo largo de la vida más que los de los asalariados con ESO. Concretamente, cada año de experiencia aumentó el diferencial salarial entre ambos grupos en 1pp para hombres y entre 3 y 5pp para mujeres.

**Tabla 2. Comprobación del paralelismo a lo largo de la vida laboral de los diferenciales de ingresos laborales entre niveles educativos.**

| | Hombres | | Mujeres | |
|---|---|---|---|---|
| | <2008 | >=2008 | <2008 | >=2008 |
| **Bachillerato o FP Media vs ESO** | | | | |
| Experiencia | 0.011*** | 0.013** | 0.052*** | 0.032*** |
| | (0.006) | (0.008) | (0.001) | (0.001) |
| Experiencia 2 | –0.001*** | –0.002 | –0.001*** | –0.001*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| **FP Superior vs Bachillerato o FP Media** | | | | |
| Experiencia | 0,020*** | –0.018*** | 0.005*** | –0.000 |
| | (0.002) | (0.001) | (0.011) | (0.001) |
| Experiencia 2 | –0.001*** | –0.000 | –0.000 | –0.000*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| **Universidad vs Bachillerato o FP Media** | | | | |
| Experiencia | 0.038*** | 0.013*** | 0.041*** | 0.005*** |
| | (0.007) | (0.008) | (0.002) | (0.001) |
| Experiencia 2 | –0.001* | –0.000*** | –0.001*** | –0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |

Fuente: Encuesta financiera de las familias 2001-2019 (Banco de España).
Nota: Se realizan regresiones por periodo y género del logaritmo del ingreso laboral anual (incluyendo salario en especie) sobre una variable ficticia correspondiente al nivel educativo superior, experiencia potencial, experiencia potencial al cuadrado y sus interacciones con el nivel educativo. En la tabla se presentan los coeficientes de las interacciones entre experiencia y nivel educativo superior. La regresión incluye los colectivos pertenecientes a los dos grupos educativos que se comparan. Se limita la muestra a personas entre 18 y 64 años. La experiencia potencial es edad-16 años para asalariados con ESO, edad-18 para Bachillerato o FP Media, edad-20 para FP Superior y edad-22 para universitarios. Se hace la media de las cinco imputaciones. En paréntesis la desviación estándar.
*** Significativo al 1%, ** al 5% y * al 10%.

Por otro lado, no se observan grandes diferencias en el perfil de los asalariados de titulación media con los que tienen los titulados de FP Superior. Si bien, la carrera laboral era algo más dinámica para los graduados de FP Superior que para los graduados medios con anterioridad a la crisis, especialmente para los hombres, no se observan diferencias positivas significativas con posterioridad a ella.

Finalmente, los universitarios presentaron un perfil salarial con mayor pendiente a lo largo de la vida que los asalariados con educación media. Esta mayor pendiente parece haberse reducido tras la crisis financiera pero aún es positiva y significativa estadísticamente. De esta forma, un año más de experiencia con anterioridad a la crisis, aumentaba el diferencial salarial con respecto a los trabajadores con educación media en 4pp tanto para hombres como para mujeres. Posteriormente, aumentó en 1pp y 0,5pp respectivamente por año trabajado. Esta evidencia sugiere un mayor perfil salarial para los universitarios respecto a los graduados medios y a los graduados de FP Superior lo que sería coherente con la evidencia anteriormente presentada en otros países para este tipo de datos. El ejercicio 8 propuesto debería servir para pensar sobre la forma en que este fenómeno hace variar las estimaciones de TIR presentadas anteriormente.

Para poder identificar cambios en la calidad de las diferentes generaciones habría que seguir en el tiempo a diferentes generaciones de graduados universitarios y de FP Superior. En este sentido, sería interesante corroborar los resultados anteriores basados en muestras de secciones cruzadas repetidas a partir de datos longitudinales. En particular, esto se podría hacer combinando los datos de inserción laboral del Ministerio de Universidades y los datos de inserción del Ministerio de Educación y Formación Profesional y alargando las series más allá de los cuatro primeros años tras la graduación.

Ejercicio 8. *Compara la TIR que sale de la regresión de Mincer sin interacciones del ejercicio 6 y una regresión con interacciones aplicada a los mismos datos. En el caso de la regresión con interacciones los coeficientes son: 8,6; un coeficiente de estudios medios de 0,1; un coeficiente de la experiencia de 0,01; un coeficiente de −0,0004 para su cuadrado; un coeficiente de 0,04 para la interacción de educación y experiencia y de −0,0005 para la interacción con su cuadrado. ¿Puedes intuir porqué existe esa diferencia?*

Ejercicio 9. *Basándote en tu propia experiencia, m*ás allá de los rendimientos *expost a lo largo de la vida, discute qué otras razones crees que pueden influir en la*

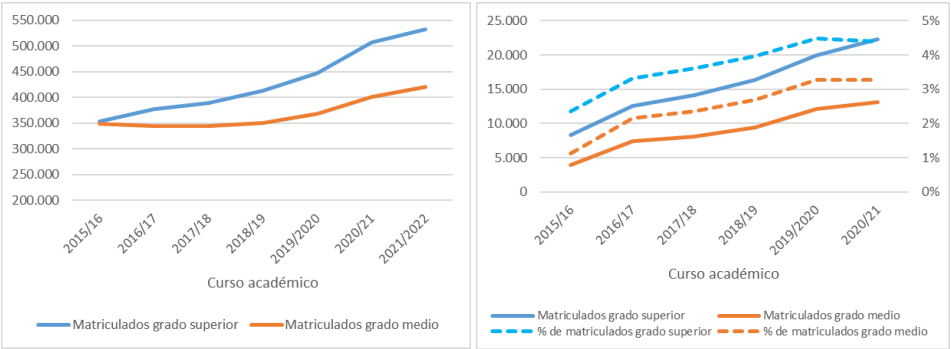
*decisión de cursar Universidad o FP Superior. Piensa en la forma y los datos que necesitarías para contrastar estos mecanismos.*

## 5. ¿FP TRADICIONAL O DUAL?

Como se ha visto en la sección 2, en los últimos años, se ha producido un aumento considerable del número de matriculados en la FP en España, especialmente concentrado en los ciclos de Grado superior. El gráfico 6 muestra la evolución del número total de matriculados en los ciclos de FP de Grado medio y superior. En el caso de los ciclos formativos de Grado superior se ha producido un incremento de casi el 50% en el número de matriculados desde el curso 2015-2016, frente a una tasa de crecimiento del 20% para los ciclos formativos de Grado medio. De acuerdo a lo anterior, una parte importante del crecimiento de la FP Superior se debe a una mayor entrada de graduados desde la FP Media, si bien también hubo un incremento del porcentaje de personas que decidieron realizar estos estudios tras el Bachillerato. El reciente auge del rendimiento de la formación profesional y el posible estancamiento del de la formación universitaria hacen suponer que en los próximos años el crecimiento relativo de la demanda de formación profesional se podría mantener.

**Gráfico 6. Matriculados en ciclos de FP (por curso académico).**

**6.1. Total.** **6.2. FP Dual.**




Fuente: Ministerio de Educación y Formación Profesional..

Esta evolución de la demanda genera un reto para la oferta[10]. Para responder al reto, el componente 20 del Plan de Recuperación, Transformación y Resilien-

[10] Véase Auciello, Segú y Lacuesta (2021) para ilustrar alguno de estos retos.

cia que el Gobierno español presentó el 30 de abril de 2021 a las autoridades europeas en el marco del programa Next Generation EU (NGEU) contempla una inversión de 254 millones de euros en los próximos años con el propósito de aumentar en 200.000 plazas la oferta de formación profesional en nuestro país. Asimismo, la Ley Orgánica de Ordenación e Integración de la Formación Profesional aprobada en 2022 pretende una transformación global del sistema de formación profesional en nuestro país para adaptarlo a este nuevo contexto.

En particular, la ley apuesta decididamente por la modalidad dual introducida en 2012. La principal característica de la FP dual es el protagonismo de las empresas en la formación. La normativa que estuvo vigente hasta la nueva ley establecía que al menos un tercio de todas las actividades formativas tenía que contar con la participación del centro de trabajo adscrito, preferiblemente en alternancia con la formación en el centro. En comparación, los alumnos en la FP tradicional –denominado FP presencial en adelante– solo disfrutaban de prácticas durante tres meses al final del segundo y último curso. Con la nueva ley, todos los ciclos de Grado medio y superiores se ofrecerán únicamente en la modalidad dual. En el caso de la FP dual general, los alumnos pasarán entre el 25% y el 35% de todas las horas en su empresa de prácticas, y ésta tiene que responsabilizarse de un mínimo del 20% de todas las actividades formativas. Por el contrario, en el caso de la FP dual intensiva la formación en la empresa debe ser superior al 35% de la duración de los estudios y las empresas tienen que colaborar en un mínimo del 30% de todas las actividades formativas. Además, en esta modalidad las empresas están obligadas a contratar a los aprendices mediante un contrato formativo.

La creación de la modalidad dual persiguió varias finalidades: reducir la tasa de abandono temprano, mejorar la motivación de los alumnos y fomentar su inserción laboral, además de mejorar el vínculo entre la oferta de formación y las necesidades de las empresas. La apuesta por la FP dual fue una respuesta a una recomendación de la Comisión Europea y está avalada por la buena experiencia en los países centroeuropeos, donde la FP dual es la principal vía de acceso al mercado laboral. Alemania, Suiza y Austria no solo están entre los países con la tasa de desempleo juvenil más bajas de Europa. También figuran entre los países con las tasas de abandono prematuro más bajas. Sin embargo, la implantación a gran escala de la FP dual es un reto y no está exenta de riesgos.

El gráfico 6.2. muestra la evolución del número de alumnos matriculados en los ciclos de FP dual de Grado medio y superior y su respectivo peso en el

total de alumnos matriculados. Como puede observarse, la implantación de la FP dual es mayor, tanto en términos absolutos como relativos, en los ciclos de Grado superior, pero en ambos casos su peso no alcanza el 5%. Además, los datos del ministerio dejan patente la desigual implantación de la modalidad dual entre regiones. Según los últimos datos disponibles relativos al curso 2020-2021, Andalucía, Cataluña y Madrid representan, en conjunto, el 50% de todos los matriculados en ciclos de FP dual de Grado medio y el 70% superior. Además, en términos relativos, destaca el caso del País Vasco, donde la FP dual representa el 16,7% de los matriculados en ciclos superiores. Más allá de las diferencias en el Grado de implantación de la FP dual, cabe destacar también la variedad de modelos. Mientras que el País Vasco optó por un modelo en que las empresas contratan a los aprendices mediante un contrato para la formación y el aprendizaje, otras regiones, como Madrid, optaron por un modelo en el que los aprendices son becarios.

Esta sección pretende ofrecer un breve repaso de las ventajas y desventajas de la FP dual y la evidencia sobre el impacto de la FP dual en la inserción laboral comparado con la FP tradicional o presencial.

### 5.1. Marco teórico

Las ventajas de la formación en alternancia están ampliamente documentadas[11]. La alternancia permite a los alumnos aplicar sus conocimientos y adquirir habilidades y competencias pertinentes para su ocupación, así como hábitos de trabajo y otras habilidades interpersonales, en un entorno profesional. Además, la combinación de estudio y trabajo puede ayudar a mejorar la motivación de los alumnos más rezagados y/o en riesgo de abandono. Las empresas, por su parte, tienen la oportunidad de observar el desempeño de los aprendices durante un periodo de tiempo relativamente largo. Este mayor conocimiento puede favorecer la contratación (o retención) de los aprendices por su empresa de formación.

Sin embargo, estas ventajas deben contraponerse con las desventajas de un menor número de horas de formación en el centro de formación. La sustitución de la formación en el aula por la formación en la empresa puede erosionar la

[11] Para una revisión de la literatura sobre el impacto de distintos sistemas de formación profesional y la FP dual en particular, vean Wolters y Ryan (2011), Ryan (1998, 2001) o European Commission (2013).

calidad de la formación si los tutores de empresa están menos capacitados que los profesores de los centros. Por otra parte, podría producirse un desequilibrio entre las competencias generales y específicas, ya que las empresas podrían verse tentadas a hacer que la formación fuera demasiado específica para la empresa.

De hecho, los incentivos de las empresas son un elemento clave que las autoridades deben tomar en consideración a la hora de diseñar un sistema de FP dual. La teoría del capital humano muestra las dificultades para las empresas de rentabilizar la inversión en capital humano generalista. En concreto, en un mercado laboral de competencia perfecta la inversión óptima por parte de las empresas en capital humano generalista sería nulo. Para que las empresas tengan incentivos para participar en la FP dual y fomentar la adquisición de capital humano generalista tiene que existir algún tipo de mecanismo que permita recuperar la inversión. Una posibilidad es que las empresas recuperen el coste a posteriori pagando salarios por debajo de la productividad. La otra opción es que los aprendices financien su propia formación o que participen en actividades que generan valor para la empresa durante su formación.

A nivel internacional hay muchos modelos distintos. En Alemania las empresas alemanas suelen hacer una inversión neta en los aprendices durante su formación, que luego recuperan a través de salarios de convenio moderados. Por el contrario, en Suiza los aprendices suelen aportar valor a la empresa durante su formación por su implicación en tareas productivas y la menor dedicación de horas a la formación. El modelo alemán requiere tasas de retención relativamente altas y coordinación entre las empresas del mismo sector para evitar el "robo" de trabajadores cualificados por parte de empresas que no forman a sus propios empleados. Por el contrario, en el modelo suizo, los aprendices suelen moverse a otras empresas cuando acaban su formación. Por último, según la evidencia disponible, el beneficio neto de la formación no solo varía entre países. También se observan considerables diferencias por sectores, ocupaciones y en función del tamaño de la empresa.

### 5.2. Evidencia del impacto en la carrera laboral de la FP tradicional o dual

#### *Evidencia descriptiva*

Existe un amplio consenso en la literatura sobre el hecho de que los programas de FP dual mejoran la transición de la educación al trabajo. Al inicio de

sus careras laborales, los graduados de la FP dual suelen tener mayores tasas de empleo y puestos de trabajo de más calidad que sus homólogos en la FP presencial. En general, las diferencias en tasas de retención solo explican una parte de las diferencias, y con el tiempo, estas tienden a atenuarse a medida que los graduados de la modalidad presencial recuperan parte de la desventaja acumulada en los inicios de su carrera laboral. Por el contrario, en lo que se refiere al efecto sobre los salarios, los resultados son ambiguos, tanto a nivel teórico como a nivel empírico por las razones expuestas anteriormente.

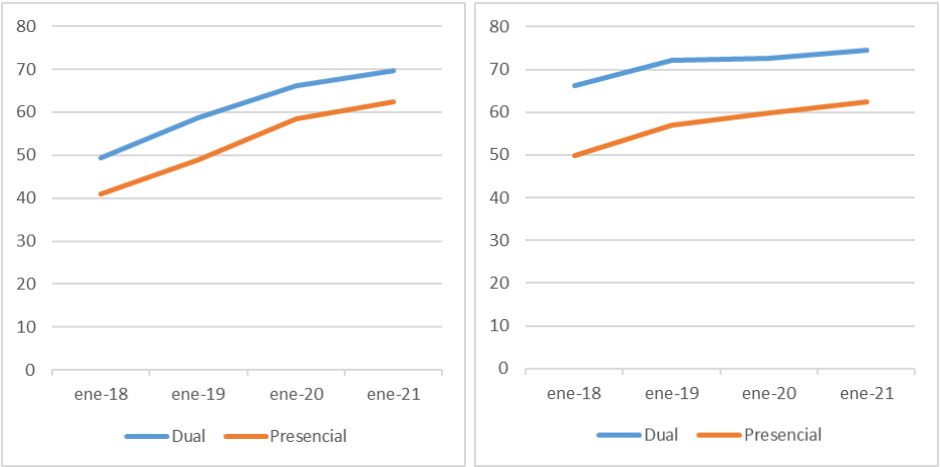
La evidencia internacional sobre el impacto de la FP en los resultados laborales es muy abundante. Sin embargo, no se considera concluyente por la falta de evidencia causal. Esta falta de evidencia causal se debe a una combinación de problemas de datos y de identificación. En concreto, en países con sistemas consolidados de FP dual resulta difícil construir un contrafactual convincente y en el resto de los países no es siempre fácil resolver los problemas de selección que se producen cuando los alumnos pueden elegir entre matricularse en la FP dual o la FP presencial. Esta elección puede depender de características no observables de los alumnos que también ejercen un efecto directo en su futuro desempeño laboral, como puede ser la motivación de los alumnos o las preferencias por una entrada temprana al mercado laboral. Además, otro factor que limita la relevancia de la evidencia internacional para el caso de España es la prevalencia de estudios para países centroeuropeos. Sus sistemas de FP dual están integrados en un marco institucional difícil de reproducir en otros países.

En el caso de España, la evidencia es escasa, pero alentadora. La principal fuente de evidencia descriptiva son las tasas de inserción que publica el Ministerio de Educación y Formación Profesional en colaboración con la Tesorería General de la Seguridad Social. Desde el año pasado, el ministerio ofrece datos desglosados para los ciclos de FP dual. El gráfico 7, muestra los datos nacionales para la cohorte 2016-2017. La tasa de inserción está definida como el porcentaje de graduados que han trabajado al menos un día en los meses de marzo en los cuatro años siguientes a su graduación. En línea con la evidencia internacional, se observa que los graduados de la FP dual obtienen mejores resultados, en promedio, que sus homólogos de la FP presencial. La diferencia en las tasas de inserción es algo mayor en el caso de los ciclos de Grado superior y bastante estable en el tiempo. Tras cuatro años, la diferencia asciende a 12 pp en el caso de los ciclos de Grado superior y 7pp en el caso de los ciclos de Grado medio. El Ministerio

también ofrece datos desagregados al nivel de CC.AA. con resultados similares para 11 de las 13 CC.AA. para las que se dispone de datos para la FP dual.

**Gráfico 7. Tasa de inserción de la generación graduada 2016-2017, media nacional (por año).**

**7.1. Grado medio.** **7.2. Grado superior.**




Fuente: Ministerio de Educación y Formación Profesional.

### *Evidencia causal*

La única evidencia causal para España proviene de Bentolila, Cabrales y Jansen (2023). Los autores tienen acceso a los expedientes académicos y los historiales laborales completos de las tres primeras cohortes de graduados en la FP dual en Madrid, todos de Grado superior, y compara sus resultados laborales con un grupo de control compuesto de individuos que se han graduados en los mismos años, centros y ciclos en la modalidad presencial.

Antes de presentar las conclusiones del estudio, reproducimos parte de su tabla 4 con las estadísticas descriptivas que resume los resultados laborales de ambos conjuntos de graduados (véase tabla 3).

Los datos muestran que los graduados de la FP dual obtienen mucho mejores resultados laborales que sus homólogos de la modalidad presencial. Consiguen acumular más días de trabajo y más renta y obtienen muchos contratos. La magnitud de las diferencias es llamativa. Por ejemplo, la diferencia en la media de los días de trabajo equivalentes al tiempo completo es de 104 días, equivalente a una mejora del orden del 35% comparada con el resultado medio de los gradua-

**Tabla 3. Resultados laborales FP dual y presencial durante los primeros dos años.**

| | Dual | Presencial | Diferencia | *p* |
|---|---|---|---|---|
| Días de empleo | 462,98 | 381,82 | 81,16 | 0 |
| Días de empleo ETC: | | | | |
| *Media* | 405,38 | 300,53 | 104,85 | 0 |
| *Mediana* | 437 | 246,22 | 190,78 | 0 |
| Duración de contratos | 147,48 | 123,95 | 23,53 | 0 |
| Empleo regular (%) | 81,74 | 75,36 | 6,38 | 0 |
| Empleo a tiempo parcial (%) | 49,23 | 61.88 | –12,65 | 0 |
| Contrato indefinido (%) | 39,38 | 32,63 | 6,75 | 0 |
| Tasa de retención (%) | 19,29 | 1,22 | 18,07 | 0 |
| Renta laboral: | | | | |
| *Media* | 16.144 | 11.063 | 5.081 | 0 |
| *Mediana* | 15.461 | 8.323 | 7.138 | 0 |
| Renta total | 17.226 | 11.931 | 5.294 | 0 |
| Observaciones | 975 | 6291 | 7.266 | |

Fuente: Bentolila, Cabrales y Jansen (2023).
Nota: Graduados en la Comunidad de Madrid durante el periodo 2014-2016 en los ciclos de Grado superior con oferta de FP dual. ETC es Equivalente a Tiempo Completo. Empleo regular se refiere a contratos con una duración mínima de tres meses. La tasa de retención es el porcentaje de graduados que obtienen un contrato con su empresa de prácticas en los doce meses siguientes a la graduación.

dos en la FP presencial. Además, la diferencia en la mediana de días de trabajo ETC es incluso mayor (190 días equivalentes a 77% de la mediana para la FP presencial). Para la renta se observa un patrón similar. En este caso se amplían las diferencias por la menor incidencia del empleo a tiempo parcial entre los graduados de la FP dual.

Otros dos aspectos llamativos son las diferencias en la tasa de retención y el porcentaje de graduados que ha obtenido un contrato indefinido. La tasa de retención en la FP dual es relativamente baja en comparación con otros países como Alemania o Francia con tasas de 60% y 40%, respectivamente, pero muy superior a la irrisoria tasa de retención en la FP presencial. Además, cuatro de cada diez graduados en la FP dual obtuvieron un contrato indefinido en los primeros dos años frente a un tercio de los graduados en la modalidad presencial.

Se trata de buenas cifras en un país donde casi dos tercios de los menores de 25 años suelen ocupar un contrato temporal.

El objetivo de Bentolila et al. (2023) es determinar que parte de estas diferencias tiene una interpretación causal. Su estrategia de identificación está basada en una variable instrumental que mide la diferencia en el tiempo de traslado de los alumnos hacia los centros de FP más cercano con el ciclo de su elección en ambas modalidades. Además, el estudio incluye un test de placebo para descartar la posibilidad que la oferta de la FP dual refleje diferencias anteriores a la reforma de 2012 en la calidad de la formación. Los resultados principales del estudio muestran que la graduación en la modalidad dual genera aumentos altamente significativos en el número de días trabajados, la renta acumulada y la tasa de retención. Por el contrario, no existe ningún impacto significativo en las variables que miden la calidad del empleo.

La magnitud de los incrementos en los días de trabajo y la renta es similar a la magnitud de las diferencias observadas en los datos brutos. Aproximadamente, dos tercios de estas diferencias se producen durante los primeros doce meses, y a nivel de cohorte los autores observan un aumento tendencial en los retornos de la FP dual frente a la FP presencial. Por último, contrariamente a la creencia popular, los mayores incrementos no se producen en los ciclos vinculados a los sectores manufactureros, sino en ciclos donde los alumnos obtenían resultados por debajo de la media en el periodo pre-reforma. A nivel individual, esta heterogeneidad por ciclos explica porque la graduación en la modalidad dual aumenta significativamente la posibilidad de alcanzar resultados laborales en la parte alta de la distribución por encima del cuarto decil.

Debido a las diferencias en los modelos de FP dual entre regiones no se pueden extrapolar los resultados a otras CC.AA., pero en un trabajo en curso varios de los autores encuentran resultados similares para el caso de Cataluña, aunque en esta Comunidad Autónoma la tasa de retención es el doble. Uno de los retos para la investigación es analizar si esta diferencia está relacionada con las distintas reglas para la alternancia. A diferencia de Cataluña, la Comunidad de Madrid ha convergido a un modelo donde los alumnos pasan el primer año entero en el centro de formación y el segundo en la empresa de prácticas. La ausencia de alternancia puede generar incentivos para la rotación de aprendices cuando estos generan un beneficio a las empresas.

Ejercicio 10. *Dados los argumentos de la sección 4, la ventaja de la FP dual podría desaparecer en el largo plazo. Una vía para que esto no suceda es la formación*

*continua de los trabajadores, elemento que no se ha tratado en este capítulo. Describe la situación de la formación en empresas de España en comparativa internacional y discute cuáles pueden ser los retos de este tipo de formación en particular para la economía española. (Pista: existen datos sobre este tema en Eurostat "Education and Training").*

## 6. CONCLUSIONES

El capítulo pretende mostrar algunos artículos teóricos y empíricos que pueden servir para ilustrar la decisión de seguir estudiando una vez concluida la etapa obligatoria. Para dotar el capítulo de contexto se ha recopilado algo de información sobre esta decisión durante los últimos treinta años en la economía española. En particular, se ha mostrado que las generaciones nacidas entre 1978 y 1990 experimentaron un estancamiento de su formación dadas las oportunidades laborales derivadas del boom inmobiliario para los menos formados. Con la crisis financiera, las generaciones nacidas tras 1990, volvieron a aumentar el porcentaje de población con educación superior. Finalmente, las generaciones nacidas a partir del 2000, en un entorno de recuperación económica tras la crisis financiera, aumentaron el porcentaje de población que alcanzó un grado de formación profesional en relación al de graduados universitarios.

El resto del capítulo analiza las razones que explicarían el anterior comportamiento de la decisión de seguir estudiando tras la finalización de la enseñanza obligatoria, documenta las diferencias en ingresos laborales de los graduados de formación profesional y formación académica y documenta las diferencias en ingresos laborales de los graduados en formación profesional dual o tradicional.

En relación a primer punto, la sección tres realiza un análisis con datos de la Encuesta Financiera de las Familias entre 2001 y 2019, y muestra un incremento del rendimiento de realizar estudios medios respecto a la ESO y de realizar estudios superiores respecto a los medios tras la crisis financiera. A partir de 2013, parece existir un estancamiento en los rendimientos expost de realizar estudios universitarios, especialmente entre los hombres, mientras que los rendimientos de la FP Superior aumentaron de forma considerable. Una limitación del análisis descriptivo presentado en la sección es la imposibilidad de tratar los sesgos derivados de la autoselección en diferentes grados de formación y la heterogeneidad de rendimientos en la población. De corroborarse, los resultados anteriores, en los próximos años, el crecimiento relativo de la demanda de formación profesio-

nal en relación a la Universidad se podría mantener. Esto supone un reto para la organización y redimensionamiento del sistema de formación profesional, especialmente en ámbitos que requieran una fuerte inversión en infraestructuras y la contratación de personal cualificado adicional.

La sección 4 trata la cuestión de si la ventaja que pudiera tener la formación profesional respecto a la formación académica en la inserción laboral al inicio de la carrera desaparece al cabo de unos años. La evidencia descriptiva presentada con datos de la EFF parece corroborar estos resultados, aunque en los últimos años no parece existir gran diferencia entre ambas opciones a lo largo de la vida laboral. En el caso de España, sería interesante corroborar los resultados descriptivos basados en muestras de secciones cruzadas repetidas con nuevas series largas longitudinales de inserción laboral de la Seguridad Social. Adicionalmente, la evidencia causal internacional no parece ser tan concluyente y refuerza la idea de la heterogeneidad de rendimientos de formación profesional y formación académica según el perfil de habilidad del estudiante.

Finalmente, en relación al diseño del sistema de formación profesional, la evidencia descriptiva presentada en la sección 5 a partir de datos de bases de cotización de la Seguridad Social muestra que al inicio de sus careras laborales, los graduados de la FP dual suelen tener mayores tasas de empleo y puestos de trabajo de más calidad que sus homólogos en la FP presencial. Los trabajos presentados también muestran que las diferencias en días trabajados, renta acumulada y tasa de retención son causales. Estos resultados reforzarían la estrategia del Gobierno que ha decidido generalizar la FP dual con su última reforma. En concreto, a partir del próximo curso, todos los actuales ciclos de Grado medio y superiores se ofrecerán únicamente en la modalidad dual. Sin embargo, esto no significa que la implementación de la FP dual en España no esté exenta de retos.

El éxito de la reforma dependerá de la capacidad de las empresas de generar suficientes plazas de aprendices y de ocuparse de parte de su formación. El segundo aspecto supone todo un reto para las pequeñas empresas que hasta ahora ofertaban plazas de prácticas para alumnos de la FP presencial. En este sentido conviene señalar que el Gobierno todavía no ha regulado la figura de los tutores. Además, tampoco ha aclarado si las PYMES van a recibir ayudas para financiar la formación de los tutores. Sin embargo, sin la presencia de tutores cualificados en una parte considerable de las empresas no está claro cómo se puede garantizar la calidad de la formación en las empresas. Por último, el aumento en el coste

para las empresas puede dificultar la continuidad de algunas de las empresas que participaban en la FP dual hasta ahora.

## BIBLIOGRAFÍA


Acemoglu, D. y Autor, D. (2011). Chapter 12. Skills, Tasks and Technologies: Implications for Employment and Earnings en Card, D. y Ashenfelter, O., *Handbook of Labor Economics.* Elsevier, 4, Part B, 1043-1171. https://doi.org/10.1016/S0169-7218(11)02410-5

Altonji, J. (1993). The Demand for and Return to Education When Education Outcomes are Uncertain. Journal of Labor Economics, 11(1), Part 1: Essays. *Honor of Jacob Mincer*, 48-83. https://www.journals.uchicago.edu/doi/abs/10.1086/298317

Altonji, J., Blom, E. y Meghir, C. (2012). Heterogeneity in Human Capital Investments: High School Curriculum, College Major and Careers. *Annual Review Economics, 4*, 185-223. https://www.annualreviews.org/doi/abs/10.1146/annurev-economics-080511-110908

Anghel, B., de la Rica, S. y Lacuesta, A. (2014). *The impact of the great recession on employment polarization in Spain*. SERIEs, 5, 143-171. https://link.springer.com/article/10.1007/s13209-014-0105-y

Aparicio-Fenoll, A. (2016). Returns to Education and Educational Outcomes: The Case of the Spanish Housing Boom. *Journal of Human Capital, 10*(2), 235-265. https://www.journals.uchicago.edu/doi/abs/10.1086/686154

Arpaia, A. y Halasz, A. (2023). Short- and long-run determinants of labour shortages. *Quarterly Report on the Euro Area*, Vol. 22 (1), European Union. https://economy-finance.ec.europa.eu/system/files/2023-05/ip201_en.pdf

Atkin, D. (2016). Endogenous Skill Acquisition and Export Manufacturing in Mexico, *The American Economic Review, 106*(8), 2046-2085. https://www.jstor.org/stable/43956906

Auciello, I., Lacuesta, A. y Segú, J.Ll. (2021). Oportunidades laborales y desajustes en la oferta y la demanda de plazas en las titulaciones de la formación profesional. Recuadro 6 del Informe Trimestral de la Economía española 2021/T4.

https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/Informe%20trimestral/21/Recuadros/Fich/be2104-it-Rec6.pdf

Autor, D. y Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review, 103*(5), 1553-1597. https://www.aeaweb.org/articles?id=10.1257/aer.103.5.1553

Autor, D., Katz, L. y Kearney, M. (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *The Review of Economics and Statistics, 90*(2), 300-323. https://www.jstor.org/stable/40043148

Beaudry, P., Green, D. y Sand, B. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics, 34*(S1-2), 199-247. https://www.journals.uchicago.edu/doi/epdf/10.1086/682347

Becker, G.S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy, 70*, 9-49. https://www.jstor.org/stable/1829103

Bentolila, S., Cabrales, A. y Jansen, M. (2023). Does Dual Vocational Education and Training Pay Off?, mimeo.

Bertrand, M., Mogstad, M. y Mountjoy, J. (2021). Improving Educational Pathways to Social Mobility: Evidence from Norway's Reform 94. *Journal of Labor Economics, 39*(4), 965-1010. https://www.journals.uchicago.edu/doi/epdf/10.1086/713009

Black, D. McKinnish, T. y Sanders, S. G. (2005). Tight Labor Markets and the Demand for education: Evidence from the Coal Boom and Bust. *Industrial and Labor Relations Review, 59*(1), 3-16. https://www.jstor.org/stable/25063012

Bonhomme, S y Hospido, L. (2017). The Cycle of Earnings Inequality: Evidence from Spanish Social Security Data. *The Economic Journal, 127*(603), 1217-1729. https://onlinelibrary.wiley.com/doi/epdf/10.1111/ecoj.12368

Brunello, G. y Rocco, L. (2017). The Labor Market Effects of Academic and Vocational Education over the Life Cycle. *Journal of Human Capital, 11*(1), 106-166. https://www.journals.uchicago.edu/doi/full/10.1086/690234

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica, 69*(5), 1127-1160. http://links.jstor.org/sici?sici=0012-9682%28200109%2969%3A5%3C1127%3AETRTSP%3E2.0.CO%3B2-Q

Card, D. (2018). Returns to Schooling. *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, London. https://doi.org/10.1057/978-1-349-95189-5_2574

CEDEFOP (2016). *Leaving education early: putting vocational education and training centre stage.* Vol. I. Investigating causes and extent Luxembourg: Publications Office of the European Union, 2016. https://www.cedefop.europa.eu/en/publications/5557

Cerezo, A. y Izquierdo, M. (2023). Encuesta a las empresas españolas sobre la evolución de su actividad: primer trimestre de 2023. Artículo 21 del Boletín Económico 2023/T1 Banco de España. https://repositorio.bde.es/handle/123456789/29809

Cuadrado, P. y Montero J.M. (2022). Flujos migratorios en España durante la crisis del COVID-19 y su incidencia sobre la oferta laboral, Recuadro 5 del Informe Trimestral de la economía española Boletín Económico 2022/T1 Banco de España. https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/Informe%20trimestral/22/Fich/be2201-it-Rec5.pdf

Cunha, F. y Heckman, J. (2007). Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling. *Labour Economics, 14*(6), 870-893. https://doi.org/10.1016/j.labeco.2007.06.002

Denning, J, Eide, E., Mumford, K., Patterson, R. y Warnick, M. (2022). Why Have College Completion Rates Increased? *American Economic Journal: Applied Economics, 14*(3), 1-29. https://pubs.aeaweb.org/doi/pdfplus/10.1257/app.20200525

Dustmann, C., Ludsteck, J. y Schönberg, U. (2009). Revisiting the German Wage Structure, *The Quarterly Journal of Economics, 124*(2), 843-881. https://www.jstor.org/stable/40506245

European Comission (2013). *The effectiveness and cost-benefits of apprenticeships: Results of the quantitative analysis*. DG Employment, Social Affairs and Inclusion
https://ec.europa.eu/social/BlobServlet?docId=11352&langId=en

Fortin N., Green, D., Lemieux, T. y Milligan, K. (2012). Canadian Inequality: Recent Developments and Policy Options. *Canadian Public Policy, 38*(2), 121-145.
http://www.sfu.ca/~pendakur/teaching/econ104/fortin%20et%20al%20CPP%202012.pdf

García-Miralles, E., N. Guhner y R. Ramos (2019). The Spanish personal income tax: facts and parametric estimates. SERIEs, 10, 439-477. https://doi.org/10.1007/s13209-019-0197-5

Golstein, B. y Stenberg, A. (2017). Earnings over the Life Course: General versus Vocational Education. *Journal of Human Capital, 11*(2), 167-212.
https://www.journals.uchicago.edu/doi/full/10.1086/691798

Goos, M., Manning, A., Salomons, A. (2014). Explaining job polarization: routine-biased techno- logical change and offshoring. *American Economic Review, 104*(8), 2509-2526.
https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.104.8.2509

Gosling, A., Machin, S. y Meghir, C. (2000). The Changing Distribution of Male Wages in the U.K. *The Review of Economic Studies, 67*(4), 635-666.
https://www.jstor.org/stable/2695942

Griliches, Z. (1977). Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica, 45,* 1-22.
https://www.jstor.org/stable/1913285

Hampf, F. y Woessmann, L. (2017). Vocational vs. General Education and Employment over the Life Cycle: New Evidence from PIAAC. *CESifo Economic Studies, 63*(3), 255-269.
https://doi.org/10.1093/cesifo/ifx012

Hanushek, E., Schwerdt, G., Woessmann, L., y Zhang, L. (2017). General Education, Vocational Education, and Labor market Outcomes over the Lifecycle. *Journal of Human Resources, 52*(1), 48-87.
https://jhr.uwpress.org/content/52/1/48

Heckman, J., Lochner, L.J., Todd, P.E. (2006). Earnings equations and rates of return: The Mincer equation and beyond, en Hanushek, E.A. y Welch F. (eds.), *Handbook of the Economics of Education*, Capítulo 7, North-Holland, Amsterdam, 307-458.
http://jenni.uchicago.edu/papers/Heckman_Lochner_etal_2006_HEE_v1_ch7.pdf

Hershbein, B. y Kahn, L. (2018). Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings. *American Economic Review, 108*(7), 1737-1772.
https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.20161570

Izquierdo, M. y Lacuesta, A. (2012). The contribution of changes in employment composition and relative returns to the evolution of wage inequality: the case of Spain. *Journal of Population Economics, 25*, 511-543.
https://link.springer.com/article/10.1007/s00148-011-0371-y

Krueger, D. y Kumar, K. (2004). Skill-Specific Rather than General Education: A Reason for US-Europe Growth Differences? *Journal of Economic Growth, 9*(2), 167-207.
https://www.jstor.org/stable/40212697

Lacuesta, A., Puente, S. y Villanueva, E. (2020). The schooling response to a sustained increase in low-skill wages: evidence from Spain 1989-2009. SERIEs, 11, 457-499. https://doi.org/10.1007/s13209-020-00218-0

Lindley, J. y Machin, S. (2016). The rising postgraduate wage premium. *Economica, 83*, 281-306.
https://www.jstor.org/stable/24751920

Malamud, O. y Pop-Eleches, C. (2011). School tracking and access to higher education among disadvantaged groups. *Journal of Public Economics, 95*(11-12), 1538-1549.
https://doi.org/10.1016/j.jpubeco.2011.03.006

Mincer, J. (1974). *Schooling, experience and earnings*. New York: NBER (distributed by Columbia University Press).
https://books.google.es/books/about/Schooling_experience_and_earnings.html?hl=es&id=9IGqAAAAIAAJ&redir_esc=y

Pijoan-Mas, J. y Sánchez-Marcos, V. (2010). Spain is different: Falling trends of inequality. *Review of Economic Dynamics, 13*, 154-178. https://www.sciencedirect.com/science/article/pii/S109420250900057X#:~:text=Our%20analysis%20shows%20that%20inequality,to%20understand%20this%20falling%20trend.

Ryan, P. (2001). The School-to-work transition: A cross-national perspective. *Journal of Economic Literature, 39*(1), 34-92. https://www.aeaweb.org/articles?id=10.1257/jel.39.1.34

Ryan, P. (1998). Is Apprenticeship better? A review of the relevant literature', *Journal of Vocational Education and Training, 50*, 289-329. https://www.tandfonline.com/doi/abs/10.1080/13636829800200050

Silliman, M. y Virtanen, H. (2022). Labor Market Returns to Vocational Secondary Education. *American Economic Journal: Applied Economics, 14*(1), 197-224. https://doi.org/10.1257/app.20190782

Wolter, S. y Ryan, P. (2011). Chapter 11: Apprenticeship en Hanushek, R. y Woessman, L., *Handbook of the Economics of Education*, 3, 521-576 https://www.sciencedirect.com/science/article/pii/B9780444534293000119

# Capítulo 4c
# EDUCACIÓN TERCIARIA: ACCESO, FINANCIACIÓN Y RENTABILIDAD

José García Montalvo
Departament d'Economia i Empresa
Universitat Pompeu Fabra
Barcelona School of Economics e IVIE
ORCID: https://orcid.org/0000-0003-4161-5055
Jorge Sainz
Departamento de Economía Aplicada I e Historia e Instituciones Económicas
Universidad Rey Juan Carlos
Institute for Policy Research
University of Bath
ORCID: https://orcid.org/0000-0001-8491-3154

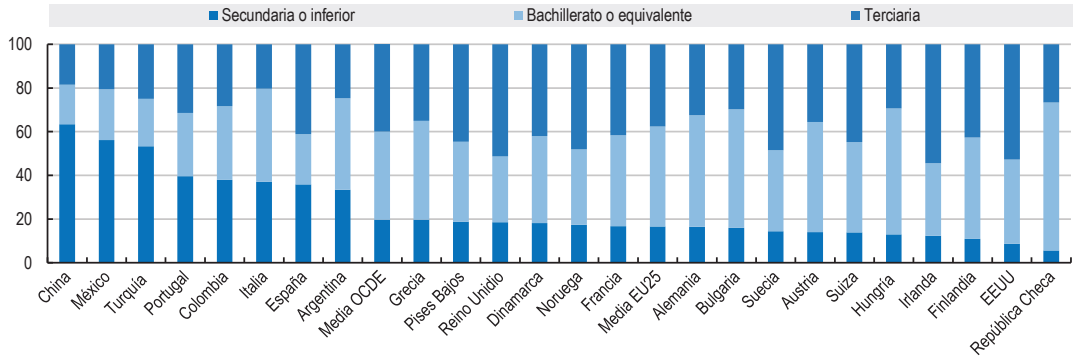
## 1. INTRODUCCIÓN

La Educación Superior no puede ser analizada aisladamente del resto del sistema educativo. Aunque la rentabilidad pública y privada del sistema universitario contribuye definitivamente al acopio del capital humano por parte de un país, añadir años de educación no genera por sí solo crecimiento en una sociedad y su aportación al crecimiento depende de las competencias que los estudiantes han cimentado a partir el conocimiento acumulado en los años previos (Hanushek, 2016). Aun así, Valero y Van Reenen (2019) muestran cómo un aumento del 10% en el número de centros universitarios se traduce en un incremento del PIB per cápita de un 0,4% a través de estimular la innovación y la oferta de capital humano, además de desempeñar un papel en la promoción de la democracia, la equidad social y los comportamientos cívicos.

El gráfico 1 muestra como la Educación Terciaria se ha extendido la mayor parte de los países de nuestro entorno, y en algunos países como Japón, Irlanda, Reino Unido o Estados Unidos más de la mitad de la población adulta en edad laboral ha pasado por sus aulas. Se ha producido una generalización de los estudios superiores independientemente de la estructura de los distintos sistemas que ha afectado a diferentes aspectos del sistema universitario como puede ser la

rentabilidad privada, la homogeneidad de los centros, el coste de los estudios y la segregación dentro de los sistemas universitarios. Por ejemplo, en el caso de España, al comienzo de la democracia en 1975 había 26 universidades públicas y 4 de la iglesia católica a las que acudían cerca de 540.000 estudiantes. Hoy, casi 50 años después, el número de estudiantes universitarios ronda el 1.700.000 en 86 universidades qué incluye 50 públicas y 36 privadas o de la Iglesia católica, 6 más en el proceso de puesta en marcha, además de cientos de centros que imparten Educación de Superior no universitaria (FP Superior) y decenas de instituciones internacionales que imparten sus títulos superiores en España.

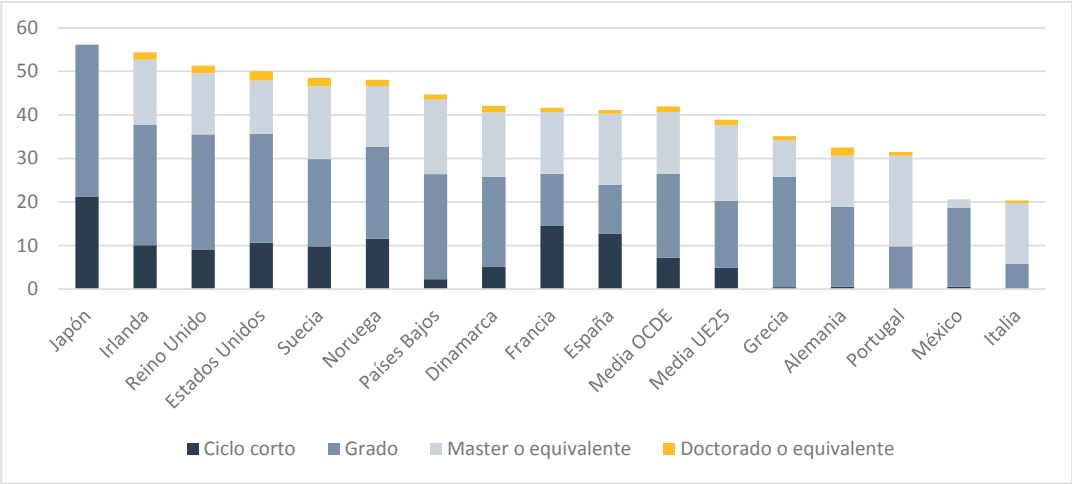
**Gráfico 1. Nivel de estudios entre 25 y 64 años (2022).**




Fuente: Education at a Glance, OCDE, 2023.

Si nos atenemos al nivel educativo alcanzado por la población con Educación Superior (gráfico 2) podemos observar la heterogeneidad la composición de la población con estudios terciarios de distintos países de la OCDE. Como tendencia general, hay un creciente número de estudiantes que acceden a la Educación Superior a través de los ciclos cortos (OCDE, 2023), mientras que el porcentaje de doctores raramente pasa del 2% de la población y las tasas de los estudios de Máster se encuentran alrededor del 14% (véase también el capítulo 1 de indicadores de este manual).

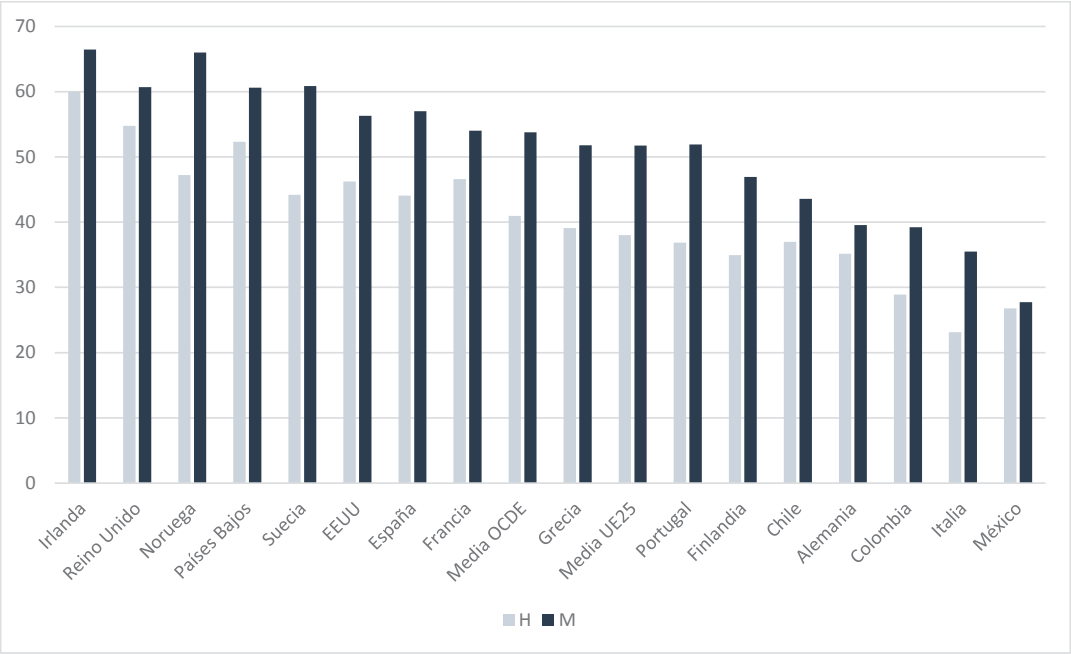
**Gráfico 2. Distribución de la población adulta con estudios terciarios por el nivel alcanzado (25-64 años) (2022).**




Fuente: Education at a Glance, OCDE, 2023.

Otro aspecto importante de la Universidad moderna es la inclusión. El gráfico 3 muestra que las mujeres representan una mayoría en la Educación Terciaria. Aunque todavía queda mucho camino por recorrer, la mayor preparación académica de las mujeres es un factor relevante para conseguir la igualdad en otros ámbitos profesionales y en la misma Universidad. Sin embargo, existe todavía en muchos países una fuerte segregación en su sistema universitario que beneficia a ciertos grupos sociales. Esta realidad debería abocar a la intervención de las Administraciones Públicas para evitar que se produzcan brechas en cuanto a la posibilidad de acceso y aprovechamiento en igualdad de condiciones que serán apuntadas a lo largo de este capítulo, como ya lo fueron en capítulos anteriores en el sistema educativo general.

**Gráfico 3. Distribución por género de la población adulta (25-34 años) en Educación Terciaria (2022).**




Fuente: Education at a Glance, OCDE, 2023.

Algunos mecanismos de funcionamiento propios del sistema universitario son, curiosamente, poco conocidos, probablemente por la falta de transparencia de los mercados en Educación Superior. Muchos de los ámbitos de la función de producción de las universidades corresponden a un híbrido público-privado, a decisiones condicionadas por las preferencias de los estudiantes, su talento, su formación previa, los sistemas de acceso, las características socioeconómicas de su familia, el marco regulador de cada país, la estructura del mercado universitario en cada país o región, la calidad del profesorado, la financiación universitaria, las expectativas profesionales que tienen los usuarios del sistema al terminar sus estudios o, a algo tan etéreo como la calidad de la gestión universitaria. La aportación de cada uno de estos elementos al producto final, el egresado, varía significativamente de tamaño y calidad, especialmente por las complementariedades que existen entre cada uno de ellos y que pueden ser determinantes en ese resultado final. Por ejemplo, y en base a la estructura del mercado universitario, Aghion et al. (2010) demuestran causalmente que un aumento exógeno del pre-

supuesto de una universidad genera más patentes o publicaciones si tiene mayor autonomía o competencia.

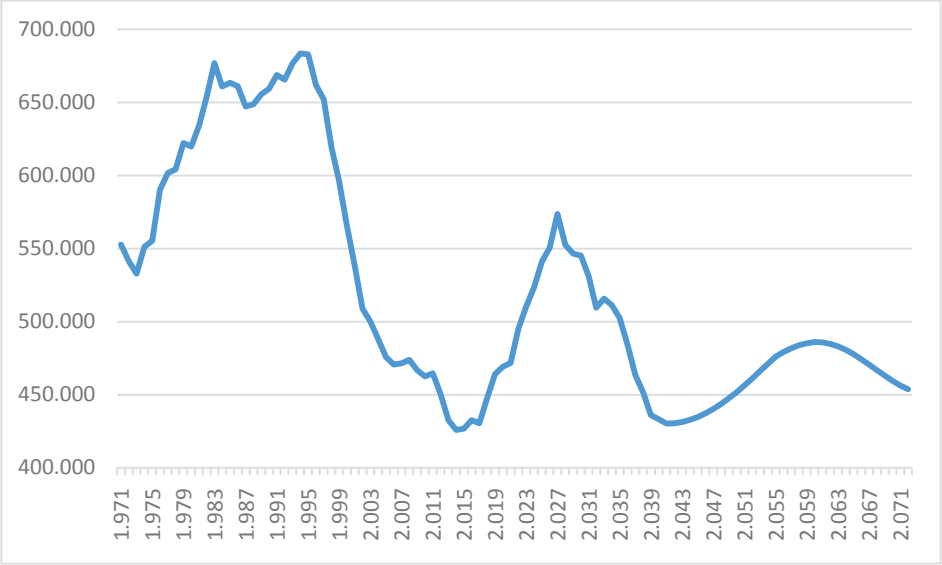
Los mismos autores en un documento previo (Aghion et al., 2008), y que sigue siendo actual (desgraciadamente), plantean para Europa la necesidad de reformas en aspectos como el incremento de la movilidad de los estudiantes, la flexibilización del currículum y la selección de entrada de profesores e investigadores. Para conseguir que las universidades europeas sean competitivas deben incrementar la competencia entre ellas y con universidades internacionales, lo que se consigue a través de la movilidad y la mayor autonomía, que por supuesto debe ir acompañada de una mayor transparencia. Por último, también es necesaria una mejor financiación del sistema.

Tres lustros después algo se ha avanzado en alguno de estos ámbitos. Programas como las universidades europeas tratan de dotar al sistema universitario europeo de mayor movilidad y competencia, en investigación, los sistemas de financiación el programa Horizonte de la comisión están creando los incentivos necesarios para una mejor investigación. Sin embargo, a nivel de país la heterogeneidad sigue siendo dominante. El pretencioso objetivo final de estas páginas es dotar a gestores e investigadores de instrumentos para seguir profundizando en ese cambio.

## 2. LOS MERCADOS DE ENSEÑANZA SUPERIOR Y LA ELECCIÓN EDUCATIVA

La estructura demográfica de los países occidentales está llevando a la paradoja de que el porcentaje de estudiantes que acuden a la Educación Superior está llegando a sus máximos, mientras que la capacidad instalada en las universidades sigue creciendo. Por ejemplo, en España, las proyecciones del INE sobre la población en edad típica de entrar a la Universidad, los 18 años, tiene una tendencia claramente a la baja (ver gráfico 4). Mientras, la ratio de población que acude a la Universidad dentro del grupo de edad está en máximos (para más detalle ver el capítulo 4b de Lacuesta y Hansen en esta obra). La conclusión es que en las sociedades en nuestro entorno en general, y España en particular, puede encontrarse ya pasado el punto de máximo de población en las aulas.

**Gráfico 4. Evolución y proyección población española de 18 años.**




Fuente: INE.

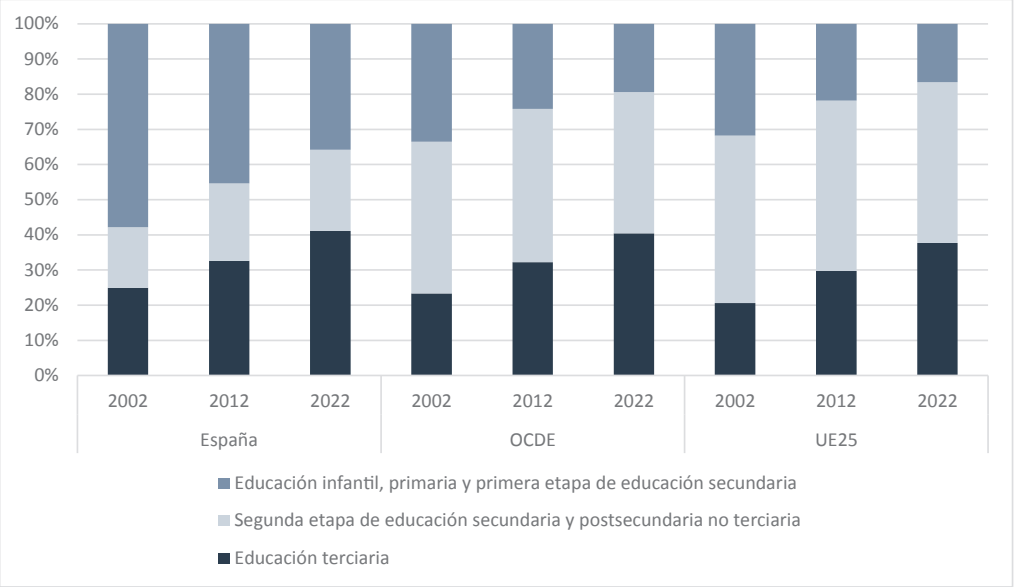
En España, a estos factores hay que añadir dos aspectos; la irrupción de la Educación Terciaria no universitaria (Formación Profesional Superior) y frente a la estabilidad en el número de universidades públicas (50 desde que en 1998 se creó la Politécnica de Cartagena), el crecimiento de los centros privados (42 universidades, entre privadas, católicas y concordatarias, 26 de ellas en los últimos 25 años) que compiten por una población menguante (Kurban et al., 2015; OECD, 2022b). El objetivo de las próximas secciones es ofrecer una semblanza completa, pero no exhaustiva, de la literatura existente del comportamiento de los estudiantes con el objetivo de apoyar el análisis de por qué y dónde estudian nuestros jóvenes universitarios.

## 2.1. El mercado de Educación Superior: FP vs Universidad

Los diversos cambios legislativos, pero en especial la LOMCE (2013) y la Ley de Formación Profesional (2022), han consolidado la Educación Superior profesionalizante como una opción real tras la Educación Obligatoria para un porcentaje importante de estudiantes españoles (gráfico 5), incrementando la diversidad de opciones para adquirir las competencias necesarias para contribuir a la sociedad (Durazzi, 2019; OECD, 2022a). La creación de un

sistema de pasarelas efectivo entre la FP y la Universidad, la generalización de la dualidad y la alta empleabilidad han incrementado el porcentaje de jóvenes que eligen estudios superiores de corta duración frente a los estudios tradicionales universitarios, en una tendencia similar a otros países de nuestro entorno (Marginson, 2016).

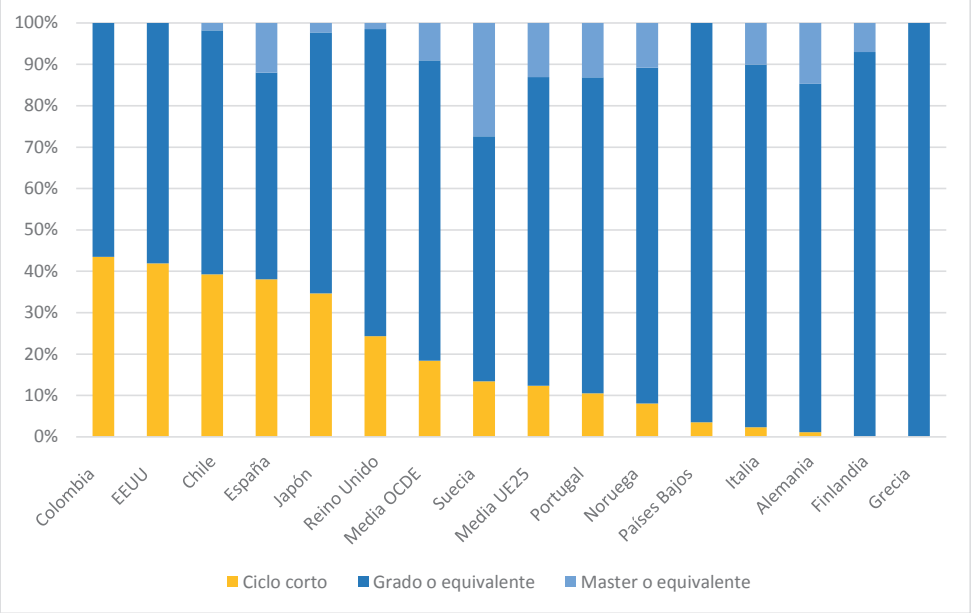
**Gráfico 5. Evolución del nivel de formación de la población adulta (25-64 años).**




Fuente: Education at a Glance, OCDE, 2023.

Por lo tanto, ya no estamos hablando de Educación Superior sí o no, sino de cómo, cuándo y dónde ya que la decisión plantea cuestiones vinculadas a la estratificación social, salarios futuros y movilidad entre clases (Hoxby, 2004; Marginson, 2016). El gráfico 6 muestra la evolución del porcentaje de población egresada en los distintos niveles de Educación Superior y la tendencia al incremento de la FP Superior (Educación Superior de ciclo corto), como ya se comenta en la introducción descriptiva del capítulo 1 y del capítulo 3b, vinculada a la demanda de competencias específicas, la recuperación del prestigio de la formación profesionalizante, en gran parte por la mayor empleabilidad (Martínez-Morales y Marhuenda-Fluixá, 2020).

**Gráfico 6. Distribución de las personas que acceden por primera vez a Educación Terciaria, por nivel de educación (2021).**




Fuente: OCDE, Education at a Glance, 2023.

Aunque existen diferencias importantes, es un buen momento para sacar lecciones de los Community College (CC) estadounidenses, que se han transformado en una opción de ciclo corto de Educación Superior profesionalizarte, sin abandonar su paso previo como formación académica conducente a un Grado académico en un centro universitario. Comenzando desde puntos de partida distintos, el resultado final ha convergido con el cambio que se ha producido en la demanda educativa en busca de nuevas competencias y flexibilidad en la formación superior corta (Kasper, 2003; OECD, 2018, 2022a) y la necesidad de los potenciales trabajadores de corregir el estigma que supone en muchos países el desempleo (Ayllón et al., 2022).

Triventi (2013) apunta uno de los problemas vinculados a la existencia de distintos niveles de Educación Superior: la estratificación social que se puede generar según los niveles educativos y que puede traducirse en reducción de la movilidad social. Aunque el artículo no presenta causalidad, sirve como referencia para el riesgo que puede suponer la Formación Profesional Superior como la opción para que las familias con menos formación académica y rentas más bajas, al menos en el caso de España.

En Estados Unidos sí que podemos ver ejemplos de causalidad vinculados a este fenómeno. Mountjoy (2022) utiliza datos de bajadas de tasas en todos los niveles de Educación Superior del estado de Texas como variable instrumental que resuelva los sesgos de selección en el acceso a la Universidad para medir el éxito académico y las ganancias de cuatro cohortes de estudiantes y su acceso o no a la Educación Superior y en su caso, ciclo corto o el equivalente al Grado. Sus resultados señalan que aquellos estudiantes que no tenían pensado acudir a la Educación Superior, pero que finalmente asistieron a CC aumentan sus salarios en un 18% por encima de la media, su educación media en 1,7 años y tiene un 27% más de posibilidades de cursar un Grado en la Universidad. El tercio restante, aquellos estudiantes que dejan de acudir a la Universidad a cursar un Grado de 4 años y se centran en estudios profesionalizantes, reduce en un 20% su posibilidad de obtener el Grado con la consecuente caída en sus ganancias futuras.

Otros artículos que analizan intervenciones similares obtienen resultados que igualmente muestran mejoras en el acceso a la Educación Superior, directamente a través de los ciclos cortos o utilizándoles como un paso intermedio hacia los Grados. Denning (2017), utilizando una metodología de diferencias en diferencias, obtiene que una bajada de las tasas de 1.000 dólares incrementa la participación en 5,1% en Educación Superior sin que se vean reducciones de las tasas de graduación, mientras que Carruthers et al. (2023), comparando de forma sintética cohortes, obtienen que una eliminación de las tasas se traduce el incremento del 24% en la graduación en el ciclo corto, pero no tiene ningún efecto sobre el acceso a los Grados y tampoco sobre los salarios.

Otro aspecto interesante por su traslación a España, especialmente por la brecha que el desempleo genera sobre los más jóvenes (Ayllón et al., 2022), es la relación que existe entre el salario mínimo y la matriculación en los estudios profesionalizantes. Schanzenbach et al. (2023) utilizando una estrategia de sucesos apilados (Autor et al., 2006) encuentran un descenso de un 4% en la matriculación de los programas de ciclo corto en el primer año, que además se mantiene a lo largo del tiempo, mientras que en los programas de Grado no existen variaciones significativas. El resultado implica que a pesar de ser estudios en los dos casos superiores y, hasta cierto punto, sustitutivos, la demanda de uno y otro es distinta, generando comportamientos distintos (Betts y Mcfarland, 1995).

## 2.2. Elección de área de estudio

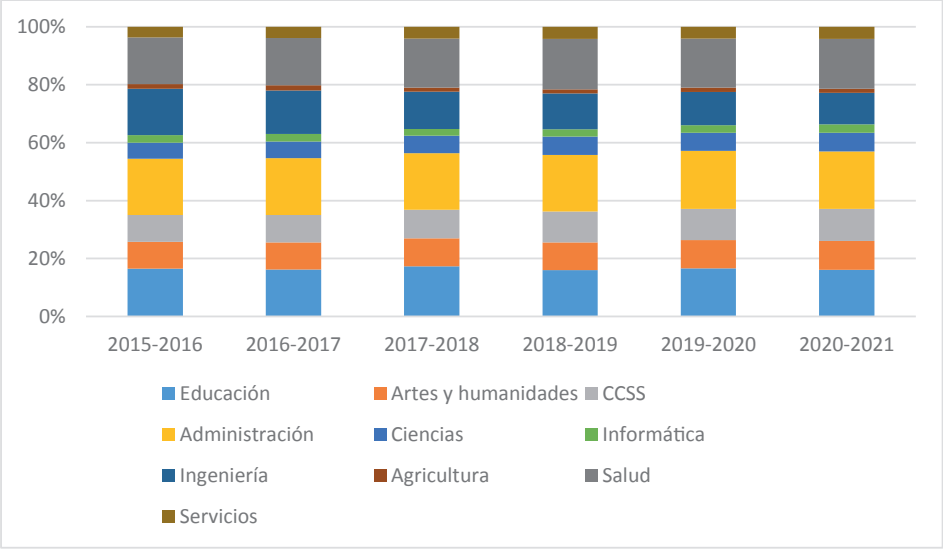
La elección de carrera es, probablemente, el primer momento más complicado de la vida adulta para muchos jóvenes españoles. La decisión, como ponen de manifiesto los datos del estudio de inserción laboral de los graduados españoles, realizada por el Ministerio de Universidades, tienen un impacto a través de toda la vida del joven.

Hay dos aspectos que hay que tener en cuenta a la hora de analizar cómo los estudiantes realizan esa elección. Primero, la elección está determinada en la Universidad pública y en las privadas de mayor calidad, y en general en todas en carreras con mayor demanda, como Medicina, por la prueba de acceso a la Universidad, es decir, exclusivamente por méritos académicos. Este sistema tiene el riesgo de excluir a estudiantes con capacidades extraordinarias pero que no se traduce en resultados académicos convincentes en todas las asignaturas evaluadas, afectando a la diversidad de talento en las aulas (Kamis et al., 2023). Además, existen críticas a la diversidad de nivel educativo por Comunidades Autónomas, consistentes con la brecha de nivel mostradas en PISA (OECD, 2019), que en un marco de distrito único pueden afectar a las posibilidades de algunos estudiantes a acceder a las carreras con mayor demanda como son las Ciencias de la Salud y algunas ingenierías (Ruiz-Lázaro, 2021).

El gráfico 7 muestra la evolución del porcentaje de egresados según ámbitos de estudio. La tendencia consolidada en nuestro país es un aumento de las carreras de Ciencias Sociales y simultáneamente la reducción de las carreras de Ciencias, comportamiento consistente con las tendencias de otros países de nuestro entorno. Sin embargo, siguen siendo las carreras de Ciencias las que tienen un mayor nivel de empleabilidad y unas mejores condiciones laborales y sueldos (gráfico 8), lo que plantea una paradoja entre las expectativas de futuro de los jóvenes y la demanda que hacen de títulos.

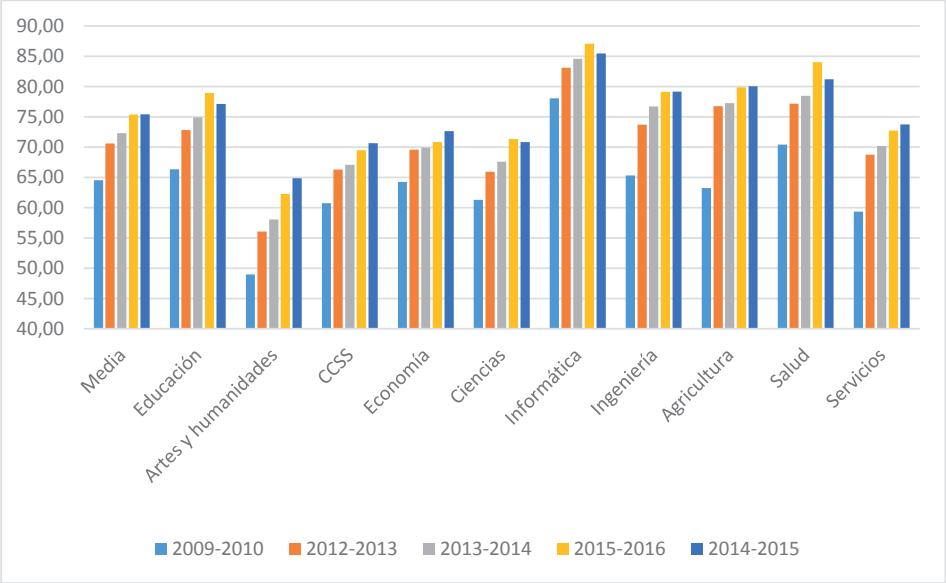
Desde el punto de vista económico las decisiones de acceso a la Universidad se suelen plantear como una elección bajo incertidumbre donde las distintas opciones de dirección de carrera dependen de la expectativa de institución y estudios que maximizan la utilidad esperada a partir de los datos disponibles y las elecciones alternativas, la experiencia familiar, el género y las habilidades cognitivas de los candidatos para el desempeño de ciertas áreas (Anelli y Peri, 2015; Arcidiacono et al., 2012; Coenen et al., 2021; Patnaik et al., 2021). Uno de problemas que surgen en este análisis es la diversidad entre las creencias y la

**Gráfico 7. Evolución del porcentaje de egresados según ámbitos de estudio.**




Fuente: INE.

**Gráfico 8. Afiliación a la Seguridad Social según ámbitos de estudio 4 años después de acabar.**




Fuente: INE.

realidad de las expectativas de las posibles salidas en las distintas carreras, lo que genera sesgos, como la ventaja absoluta en un área (no observable) puede no corresponder a la remuneración salarial, en la estimación si no son instrumentados. Estos sesgos son difíciles de tratar econométricamente (Altonji et al., 2016; Cardoso et al., 2016).

Hastings et al. (2013) presentan el análisis de elección en el modelo chileno, que es similar al español, consistiendo en un sistema centralizado de acceso a la Universidad en base a un examen académico análogo a la EVAU que determina las posibilidades de los estudiantes para acceder a los más de 1.000 títulos que ofertaban en 2012 las universidades chilenas. Los estudiantes eligen entre 8 de esas titulaciones en función de su nota y son elegidos para la realización de ese título en la Universidad seleccionada la función de un algoritmo similar al que utilizan las distintas Comunidades Autónomas.

Esto permite aplicar regresiones de discontinuidad en los niveles de corte en función de los datos demográficos y las expectativas de ganancias en el futuro. Como resultado destacar la diferencia de retorno salarial entre los títulos en el cuartil de menor demanda (2%) respecto a los títulos en el cuartil de mayor demanda (9,2%). Específicamente, dentro del mismo cuartil, entre los estudios de Ciencias de la Salud y los de Ciencias Sociales. La diferencia entre cuartiles es mucho mayor, llegando a ser las diferencias entre el primer y cuarto cuartil de un 24,7%. Esos resultados son consistentes independientemente del nivel social de origen de los estudiantes. La única excepción son los títulos de Administración de Empresas cuya empleabilidad vienen dadas en Chile por la pertenencia a redes de conocidos, y por la presencia de habilidades no académicas lo que genera resultados no consistentes con el resto del análisis.

Tampoco está claro cómo se construyen las preferencias de los estudiantes. Wiswall y Zafar (2015), a través de un amplio panel de información personal que tanto el salario esperado como las habilidades del estudiante juegan un papel importante mientras que algunas características familiares y algunos aspectos vinculadas a la oferta también son relevantes. Sin embargo, la mayor parte de la decisión sigue correspondiendo a características heterogéneas vinculadas a preferencias personales difíciles de describir, que pueden ser innatas o construidas endógenamente a lo largo de la vida y donde el género puede tener cierto peso.

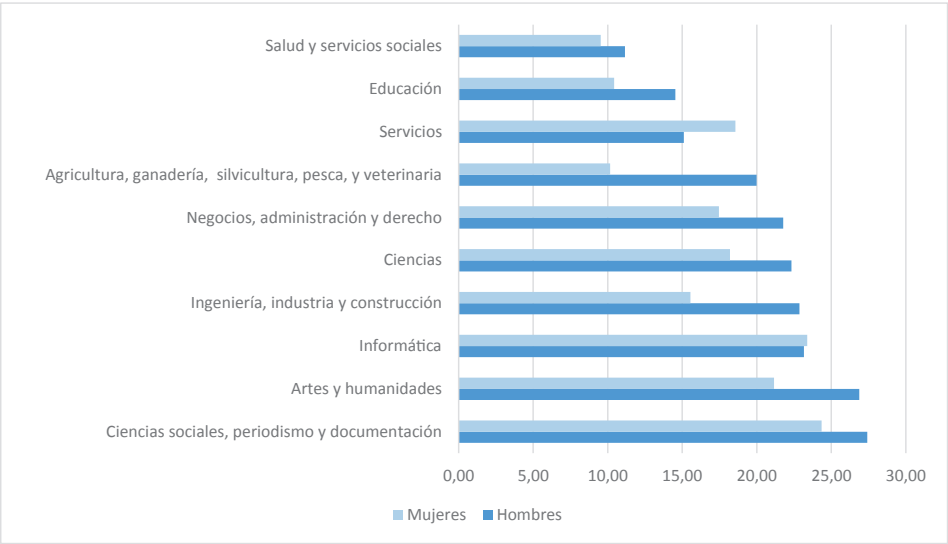
Otro aspecto que considerar es el sistema de acceso. En países como Chile, China, España, Francia o Noruega se produce simultáneamente la elección de Universidad y de carrera. Por esa razón, los resultados de Bordon y Fu (2015)

para Chile, son interesantes. Para analizarlo contraponen un modelo de equilibrio donde un planificador reparte las plazas de carrera y Universidad y compara los resultados con un juego de Stackelberg donde el planificador elije la Universidad y el estudiante el Grado. Los resultados muestran que en el segundo se produce un incremento de la matriculación en los tres cuartiles inferiores, pero especialmente para el cuartil más bajo, un incremento de las tasas de retención en todos los grupos. Como corolario, el análisis muestra que mujeres y los estudiantes de familias de rentas bajas tienen más probabilidades de beneficiarse por el cambio de sistema, unos resultados que son consistentes con las recientes reformas en el acceso a la Universidad en China (Ma et al., 2023).

### 2.3. Abandono en la Universidad

Si como hemos visto en la introducción, terminar los estudios de Educación Superior es ventajoso para los estudiantes ¿quiénes y por qué dejan atrás las aulas? El gráfico 9 da una primera visión con los últimos datos disponibles del Ministerio de Universidades (SIIU, 2023). Se observa un fuerte componente de género y de área de estudio con respecto al 11,7% de jóvenes que abandonan el primer año de estudio en la Universidad y un 22,85% de abandono a lo largo

**Gráfico 9. Tasa de abandono del SUE el primer año por sexo (2019-2020).**




Fuente: SIIU, 2023.

del Grado del Sistema Universitario Español (técnicamente, aquellos que, tras haber iniciado la carrera, no se matriculan en ningún estudio universitario oficial de ninguna Universidad española durante dos cursos seguidos ni se titulan).

También se puede ver un importante componente geográfico. A los 3 años de haber iniciado estudios, el abandono en comunidades como Canarias (20,5%) o Baleares (20,4%) es casi el doble que en Madrid (10,8%) o País Vasco (11,8%). Sin embargo, resulta curioso que mientras los estudios más selectivos, que corresponde en estos momentos son Salud y Ciencias tienen bajas tasas de abandono reducidas, los estudios de educación, las medias de entrada son relativamente bajas, se caracterizan por una tasa relativamente baja de abandono.

Existe una amplia literatura sobre las causas del abandono universitario (por ejemplo, de Aina et al., 2018, 2022; Behr et al., 2020; Kehm et al., 2020). Un breve resumen de la literatura señala, entre los principales aspectos que determinan la continuidad o no: a) condiciones de estudio en la Universidad; b) integración académica y social en la Universidad; c) esfuerzos y motivaciones personales para estudiar; d) información y requisitos de admisión; e) rendimiento académico previo; f) coste y financiación de la Universidad; g) características personales; h) características sociodemográficas; e i) condiciones externas. Parte de estas características son subjetivas al individuo y plantea dificultades en su estimación y evaluación Altonji et al. (2016) o Braga et al. (2014) señalan la importancia que tiene la calidad del profesor para el rendimiento de los estudiantes, medido por el rendimiento en los cursos posteriores, sin embargo, sus resultados también señalan que las evaluaciones de los estudiantes no están correlacionadas con el aprendizaje, ya que pueden influir muchos aspectos como el sesgo de selección de los estudiantes, las expectativas de notas en la asignatura, o la propia inflación de notas dentro de las distintas universidades.

Para el caso español, Lassibille et al. (2008) estiman modelo logit que establece la probabilidad de abandono de los estudios universitarios es mayor para la carrera de Derecho o Ciencias frente a los de Ciencias de la Salud. Los datos de genero no son estadísticamente significativos, pero si encuentran una mayor tasa de abandono para los estudiantes que retrasan la continuación de sus estudios tras el Bachillerato, han repetido un curso durante Secundaria o provienen de la Formación Profesional. También es un factor importante para seguir en las aulas el hecho de recibir beca, si el estudiante está cursando su primera elección de estudios o el hecho de que los padres tengan carrera universitaria, resultados consistente con las conclusiones causales para España de Montalbán (2023) en

un contexto distinto o los de Backes-Gellner et al. (2021) para Suiza; Modena et al. (2018); Vergolini y Vlach (2017) para Italia, etc.

Por último, señalar la paradoja de las ingenierías, y las carreras STEM y su elevada tasa de abandono a pesar de ser selectivos en las competencias académicas que caracterizan a sus estudiantes (Lassibille et al., 2008), y la demanda de sus habilidades en una sociedad digital (Autor et al., 2003; Bresnahan et al., 2002) vinculada a la falta de motivación de los estudiantes a pesar de la alta empleabilidad, elevado retorno o notas de ingreso (Tayebi et al., 2021).

### 2.4. Epítome: mercados que no existen pero que deberían existir

El mercado universitario en España es un mercado incompleto. No existe un mercado de contratación de académicos como se entiende en la mayor parte del mundo civilizado y es una rareza la existencia de un mercado de contratación de profesores jóvenes y la movilidad de los más senior (García Hombrados, 2022). El marco de la nueva Ley de Universidades aprobada en 2023 (LOSU) genera oportunidades, pero también incertidumbres que según como se resuelvan permitirá que en una segunda edición de este libro se puede hablar de incentivos, discriminación en la contratación, estratificación, jerarquía, redes, etc. (Carlsson et al., 2021; Checchi et al., 2021; Ginther y Kahn, 2021; Musselin, 2009; Terviö, 2011; Wapman et al., 2022; Williams y Ceci, 2015).

Tampoco hemos hablado de los títulos universitarios de 3 años, que permitidos a partir del año 2015, desaparecen de forma generalizada en la LOSU, a no ser que sean vinculados al programa de las universidades europeas como en el caso de Italia. Cappellari y Lucifora (2009) encuentran evidencia empírica mediante el uso de una metodología de diferencias-en-diferencias donde las discrepancias entre cohortes se deben exclusivamente a la implantación de Bolonia, que en el país transalpino sí significó pasar a tres años de estudios en la titulación de Grado. Los estudiantes que egresaron de Secundaria tras el cambio tienen un 15% más de probabilidades de asistir a la Universidad debido a una decisión en la inversión en capital humano y no por la tendencia subyacente a largo plazo en la asistencia a la Universidad. También hay un componente social: la mayor parte del aumento se da entre los alumnos brillantes procedentes de entornos con rentas bajas, aunque si constatan un pequeño incremento en las tasas de abandono universitario, consistente con los resultados anteriores.

## 3. ELITISMO, DISCRIMINACIÓN Y ACCESO

El filósofo Michael Sandel critica duramente en su libro de 2020, "La tiranía del mérito" (Sandel, 2020), el sistema de acceso a las universidades americanas. Apoyado en el escándalo de las 33 familias cuyos hijos consiguieron acceder a universidades situadas en el primer decil de los rankings[1], plantea la consideración de élites basada en factores que escapan al control del individuo, como las condiciones socioeconómicas y la valoración que el mercado otorga a ciertas habilidades y destrezas sobre otras. Por lo tanto, es crucial comprender que el azar desempeña un papel importante en el concepto de "éxito".

El antecedente empírico de ese tipo de planteamiento corresponde a Dale y Krueger (2002, 2014), artículos que se analizarán con detalle en la siguiente sección al analizar la rentabilidad de la educación universitaria, que desmitificaron lo que hasta entonces había sido un dogma: conseguir entrar en una Universidad de la Ivy League aseguraba una vida de "éxito". Las diferencias salariales entre los egresados de universidades más selectivas frente a los de aquellas menos selectivas parecía justificar el argumento. Sin embargo, Dale y Krueger plantean que las diferencias existentes no corresponden a la asistencia a una Universidad a otra, sino de características previas no observables, que influyen en el proceso de admisión, y que determinan los ingresos tras la graduación (diligencia, inteligencia, iniciativa).

Hoxby (2009) crítica su estrategia de identificación y caracteriza la selección de la Universidad como un proceso más fácil para la mayoría de los estudiantes, mientras que los estudiantes más capaces, debido al cambio tecnológico que ha facilitado su movilidad, compiten por un número de plazas que se ha mantenido relativamente constante, en las mejores universidades. Esto genera una dicotomía entre los estudiantes que quieren acceder a las universidades de élite y aquellos que simplemente quieren ir a la Universidad.

Chetty et al. (2023), que también será analizado en la siguiente sección, estudian porqué los hijos de las familias con mayor renta tienen más probabilidad de acceder a las universidades privadas más selectivas de Estados Unidos (Ivy

[1] De hecho, los mecanismos utilizados por estas familias para que sus hijos accedieran a estas universidades lo que muestra es la necesidad de más meritocracia en lugar de menos. Esto no podría suceder en un sistema como el español.

League, Stamford, MIT, etc.). Los datos son concluyentes, dentro del grupo de estudiantes que están en el percentil 99 de mejores resultados en el examen de acceso a la Universidad, más del 30% vienen de familias que están en superior de la renta y acaban en las aulas más selectivas (un 9,3% más de posibilidades de las que les correspondería) y solo el 10% pertenecen a la clase media (percentiles 70-80). Corrigiendo por la nota de acceso, los autores encuentran que los estudiantes con familias en el percentil 99-99,9 tienen un 43% más de probabilidades de ser admitidos en una de estas instituciones que los de clase media, mientras que, en comparación, las tasas de admisión en universidades públicas selectivas son parecidas para toda la distribución de ingresos[2].

Pero ¿cuáles son las causas que determinan esta mayor participación universidades de élite? La primera que los padres del futuro estudiante hayan sido antiguos alumnos de la Universidad, la segunda, las credenciales que consiguen presentar. Los estudiantes del percentil 99 no tienen mejores credenciales académicas que los de clase media pero sí en actividades extracurriculares y también mejores cartas de recomendación de sus profesores en el colegio, especialmente si son también selectivos. En resumen, dos tercios de las mayores tasas de asistencia a centros selectivos del 1% superior en relación con la clase media se explica por una ventaja de admisión que se deriva de tres factores: el deporte universitario, las preferencias por hijos de exalumnos y la mayor puntuación de aspectos extraacadémicos.

Grossman et al. (2023) analizan otro aspecto controvertido en el proceso de admisión estadounidense, como es la discriminación de los estudiantes de origen asiático (pero estadounidenses) en el acceso a las Universidades, especialmente a las de élite. Sus conclusiones señalan que individuos con origen en el sudeste asiático observan una disminución significativa en sus posibilidades de admisión, un 49% en comparación con sus homólogos blancos con valores equiparables en cuanto a calificaciones en exámenes, promedios de calificaciones de Bachillerato y participación en actividades extracurriculares. Tras controlar por solicitudes anticipadas, la institución de procedencia del solicitante y padres

[2] Hoekstra (2009), utilizando una metodología de discontinuidad en la regresión, encuentra que aquellos estudiantes que consiguieron acceder en el límite a universidades públicas de alta calidad frente a los que se quedaron fuera del corte por décimas, obtenían diferencias salariales de hasta un 21%.

exalumnos, encuentran que los estudiantes del sudeste asiático logran tasas de admisión similares a las de sus pares blancos. En contraste, los estudiantes de ascendencia en India y Pakistán experimentan una disminución del 10% en sus posibilidades de admisión en comparación con los estudiantes blancos que presentan similitudes en las variables mencionadas.

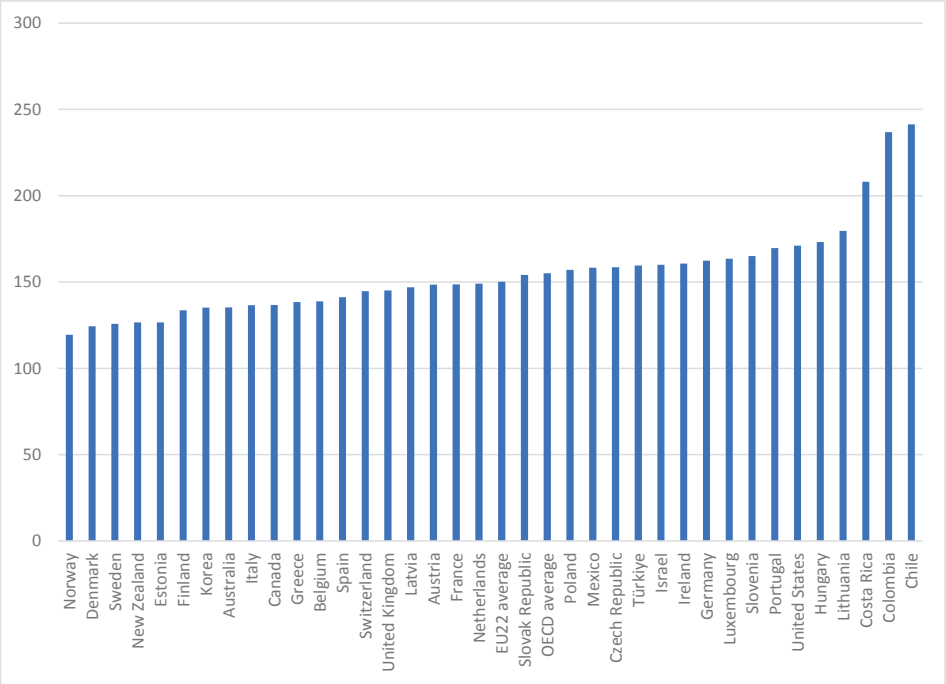
El análisis anterior nos lleva irremediablemente al efecto de la regulación de discriminación positiva en el acceso a la Universidad. Bleemer (2022) utilizando un diseño de investigación basado en diferencias, junto con una base de datos longitudinal de los solicitantes de la Universidad de California de los años 1994 a 2002 y sus posteriores trayectorias educativas y niveles de ingresos, analiza las repercusiones de la interrupción de la discriminación positiva. En concreto, demuestra que la eliminación de esta política en el estado de California dio lugar a que los solicitantes pertenecientes a minorías se matricularon en instituciones académicas de nivel inferior. El marco conceptual de la "hipótesis del desajuste" sugiere que este cambio en la matriculación conferiría aparentemente ventajas educativas netas a los solicitantes de grupos minoritarios. Sin embargo, el análisis revela un resultado opuesto, ya que la obtención general de títulos por parte de los solicitantes de grupos infrarrepresentados de menor cualificación académica experimentó un descenso, sobre todo en las disciplinas STEM. El cese de la discriminación positiva también disuadió a un número considerable de estudiantes de minorías cualificados de presentar solicitudes en los prestigiosos campus de la Universidad de California. Los análisis complementarios, que abarcan la discontinuidad de la regresión y las evaluaciones del valor añadido institucional, sugieren que los beneficios educativos y salariales netos de la discriminación positiva para los solicitantes minoritarios superan los costes netos para los solicitantes blancos y asiáticos al margen de la admisión.

En Europa hay que distinguir entre las evidencias del Reino Unido, donde la asistencia de universidades de élite (Oxbridge y, en menor medida, el Russel Group) conforman un grupo donde el género, las redes y el origen social, y el origen étnico son determinantes para el acceso (Richardson et al., 2020), y la creciente heterogeneidad en la calidad de las universidades incluso en países donde se considera una relativa homogeneidad, como Italia (Pietro y Cutillo, 2006), o en España, donde los efectos de la EBAU se detallan en el capítulo 5a de esta misma obra elaborado por Andreu y Casamiglia.

## 4. LA RENTABILIDAD PRIVADA DE LA EDUCACIÓN UNIVERSITARIA

La situación laboral de los graduados de Educación Terciaria es, en la mayoría de los países, sustancialmente mejor que la de los graduados de otros niveles educativos inferiores. Tienen menores tasas de desempleo y mayores salarios. En 2021 los graduados de Educación Terciaria de la OCDE entre 25 y 64 años tenían una tasa de empleo 10 puntos porcentuales superiores a los graduados de Educación Secundaria Superior o post-Secundaria no Terciaria. La tasa de los que no estaban estudiando ni trabajando entre los que tenían entre 25 y 29 años era un 12%, claramente inferior de la tasa de Secundaria Superior/post-Secundaria no Terciaria. También tienen tasas de desempleo de largo plazo inferiores al resto de niveles educativos. En media, los graduados de enseñanza superior en la OCDE con un empleo continuo a tiempo completo ganan aproximadamente un 55% más que los graduados de enseñanza Secundaria Superior.

**Gráfico 10. Ganancia relativa de los graduados de Educación Terciaria frente a Educación Secundaria Superior (a tiempo completo). 2020.**




Fuente: Education at a Glance, OCDE (2017).

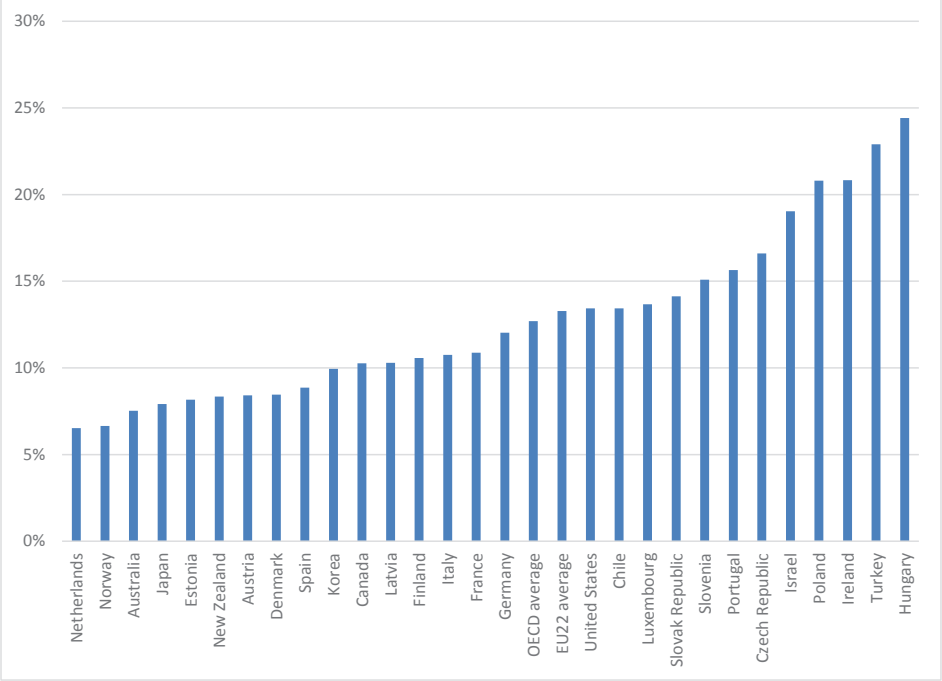
No obstante, existenten grandes diferencias entre países. Una forma simple de analizar la rentabilidad de la Educación Terciaria es comparar los salarios obtenidos frente a los graduados de Educación Secundaria Superior. El gráfico 10 presenta esta comparación por países. En la media de la UE la ventaja es un 50%, algo inferior a la media de la OCDE. En el caso español la ventaja es del 41%, algo por debajo de la media de la UE.

Una forma alternativa es calcular la tasa de rentabilidad interna de la Educación Terciaria. OCDE (2017) calcula los ingresos brutos de la obtención de estudios terciarios (teniendo en cuenta el efecto de los periodos de desempleo en cada país) menos el impuesto sobre la renta, las contribuciones a la Seguridad Social y el efecto de las transferencias. A esto se añaden los beneficios del seguro de desempleo. De los beneficios netos se restan los costes que incluyen los costes directos (precio de la matrícula por el número de años de estudios más gastos de manutención) y los ingresos no obtenidos (el salario que el estudiante podría haber obtenido en el mercado laboral teniendo en cuenta la tasa de desempleo para dicho grupo de edad). En países donde es normal estudiar y trabajar el coste de oportunidad se calcula como la diferencia entre lo que podría ganar en el mercado laboral y lo que gana compaginando el trabajo y los estudios universitarios. La diferencia entre ingresos y gastos se descuenta a una tasa del 2%, aunque aplicaciones para algunos países han utilizados tasas entre el 2% y el 5%. El gráfico 11 muestra el resultado del cálculo de la tasa interna de rendimiento de la Educación Superior en los países de la OCDE. En concreto el gráfico muestra los costes y beneficios privados de un hombre graduado de Educación Terciaria comparado con uno graduado de Educación Secundaria Superior en dólares equivalentes convertidos usando la paridad del poder de compra para el PI[3]. El cálculo muestra que en países donde el coste directo de la educación universitaria es bajo (Alemania, Noruega) el coste total de los estudios es elevado por el alto coste de oportunidad. Por su parte, en los países donde los graduados universitarios reciben salarios más altos también pagan más impuestos, lo que reduce las diferencias entre países cuando se calculan los beneficios netos. La tasa de rentabilidad media tanto de los países de la OCDE como de la UE es el 13%. En

[3] Diris y Ooghe (2018) presentan un cálculo similar, aunque con algún cambio en los supuestos. Por ejemplo, la OCDE utiliza salarios mínimos para calcular el coste de oportunidad de los años de estudios universitarios mientras Diris y Ooghe (2018) utilizan una estimación de la relación entre edad e ingresos para imputar el salario medio de los trabajadores de 20 años.

España la rentabilidad es menor (9%). En el caso de las mujeres graduadas con Estudios Terciarios la rentabilidad media de los países de la OCDE y de la UE es menor, el 11%. Sin embargo, las graduadas españolas tienen una rentabilidad superior a la media (13%).

**Gráfico 11. Tasa de rendimiento interna de la Educación Terciaria.**




Fuente: Education at a Glance, OCDE (2017).

Otra forma de calcular la rentabilidad privada de la Educación Terciaria es utilizar una regresión Minceriana (Card, 2001). La referencia clásica de este cálculo para comparaciones entre países es el trabajo de Psacharopoulos (1994) según el cual cada año adicional de Educación Superior proporciona una rentabilidad del 8% en países desarrollados. Psacharopoulos y Patrinos (2018) actualizan esas estimaciones y obtienen una rentabilidad del 9% para el conjunto de los 135 países analizados, aunque en los países desarrollados la rentabilidad sigue anclada en el 8%. Bhuller et al. (2017) proponen superar las limitaciones de las regresiones Mincerianas por el incumplimiento de los supuestos clave de esta metodología. Bhuller et al. (2017) utilizan tres procedimientos alternativos (regresión de discontinuidad, variables instrumentales y diferencias en diferencias) para superar

los problemas de selección muestral y estimar la rentabilidad de la educación utilizando una muestra longitudinal de ciudadanos noruegos. Los resultados muestran que las regresiones tipo Mincer infraestiman la tasa interna de rentabilidad de la educación que, según las estimaciones preferidas de Bhuller et al. (2017) alcanza el 11%. Los autores se preguntan cómo con unas tasas muy superiores al tipo de interés (incluso si se tienen en cuenta unos impuestos progresivos y el sistema de pensiones) los individuos no adquieran mayores niveles de educación.

La mayoría de los cálculos anteriores suponen que la rentabilidad de la Educación Terciaria es homogénea y, en particular, no depende de la Universidad elegida para seguir los estudios (siempre que sean mínimamente selectivas), del nivel socioeconómico de la familia o del Grado estudiado. Dale y Krueger (2002) analizan la rentabilidad de estudiar en universidades más selectivas frente a universidades comparables, pero menos selectivas. Utilizando una regresión con variables control observadas por los investigadores el resultado es que la rentabilidad de las universidades más selectivas es superior. Sin embargo, cuando Dale y Krueger (2002, 2014) corrigen las decisiones de admisión por las habilidades latentes de los estudiantes y por otros factores, los resultados muestran que asistir a una Universidad más selectiva en los Estados Unidos tiene poco impacto sobre los futuros ingresos de los graduados. Chetty, Deming y Friedman (2023) utilizan una metodología similar y encuentran unos resultados consistentes con los de Dale y Krueger (2002): los graduados de universidades de la Ivy leage ampliada no tienen unos ingresos superiores a los que cursaron sus estudios en una buena Universidad pública. No obstante, sí que tienen una probabilidad superior de situarse en el 1% superior de la distribución de la renta. Chetty, Deming y Friedman (2023) encuentran que no hay heterogeneidad en el efecto causal de asistir a una Universidad de la Ivy Leage ampliada[4] frente a una buena Universidad en términos de la renta de los padres, las puntuaciones del SAT u otras características de los solicitantes de admisión. Por contraposición, Dale y Krueger (2002, 2014) encuentran que los estudiantes de familias con menor nivel de renta obtienen una rentabilidad superior en universidades muy selectivas incluso ajustando por las características no observables de los estudiantes. También que existen diferencias significativas en la rentabilidad de la Educación Terciaria en función del Grado seleccionado. Kirkboen et al. (2016) usan una base de datos muy detallada del sistema de educación post-Secundaria

[4] Ivy Leage (Harvard, Princeton, Yale, Cornell, Columbia, Pensilvania, Brown y Dartmouth College) más MIT, Stanford, Duke y Chicago.

noruego para analizar el efecto de la elección del Grado. Estos autores concluyen que los diferentes campos de estudio tienen unos beneficios en el mercado laboral muy diferentes incluso condicionado a la institución y la calidad de los compañeros. En segundo lugar, muestran que el efecto de asistir a una Universidad más selectiva es pequeño comparado con el efecto de seleccionar un campo de estudio. Finalmente, Kirkboen et al. (2016) concluyen que el campo de estudio elegido por los alumnos es consistente con las temáticas en las que los alumnos tienen ventaja comparativa.

En el caso español, como mostraban los gráficos anteriores, la tasa de rentabilidad interna de la Educación Universitaria se sitúa por debajo de la media de la OCDE y la UE, pero sigue siendo bastante elevada. De esta forma también en España existen elevados incentivos privados para que los estudiantes inviertan en Educación Universitaria al igual como en la mayoría de los países.

## 5. LA FINANCIACIÓN DE LOS ESTUDIOS UNIVERSITARIOS

El gasto en Educación Superior ha aumentado sustancialmente en los últimos años por la expansión de la demanda de Educación Terciaria. Entre 2008 y 2019 el crecimiento en términos reales en la OCDE ha alcanzado el 22%, superando en más de cuatro puntos el crecimiento del presupuesto de las instituciones de educación no universitaria. En algunos países como Chile, México, Israel o Noruega el crecimiento ha sido superior al 40% en este periodo.

Hay dos modalidades fundamentales de financiación de las enseñanzas universitarias. La primera es la utilización del sistema impositivo (bien con impuestos generales o mediante impuestos a los graduados universitarios[5]) para financiar transferencias a los estudiantes (becas universales o condicionadas a la renta o las calificaciones) y a las instituciones de enseñanza superior para mantener el precio de la matrícula bajo[6]. La segunda modalidad es el pago directo de los pre-

---

[5] Este procedimiento se conoce como "graduate tax system" e implica un subsidio público a la educación que hace que el repago sea contingente a la renta, pero solo cuando las ganancias superan el coste de la educación. Esta diferencia es el subsidio púbico. En este caso, por tanto, el subsidio no proviene de la imposición general.

[6] Existe una enorme literatura académica internacional sobre el efecto los precios universitarios y las becas sobre las tasas de matriculación y las tasas de abandono. Ver resumen en García-Montalvo (2020).

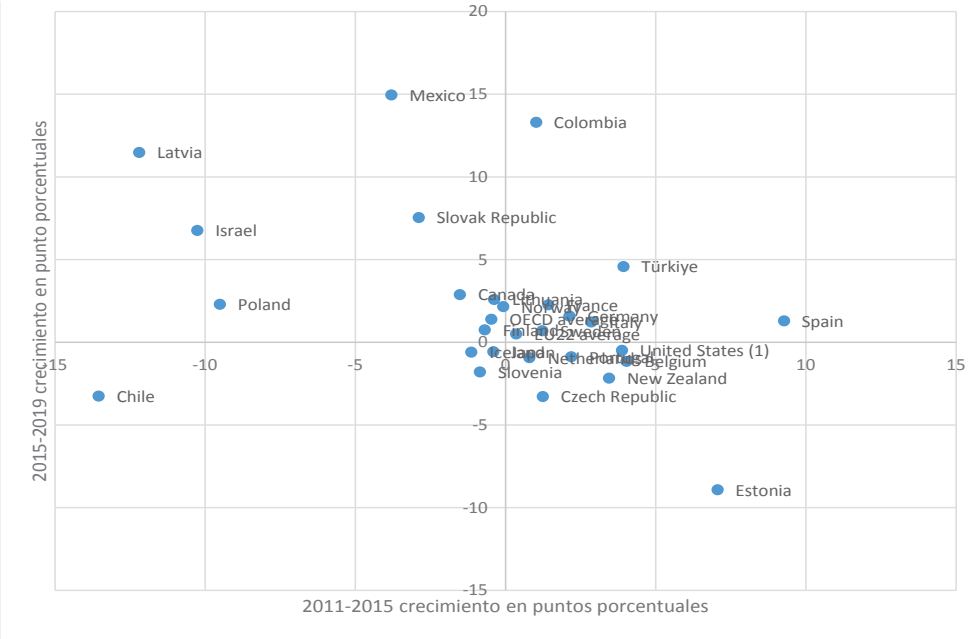
cios públicos de la Universidad bien directamente a partir de los ingresos de los estudiantes y sus familias o con la utilización de créditos. Estas dos modalidades no son excluyentes pues, en la mayoría de los casos, la financiación de la Universidad se produce de una combinación de subsidios públicos y pagos privados. A diferencia de la educación anterior a la Terciaria, donde aproximadamente el 90% de la financiación proviene de los presupuestos públicos, en la media de la OCDE la financiación privada representa una tercera parte de los recursos de las instituciones de enseñanza Terciaria.

Cada una de estas modalidades puede tener diferentes especialidades. Los subsidios derivados de la financiación impositiva pueden depender de la renta de los estudiantes y sus familias, del nivel de calificaciones de los estudiantes o de ambas. Por ejemplo, el coste de la matrícula puede ser función de las veces que se suspende una asignatura o la beca contingente a superar un número mínimo de créditos u obtener una calificación excelente. Por su parte la financiación crediticia tiene también diferentes versiones. El sector público podría proveer directamente los créditos, garantizarlos o ser créditos privados. Por su parte los créditos podrían ser no condicionados o contingentes a la renta obtenida tras la graduación (en su acepción inglesa "income-contingent loan" o ICL).

La distribución entre la financiación pública y privada de los estudios universitarios tiene una enorme carga política y es utilizada en multitud de ocasiones como un elemento de confrontación entre los partidos políticos. En esta sección se hace un análisis del impacto de los diferentes mecanismos de financiación de la Universidad en términos de su eficacia y equidad y se discute la economía política detrás de cada una de las opciones. Es importante notar que la discusión en esta sección no tiene en consideración la distribución de la financiación entre diversos niveles educativos. Por ejemplo, podría argumentarse que dado que la rentabilidad social de la educación de 0 a 3 años es muy superior a la obtenida con los estudios universitarios, la financiación pública debería dirigirse prioritariamente a este segmento educativo en detrimento del nivel educativo donde la rentabilidad privada es dominante. Este tipo de argumento queda al margen de la discusión en este capítulo.

La distribución entre financiación pública y privada de los estudios universitarios cambia a lo largo del tiempo y en función de los países. En general tras el comienzo de la crisis financiera de 2007-2008 se produjo una creciente tensión en los presupuestos públicos que redujo la financiación pública de las universidades y tendió a desplazar una parte sustancial de los costes de estos estudios

**Gráfico12. Crecimiento de la contribución privada 2011-2015 versus 2015-2019.**




Fuente: Education at a Glance (2023).

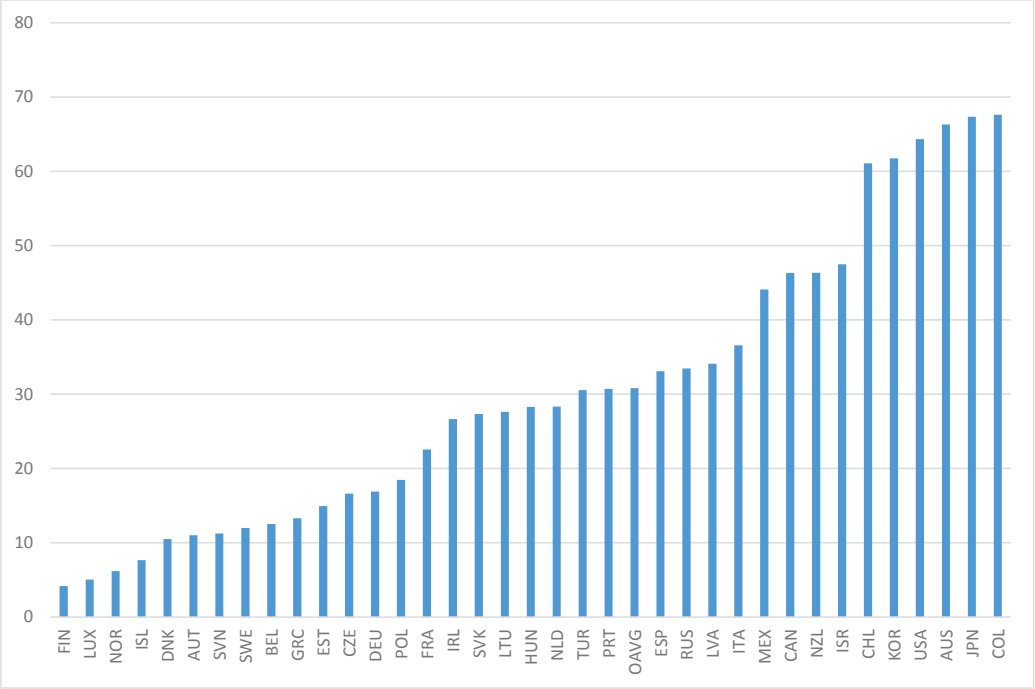
hacia los estudiantes y sus familias. El gráfico 12 muestra como en la mayor parte de los países aumentan la proporción de gasto privado en los primeros años de la crisis financiera. A partir de 2015 algunos países que habían reducido la participación privada la aumentan con excepción de Chile que sigue reduciéndola. Esta tendencia fue moderándose, e incluso revertiéndose, en algunos países a finales de la segunda década del siglo XXI.

Respecto a la clasificación por países (gráfico 13) se observa que hay un grupo, formado principalmente por países anglosajones, que de forma consistente tiene una financiación privada superior al 60% (Estados Unidos, Reino Unido, Corea del Sur, Australia, Japón y Chile) mientras otros grupos que forman principalmente los países del norte de Europa, se sitúa en niveles inferiores al 15% (Finlandia, Noruega, Islandia, Dinamarca, Austria y Suecia)[7]. El gráfico 14

[7] Es importante notar que la OCDE, la fuente de estos datos, incluye en la Educación Terciaria la formación profesional superior y otros programas. Por este motivo, por ejemplo, la proporción de financiación privada en España se sitúa en torno al 30%, cuando la financiación privada de la Universidad no alcanzaba el 20%.

muestra que no existe una relación evidente entre el gasto en Educación Terciaria sobre el PIB y la distribución de la financiación entre privada y pública[8]. El gráfico 15 desglosa el total del coste de la Educación Terciaria entre la subvención pública a las instituciones, las transferencias públicas a privados y el gasto privado.

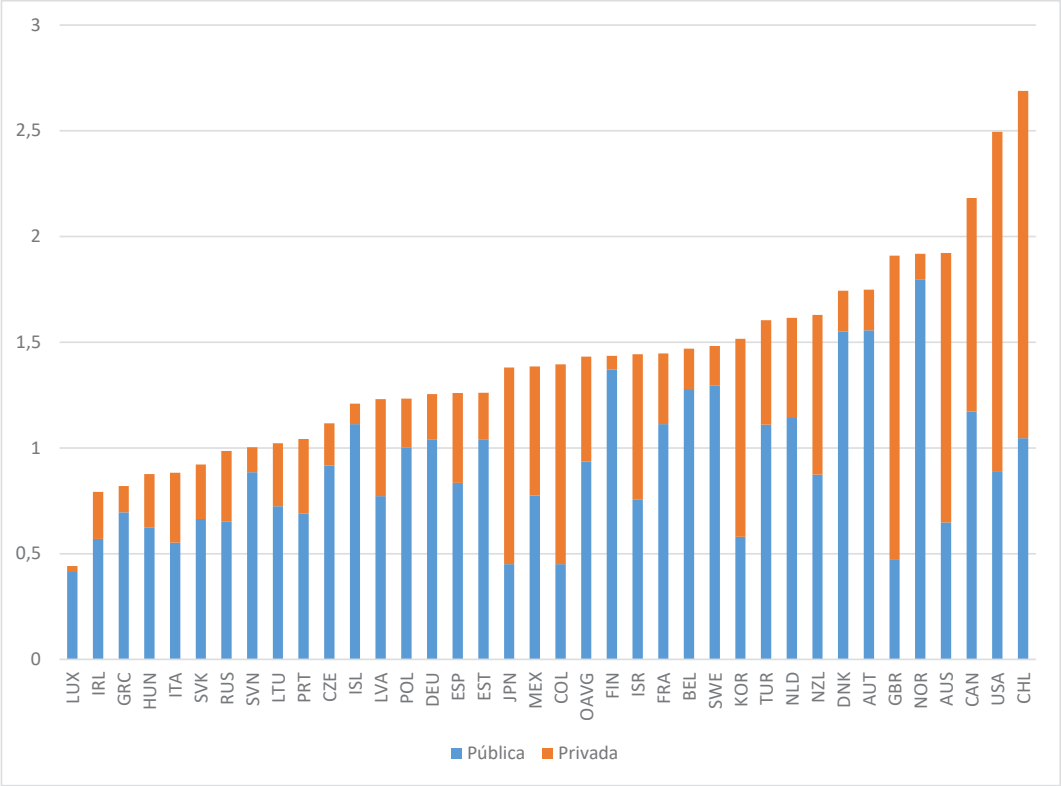
**Gráfico 13. Proporción del gasto privado sobre el total de gasto en Educación Terciaria (2019).**




Fuente: Education at a glance, OCDE (2023).

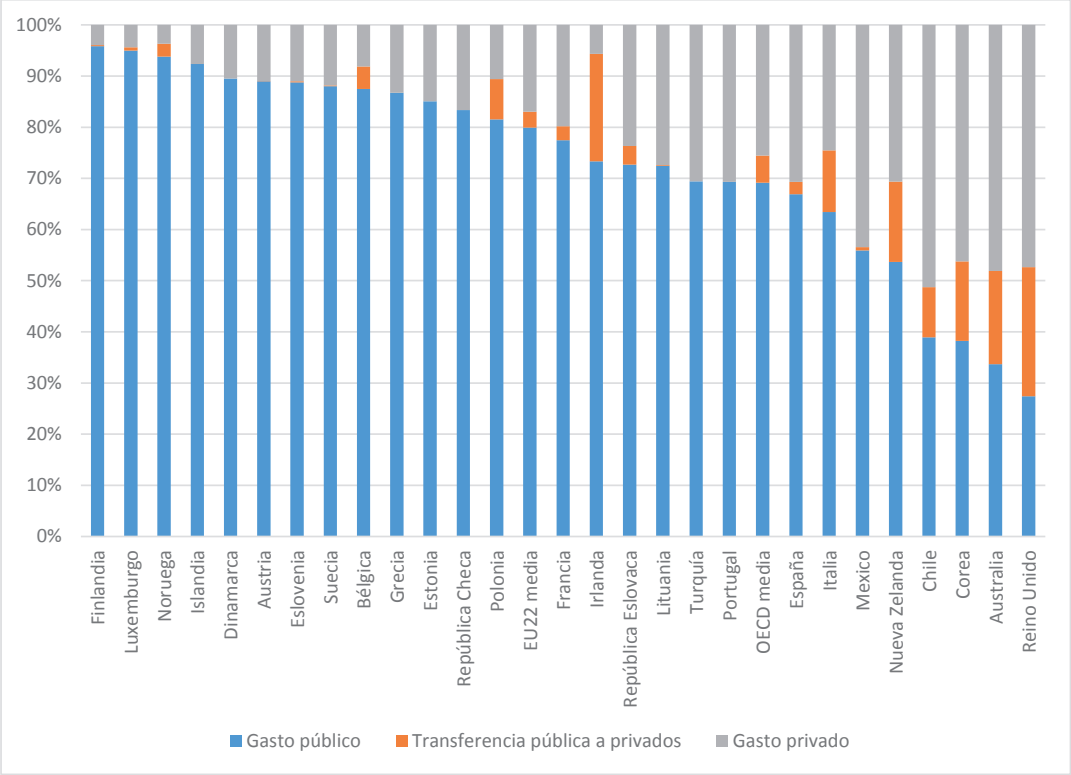
[8] La correlación entre el gasto en Educación Terciaria y la proporción de financiación privada es –0.13.

**Gráfico 14. Proporción de financiación pública y privada sobe el PIB (2019).**




Fuente: Elaboración propia a partir de OCDE (2023).

**Gráfico 15. Proporción de gasto público, privado y transferencias.**




Fuente: Education at a Glance, OCDE (2022).

La literatura recoge multitud de razones que justifican un aumento de las contribuciones privadas a la financiación de las universidades. La rentabilidad privada de la educación universitaria produce grandes incentivos para los estudiantes en términos de mejora los salarios y reducción de la probabilidad de desempleo como se comentó en la sección anterior. Esta rentabilidad es privada y apropiable por el propio graduado. Los valores son relativamente altos lo que genera un fuerte incentivo a cursar estudios universitarios. En segundo lugar, una financiación basada fundamentalmente en impuestos en una sociedad donde los alumnos de las clases más desfavorecidas tienen una menor propensión a ir a la Universidad, pero pagan impuestos sobre la renta, implica una transferencia regresiva de recursos de las familias con menores rentas a las familias con mayor renta.

Lógicamente pueden existir fallos de mercado que justifiquen la intervención pública. Un ejemplo en el sector de educación universitaria es la existencia de restricciones de crédito, derivadas de imperfecciones en el mercado de capitales, que impiden que estudiantes brillantes de familias de bajos recursos puedan pagar el precio de la Universidad. En este caso un coste privado alto podría impedir el acceso a la Universidad a estudiantes de familias con restricciones de liquidez a pesar de que los estudios pudieran tener una elevada rentabilidad privada. Asimismo, existen riesgos de la inversión en Educación Universitaria (incremento de la renta futura, tiempo de estudio, etc.) que difícilmente pueden asegurarse sin la intervención del Estado. Otra potencial justificación es la existencia de externalidades que proporciona la Universidad y que provocarían una inversión inferior a la óptima[9]. Lawson (2017) calcula el subsidio óptimo de la matrícula balanceando el impacto de las externalidades positivas por la existencia de una base impositiva mayor (los graduados cobran mayores salarios) con las restricciones de liquidez[10]. La calibración del modelo implica que el subsidio óptimo en las universidades públicas de Estados Unidos debería ser 3.000 dólares mayor por año. Este resultado está basado en la intensidad de las externalidades fiscales y la insensibilidad del subsidio óptimo a las restricciones de liquidez.

En general los beneficios privados superan por mucho a las externalidades pecuniarias. Algunas externalidades no pecuniarias han sido descritas en términos de mayor propensión a votar o mayor nivel de confianza, aunque respecto al comportamiento criminal no hay evidencia. Diris y Oogher (2018) discuten la posibilidad de la existencia de efectos negativos ligados a la utilización de la Educación Universitaria como forma de señalización (sobrecualificación) o la presión social para cursar estudios superiores cuando la opción más apropiada sería un nivel educativo inferior. En cualquier caso, el efecto de las externalidades tiene poco soporte empírico.

En el resto de la sección se comentan diferentes factores que determinan cuál es la combinación óptima entre financiación pública y privada de los estudios universitarios desde la perspectiva de la rentabilidad social y, por tanto, combina elementos de eficiencia y equidad.

---

[9] De hecho, la política pública debería basarse en la rentabilidad marginal social que proporcionan los subsidios.

[10] Este trade-off no tiene en cuenta consideraciones de equidad y redistribución.

### 5.1. Elementos redistributivos de la financiación mediante tributación general

Para analizar el efecto redistributivo de los subsidios a la Educación Terciaria es necesario analizar quién paga y quién se beneficia de los subsidios a la Educación Terciaria. Es evidente que si el contribuyente medio tiene una renta neta a lo largo de la vida (suma descontada de los ingresos laborales futuros menos impuestos y coste privado de la educación) inferior a un graduado universitario medio, un subsidio a la Educación Superior a partir de impuestos generales supone una redistribución de los pobres a los ricos. Los subsidios financiados con la imposición regular siempre generan un trade-off entre eficiencia y equidad: el subsidio aumenta la inversión en capital humano pero, a la vez, genera redistribución reversa. Solo en casos muy especiales y contrafactuales (por ejemplo, si los trabajadores sin estudios universitarios tuvieran una renta tan baja que evitara que pagaran impuestos, o todos los trabajadores tuvieran estudios universitarios, o existieran fuertes complementariedades entre trabajadores poco cualificados y muy cualificados) se podría evitar este trade-off.

García-Peñagolosa y Wälde (2000) construyen un modelo de generaciones sucesivas con dos tipos de trabajadores (poco cualificados y muy cualificados) para analizar el impacto de las distintas formas de financiación de la Educación Superior. Los subsidios financiados con impuestos generales implican siempre un trade-off entre eficiencia y equidad. El subsidio que permite a un número óptimo de trabajadores seguir estudios universitarios resulta en mayor renta para los trabajadores cualificados que para los no cualificados. Un aumento del subsidio aumenta el número de estudiantes universitarios y tiende a igualar las rentas a lo largo de la vida, pero en ese caso los trabajadores están sobrecualificados lo que produce una pérdida de output potencial. Los créditos estatales para estudiar, el impuesto para graduados y los sistemas de préstamos contingentes en la renta producen el mismo resultado si los ingresos futuros son ciertos. Si existe incertidumbre los créditos no contingentes suponen siempre el repago[11] mientras que los sistemas de repago contingente solo lo producen si el estudiante tiene éxito en el empleo. El impuesto a los graduados aumenta el pay-off para los graduados que no tienen éxito y lo reduce para los que tienen éxito lo que

[11] La cuestión de los impagos de este tipo de créditos es también un tema que ha recibido cierta atención de la literatura.

produce un sistema de seguro parcial que, en caso de agente con aversión al riesgo, aumenta la eficiencia. Los préstamos contingentes a la renta producen un resultado similar al de los impuestos para graduados, pero implica algún tipo de financiación tributaria para subsidiar a los estudiantes que no tuvieron éxito en sus empleos. De esta forma genera una redistribución reversa que está ausente en la modalidad de impuestos a los graduados[12].

Es difícil encontrar análogos empíricos para estudiar quiénes se benefician y quiénes pierden con la financiación de subsidios a partir de impuestos generales. Diris y Ooghe (2018) proponen calcular el coste fiscal neto desde una doble perspectiva: el punto de vista de los padres (visión de corte transversal) y el punto de vista de los estudiantes (visión longitudinal). Desde la perspectiva de los estudiantes se produce regresividad por construcción: los trabajadores con alta cualificación reciben los subsidios de la Educación Superior mientras que todos los contribuyentes futuros pagan esos subsidios. Los datos e Diris y Ooghe (2018) muestran que el coste neto fiscal desde esta perspectiva es casi cero para trabajadores de alta y baja cualificación en Chile porque los ingresos de los poco cualificados son tan bajos que no pagan impuestos. El coste fiscal neto de los países escandinavos es el más alto. Desde la perspectiva de los padres el resultado depende de la transferencia intergeneracional de la educación (probabilidad relativa que el hijo de unos padres universitarios, frente a padres con nivel educativo inferior, acuda a la Universidad), del tamaño del subsidio y de la progresividad del sistema impositivo. Los cálculos de Diris y Ooghe (2018) muestran que los países nórdicos (Suecia, Dinamarca, Noruega) tienen los sistemas más regresivos desde la perspectiva paterna (elevados subsidios y baja progresividad impositiva) mientras Países Bajos, Irlanda y Chile tiene los sistemas menos regresivos.

### 5.2. Restricciones de crédito

Otra justificación para la utilización de subsidios en la Educación Superior es la existencia de imperfecciones en el mercado de capitales. La existen-

[12] Johnson (1984) utiliza un modelo con fuertes complementariedades entre trabajadores cualificados y no cualificados para argumentar que no existe problema de redistribución reversa. Usando este argumento Borck y Wimbersky (2014) señalan que los pobres pueden preferir el sistema de impuestos y subsidios al resto de mecanismos de financiación.

cia de riesgos no asegurables y las restricciones de crédito pueden generar pérdidas de eficiencia. La selección adversa y el riesgo moral pueden producir un nivel de aseguramiento insuficiente en mercados privados. Algunos riesgos son difíciles de asegurar sin la intervención del Gobierno. Este es el caso del riesgo asociado a la renta a lo largo de la vida o el riesgo de impago del crédito. La falta de aseguramiento (por ausencia de colaterales o avalistas) puede provocar que los mercados no produzcan suficiente crédito y estudiantes con pocos recursos, pero académicamente preparados, no puedan acceder a la Universidad.

Contrastar empíricamente el impacto de las restricciones de crédito sobre el acceso a la Educación Terciaria es complejo. La evidencia de que existe una correlación entre la renta familiar y acceso a la Universidad no es evidencia de la existencia de restricciones de crédito. Es posible que alumnos de familias con nivel alto de renta no solo tengan mejor acceso a los mercados de crédito, sino que también tengan mejor preparación académica, mayores habilidades o preferencias más fuertes por asistir a la Universidad. Por este motivo los contrastes empíricos basados en medidas indirectas de acceso al crédito han generado una literatura con resultados muchas veces inconsistentes.

Solis (2017) presenta el primer análisis causal del impacto del acceso al crédito sobre el acceso a la educación universitaria. Solis (2017) utiliza dos programas de créditos universitarios introducidos en Chile que proporcionan financiación a estudiantes que obtienen una puntuación superior a un cierto nivel en el test nacional de admisión a la Universidad. Comparando los estudiantes que quedaron justo por encima del nivel de elegibilidad con los que quedaron justo por debajo, Solis (2017) propone un diseño de discontinuidad que evita los problemas de variables omitidas y selección muestral que afectan a la gran mayoría de los estudios sobre el efecto de las restricciones de crédito sobre el acceso a la Universidad. Los resultados muestran que los dos programas de créditos aumentaron la probabilidad de matriculación inmediata en 18 puntos porcentuales lo que supone doblar la tasa de matriculación del grupo que se quedó justo por debajo de la puntuación para acceder a los créditos. La mejora fue mayor para los estudiantes en el quintil más bajo de la distribución (un 140%) frente al 15% de los alumnos justo debajo del límite para acceder a los préstamos. Finalmente, Solis (2017) muestra que el acceso a estos programas (estudiantes por encima del límite de puntuación) elimina el gap de matriculación entre familias de diferente nivel de renta, mientras que en el grupo que quedó justo por debajo del límite la

tasa de graduación de los estudiantes del quintil más alto de la renta es el doble que la tasa entre los estudiantes del quintil de renta más pobre.

Azmat y Simion (2020) analizan tres reformas realizadas en el Reino Unido que implicaron un aumento de los precios públicos, la implantación de un sistema de préstamos contingentes a la renta y el incremento de las becas para alumnos de familias con renta baja. La evaluación de Azmat y Simion (2020) muestra que, a pesar de ese cambio tan sustancial en la financiación universitaria, no tuvo un efecto relevante sobre la matriculación y el pequeño efecto observado implicó una pequeña reducción de la matriculación de los estudiantes de familias más favorecidas, reduciendo el gap de matriculación entre los diferentes niveles socioeconómicos.

### 5.3. Efectos comportamentales en las decisiones

Una línea relevante de la literatura argumenta que una parte de la diferencia entre la tasa de acceso a la Educación Superior de las familias de renta alta frente a las de menor renta se explica por elecciones de los estudiantes que no son acordes al modelo clásico de decisión económica. La falta de información sobre el coste de la Universidad, o la capacidad del alumno para seguir los cursos en una institución muy selectiva, puede reducir la inversión educativa de los estudiantes de nivel socioeconómico bajo. El sesgo del presente (present bias), la aversión a la deuda, o el exceso de peso de la rutina y las opciones por defecto en las decisiones es particularmente importante entre los estudiantes con baja renta o en decisiones complejas e inciertas como es el caso del proceso de admisión o el funcionamiento del sistema de ayudas y becas. En esta situación pequeños cambios pueden producir grandes modificaciones de decisiones. Dynarski et al. (2021) usan un experimento randomizado para analizar si reduciendo la incertidumbre sobre la ayuda financiera se puede modificar la decisión de los estudiantes con un nivel de renta bajo. Los investigadores identificaron 2.000 estudiantes con alto rendimiento y baja renta en colegios públicos. El tratamiento se randomizó al nivel del colegio con la mitad de los mismos asignados al grupo tratado y la otra mitad al grupo de control. El tratamiento consistió en el envío de cartas personalizadas para incentivarles a aplicar a la Universidad de Michigan (muy selectiva) y con la promesa de cuatro años de matrícula gratuita para aquellos que fueran admitidos. Por tanto, la beca HAIL proporciona una garantía anticipada y no condicionada de gratui-

dad de la matrícula reduciendo la incertidumbre sobre el coste de la educación universitaria al garantizar el pago de la matrícula. La naturaleza condicional de la ayuda evita tener que rellenar formularios para verificar que el estudiante está cualificado para la beca.

Dynarski et al. (2021) encuentran un impacto grande de la intervención sobre las tasas de solicitud y matriculación a la Universidad de Michigan. La tasa de solicitud pasó del 26% en el control al 68% en el grupo tratado. La tasa de matriculación incrementó desde el 12% de los controles hasta el 27% de los tratados. Un 25% del efecto sobre la matriculación fue causado por estudiantes que no habrían asistido a la Universidad. El resto son alumnos que habrían ido a un "community colleg" o a una Universidad menos selectiva sin afectar a universidades tan o más selectivas que Michigan.

### 5.4. La economía política de la financiación de las Universidades

Las secciones anteriores muestran que un exceso de subsidios públicos o unos precios privados muy altos pueden producir efectos regresivos o infra-inversión en educación universitaria respectivamente. En ambos casos un desplazamiento de la financiación hacia sistemas de créditos contingentes en los ingresos futuros o impuestos para los graduados supondrían mejoras. Por ejemplo, en el caso de los países nórdicos, cuyo elevado nivel de subsidios financiado con impuestos para el conjunto de la población lo convierten en un sistema muy regresivo, el movimiento hacia créditos contingentes a los salarios futuros o impuestos para graduados reduce la regresividad (impide que graduados que no se benefician de la educación universitaria acaben financiando a los que se benefician) y proporciona un seguro reduciendo el azar moral durante los estudios. En los países con contribuciones privadas muy elevadas las restricciones de crédito pueden dejar a buenos estudiantes de familias con pocos recursos fuera de la Educación Superior. De nuevo, un sistema de créditos condicionados o de impuestos para graduados proporcionaría claras ventajas (por ejemplo, un seguro frente al desconocimiento del impacto de la Educación Superior sobre la renta futura), aunque en algunos casos estas ventajas tienen que balancearse con el potencial impacto regresivo del cambio.

Ante la evidencia de las ventajas de aumentar el peso de los préstamos condicionados y los impuestos a los graduados frente a los procedimientos tradicionales (subsidios y financiación privada) que supone mejoras de eficiencia

y equidad es difícil entender por qué tan pocos países se han movido en esta dirección. Entre las excepciones se encuentran Países Bajos, Hungría, Nueva Zelanda, Reino Unido o Australia. Estos dos últimos junto con un aumento del precio de la matrícula establecieron un sistema de créditos contingentes en la renta futura[13]. Azmat y Simion (2020) muestran que, en el caso del Reino Unido, el paso progresivo de un sistema de matrícula gratuita pagada con impuestos a un sistema con precios de matrícula altos complementado con becas basadas en la renta de los estudiantes y créditos contingentes no tuvo ningún efecto negativo en la participación de los estudiantes de familias desfavorecidas, con un efecto muy limitado sobre la elección de grados y Universidad. En el caso de Australia Chapman y Ryan (2005) muestran que el paso al sistema de créditos contingentes no redujo la participación en el sistema de Educación Superior, en general ni entre los estudiantes de familias con menos recursos.

A pesar de que la evidencia empírica muestra que el movimiento hacía sistemas basados en créditos contingentes no produce una caída de la matriculación en los grupos de estudiantes económicamente más desfavorecidos, la economía política de la financiación universitaria limita la extensión de estos sistemas. De hecho, muchos partidos políticos de algunos países que incrementaron el coste de la matrícula después de la crisis financiera y sus consecuencias sobre los presupuestos públicos están proponiendo la vuelta a sistema con subsidios casi totales de la matrícula universitaria. El trabajo seminal de Fernández y Rogerson (1995) muestra como una coalición de la clase media y alta puede preferir mantener los subsidios altos de forma que las clases menos favorecidas no puedan acceder a la Universidad (por restricciones de crédito), pero paguen, vía impuestos, los subsidios de los que se benefician los estudiantes universitarios[14].

---

[13] En el caso del Reino Unido también se incrementó sustancialmente el apoyo financiero a estudiantes de familias con bajos recursos y el repago de los créditos entre los más pobres era incluso menor en el nuevo sistema que con los precios de matrícula menores anteriores a la reforma.

[14] Borck y Wimbersky (2014) construyen un modelo con consumidores con aversión al riesgo y heterogeneidad de la renta y argumentan que un apoyo mayoritario al sistema de créditos contingentes o pago por los graduados es más probable si aumenta la aversión al riesgo o se reduce la desigualdad.

### 5.5. La financiación universitaria en España

España pertenece al grupo de países que tienen un nivel de subsidio público elevado de las matrículas universitarias. El sistema proporciona matrícula gratuita a los estudiantes de familias desfavorecidas con un nivel de renta límite que depende del tamaño de la familia. Por encima de ese límite los estudiantes pagan entre el 10% y el 15% del coste del coste de los estudios universitarios. En 2012, y ante la presión de los mercados sobre la deuda pública española, el Gobierno realizó ajustes presupuestarios que afectaron a los recursos para financiar las Universidades públicas. El Ministerio de Educación, Cultura y Deportes propuso de la CC.AA. un aumento de las tasas universitarias limitando el máximo del precio a pagar por los alumnos al 25% del coste de la plaza universitaria. Además, aumentó considerablemente el precio de las segundas, terceras y cuartas matrículas. Cada CC.AA. decidió en qué proporción aumentaba los precios públicos. Mientras algunas (Galicia o Asturias) lo aumentaron muy poco, Cataluña y Madrid optaron por un aumento muy sustancial. En el caso de Cataluña el aumento de precios fue progresivo graduado a partir de las denominadas becas equidad. El pago completo de la matrícula solo se producía en el caso de familias con ingresos superiores a 67.000 euros[15]. Para rentas inferiores pero superiores a la que daba acceso a la gratuidad de la matrícula las becas equidad proporcionaban descuentos mayores cuanto menor fuera la renta. García-Montalvo (2020) analiza el impacto de este nuevo sistema de financiación de las matrículas universitarias y no encuentra evidencia de ningún impacto adverso del cambio de política en la tasa de abandono. También analiza el impacto distributivo de la reforma, mostrando que unos precios públicos progresivos, cuidadosamente diseñados, pueden generar ingresos adicionales para las universidades sin ningún impacto negativo en las tasas de abandono de estudiantes de nivel socioeconómico bajo[16]. Por tanto, en este caso la reducción del subsidio centrada en los estudiantes de familias con mayor poder adquisitivo no tuvo ningún impacto sobre las tasas de matriculación y redujo la regresividad del sistema desde la perspectiva del estudiante.

---

[15] Para una familia de referencia con cuatro miembros.

[16] La necesidad de solicitar la beca general antes de solicitar la beca equidad hizo que bastantes estudiantes que no solicitaban tradicionalmente la beca general por considerar que la renta familiar era superior a los límites, obtuvieran la exención de matrícula.

Beneito et al. (2018) se centran en el aumento del coste de las segundas y terceras matrículas y muestran que el aumento del precio de la matrícula incrementó el esfuerzo de los estudiantes. Montalbán (2022) muestra que las becas basadas en necesidad no tienen impacto en los resultados académicos si los requisitos en términos de rendimiento académico son los habituales en la mayoría de los países. La provisión de un beca-salario pequeña tienen un impacto muy superior sobre los resultados académicos y la probabilidad de completar los estudios universitarios cuando van asociadas a unos requerimientos académicos mínimos exigentes, aunque este efecto se produce solo si la comunicación de la beca se produce a principio del curso académico. Aumentar el tamaño de la ayuda no mejora los resultados académicos. Por su parte, incrementar los requisitos académicos no produce un aumento del abandono de la Universidad.

Por su parte Cabrales et al. (2019) analizan el paso del sistema actual de financiación de las universidades en España a un procedimiento basado en créditos contingentes en la renta futura (artículo que se comentó ya en el capítulo 1 de indicadores de este manual). Las simulaciones se basan en un laboratorio virtual donde se analiza el efecto distribucional del cambio en la financiación. El análisis utiliza una simulación de los ingresos a lo largo de la vida de los graduados universitarios que reproduce la dinámica del empleo y los ingresos calculados a partir de los datos administrativos de la Seguridad Social. Sobre esta base se calcula el efecto de la introducción de un sistema de créditos contingentes sobre personas en diferentes posiciones de la distribución de la renta y el impacto para el Estado. Los resultados muestran que la estructura es muy progresiva en todas las especificaciones: el 25% superior de la distribución de la renta acaba pagando casi la totalidad del coste de la matrícula y el 10% más bajo prácticamente no paga nada. Además, la proporción que representa el subsidio sobre el coste de la educación universitaria es entre 16 y 56 puntos porcentuales menos que en el sistema actual.

## 6. EL GOBIERNO DE LA UNIVERSIDAD

Gestionar la Universidad es complicado. La mayor parte de los países occidentales reconocen, de una u otra forma la autonomía universitaria en gestión sobre todo académica. Eso genera la libertad necesaria para crear innovación por parte de los profesores y debería estar vinculada a un deseable marco de financia-

ción por objetivos, que inevitablemente lleva en la realidad a la competencia por fondos, que permita al regulador generar incentivos que optimicen la provisión de docencia e investigación (Del Rey, 2001). Como consecuencia, se podrían producir procesos, más o menos virtuosos, que favoreciesen la concentración de investigación en las universidades más productivas (De Fraja y Valbonesi, 2012), la mayor productividad de los académicos más activos (Checchi et al., 2021; de Fraja et al., 2019) o favorecer la admisión de estudiantes (Bouchard St-Amant et al., 2020). El objetivo de esta sección es analizar quién y cómo se debe liderar este proceso.

### 6.1. La gestión del talento

Becker y Gerhart (1996) ponen de manifiesto la importancia que tienen la gestión y selección del personal en los resultados económicos y la supervivencia de las instituciones. De hecho, la literatura ha evolucionado hacía la consideración de la gobernanza de los recursos humanos de cualquier empresa de forma conjunta, como un sistema que va más allá de la selección, remuneración y comportamiento de la cabeza la organización, incluyendo al conjunto de la misma, y cómo esta se traslada a su rendimiento (Bloom y Van Reenen, 2007; Boon et al., 2019).

Esta premisa no es baladí. Mccormack et al. (2014) utilizan la metodología propuesta por Bloom y Van Reenen (2007), y llegan a la conclusión de que la calidad en la gestión de la Universidad tiene un efecto importante sobre los rendimientos tanto en docencia como en investigación (que no siempre están correlacionadas). A partir de una bolsa de 248 universidades inglesas, los autores utilizan un total de 17 características de la calidad de la gestión en la Universidad divididos en cuatro grupos; operaciones, supervisión, objetivos e incentivos, evaluados a través de un cuestionario y contrastados contra el Research Excellence Framework (REF), que evalúa cada 5 años la investigación de los departamentos ingleses, la encuesta satisfacción de los estudiantes y rankings académicos. Los resultados son robustos al controlar por el tamaño de los departamentos, el número de estudiantes a los que atienden, etcétera. Un primer resultado a señalar es que las características individuales de los líderes universitarios no tienen un efecto significativo en el funcionamiento de la Universidad, por el contrario, sí encuentran una relación positiva entre la calidad de la gestión descentralizada, que afecta a los cometidos menos estratégicos, como los de Recursos Humanos o los departamentos académicos, etc.

### 6.2. El talento del gestor

Frente al enfoque holístico anterior, Goodall (2009) se centra en los rectores de Universidad. Proporciona fundamentos empíricos a través del análisis de un conjunto de datos de 400 rectores y directores de las principales escuelas de negocios estadounidenses y británicas. Su hipótesis argumenta que las universidades orientadas a la investigación experimentan un progreso sustancial cuando individuos de renombre en el ámbito académico asumen cargos de responsabilidad. Asimismo, plantea la premisa de que los académicos deberían desempeñar papeles directivos en otras organizaciones, incluyendo entidades gubernamentales encargadas de financiar instituciones educativas, como una manera de respaldar y respaldarse mutuamente. Goodall (2009) encuentra correlación positiva entre la excelencia en investigación y los logros organizativos, proponiendo que el nivel apropiado de competencia académica en un presidente variará según la posición de la Universidad en términos de su investigación y su posición en los rankings, así como sus objetivos futuros. También desarrolla un modelo empírico que emplea como variable dependiente los resultados del REF de 55 universidades británicas, llevados a cabo en los años 1992, 1996 y 2001, y como independiente las citas acumuladas a lo largo de la carrera de 157 rectores de universidades británicas en su muestra. En definitiva, hay una amplia literatura que soporta teórica y causalmente la base de su análisis; buenos gestores generan buenos resultados que persisten en el tiempo (Lazear et al., 2015; Manthei et al., 2023).

Gerashchenko, (2022) utiliza la misma metodología que Goodall (2009) para el universo de universidades rusas, pero las amplía de forma interesante con variables independientes al tener en cuenta algunos aspectos diferenciales; mejora las variables bibliométricas, caracteriza las universidades según su especialización y diferencia entre aquellos rectores que son nombrados por su capacidad de ejercer gestión y aquellos que son seleccionados políticamente. Sus resultados, que no contienen causalidad, señalan que ser un investigador de calidad tiene un efecto a largo plazo en las instituciones, y sobre la productividad del área del propio rector. Para Gerashchenko (2022) es la calidad del rector y no la forma del nombramiento la que determina el resultado final.

El sistema implantado en la LOSU, alejado de la profesionalidad o de la selección por méritos académicos, parece que está lejos de los resultados vistos en los artículos anteriores y deja espacio para posteriores mejoras en el ámbito de la gobernanza en España.

## EJERCICIOS

Ejercicio 1. *¿Por qué la financiación de la educación universitaria mediante subsidios públicos es regresiva por construcción en el enfoque del estudiante (análisis longitudinal)?*

Ejercicio 2. *Explica por qué si existe un elevado grado de complementariedad entre las skills de trabajadores de baja cualificación y trabajadores universitarios se puede romper el trade-off entre eficiencia y equidad generado por la financiación de la Educación Terciaria basada en subsidios públicos.*

Ejercicio 3. *(García Peñagolosa y Wälde, 2000). Imagina una economía con N individuos que viven 3 periodos. En el primer periodo todos los niños completan la educación obligatoria. En el segundo eligen si trabajan inmediatamente o invierten en Educación Terciaria. El coste de la Educación Terciaria es E. Todo el consumo se produce en el tercer periodo. Los padres legan a sus hijos una herencia g con una distribución f(g). La economía produce un bien de consumo con la tecnología Y=F(H,L) donde H son trabajadores universitarios y L no universitarios. Por tanto, N=H+L. Hay competencia perfecta en los mercados de productos y factores. El gobierno solo puede subsidiar la educación e ingresar impuestos, manteniendo la restricción presupuestaria*

a) *Si el capital humano no se puede usar como colateral y no hay subsidiación de la educación, ¿quiénes podrán acceder a la educación universtiaria?*

b) *Si el mercado de capitales es perfecto y no hay subsidiación, ¿cuál es el número de universitarios en equilibrio? Describe la renta neta a lo largo de la vida de un universitario y un trabajador no universitario. Usa los principios comentados en el capítulo para construir la rentabilidad interna de la educación.*

c) *Supongamos que el objetivo del Gobierno es conseguir el número óptimo de trabajadores con educación universitaria, $H^*$, y para esto utiliza subsidios públicos pero el mercado de capitales no acepta el capital humano como colateral. ¿Podrá conseguir su objetivo? Explica el impacto distribucional del equilibrio obtenido.*

Ejercicio 4. *Utiliza los datos de la encuesta del INE sobre inserción laboral de los universitarios (los microdatos se pueden conseguir en el siguiente enlace https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176991&menu=resultados&idp=1254735976597#!tabs-1254736195339*

a) *Calcula la rentabilidad interna de la educación universitaria. Utiliza los supuestos que creas necesarios para poder realizar el cálculo.*

b) *Analiza los resultados por campos de estudio y por género.*

c) *Algunos estudios (por ejemplo https://www.fbbva.es/wp-content/uploads/2023/05/FBBVA-Ivie_U-ranking_Informe_Insercion_laboral_2023_def.pdf) utilizan los datos de la Seguridad Social para analizar los resultados de la inserción laboral de los universitarios. ¿Qué ventajas e inconvenientes tienen los datos de la Seguridad Social frente al INE en el análisis de la rentabilidad de la educación de los universitarios?*

Ejercicio 6. *Utilizando el catálogo de datos del Ministerio de Universidades (https://www.universidades.gob.es/catalogo-de-datos/):*

a) *En el año 2018 Andalucía pasó a bonificar el cien por cien de las matrículas a aquellos estudiantes que habían aprobado. Analice el efecto que sobre la demanda de qué estudios universitarios tuvo en esa Comunidad Autónoma en comparación con el resto.*

b) *¿Existe en España la inflación de la nota de acceso a la Universidad? Analícelo empíricamente y justifique su posición.*

## BIBLIOGRAFÍA


Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A. y Sapir, A. (2008). *Higher Aspirations: An Agenda for Reforming European Universities* (5; Bruegel Blueprint Series). Bruegel.

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A. y Sapir, A. (2010). The governance and performance of universities: evidence from Europe and the US. *Economic Policy*, *25*(61), 7-59. https://doi.org/10.1111/j.1468-0327.2009.00238.x

Aina, C., Baici, E., Casalone, G. y Pastore, F. (2018). *Delayed Graduation: A Survey* (11421; Discussion Papers). www.iza.org

Aina, C., Baici, E., Casalone, G. y Pastore, F. (2022). The determinants of university dropout: A review of the socio-economic literature. *Socio-Economic Planning Sciences* (Vol. 79). Elsevier Ltd. https://doi.org/10.1016/j.seps.2021.101102

Altonji, J. G., Arcidiacono, P. y Maurel, A. (2016). The analysis of field choice in college and graduate school: Determinants and wage effect. *Handbook of the Economics of Education* (Vol. 5, 305-396). Elsevier.

Anelli, M. y Peri, G. (2015). Gender of siblings and choice of college major. *CESifo Economic Studies*, *61*(1), 53-71. https://doi.org/10.1093/cesifo/ifu028

Arcidiacono, P., Hotz, V. J. y Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, *166*(1), 3-16. https://doi.org/https://doi.org/10.1016/j.jeconom.2011.06.002

Autor, D. H., Donohue III, J. J. y Schwab, S. J. (2006). The costs of wrongful-discharge laws. *The Review of Economics and Statistics*, *88*(2), 211-231.

Autor, D. H., Levy, F. y Murnane, R. J. (2003). The skill content of recent technological change: an empirical exploration. *Q. J. Econ.*, *118*. https://doi.org/10.1162/003355303322552801

Ayllón, S., Valbuena, J. y Plum, A. (2022). Youth Unemployment and Stigmatization Over the Business Cycle in Europe*. *Oxford Bulletin of Economics and Statistics*, *84*(1), 103-129. https://doi.org/https://doi.org/10.1111/obes.12445

Azmat, G. y Simion, S. (2020). Charging for Higher Education: Estimating the Impact on Inequality and Student Outcomes. The B.E. *Journal of Economic Analysis & Policy, 21*(1), 175-239.

Backes-Gellner, U., Herz, H., Kosfeld, M. y Oswald, Y. (2021). Do preferences and biases predict life outcomes? Evidence from education and labor market entry decisions. *European Economic Review*, *134*, 103709. https://doi.org/https://doi.org/10.1016/j.euroecorev.2021.103709

Becker, B. y Gerhart, B. (1996). The Impact of Human Resource Management on Organizational Performance: Progress and Prospects. *The Academy of Management Journal*, *39*(4), 779-801. https://about.jstor.org/terms

Behr, A., Giese, M., Teguim Kamdjou, H. D. y Theune, K. (2020). Dropping out of university: A literature review. *Review of Education*, *8*(2), 614-652. https://doi.org/10.1002/rev3.3202

Beneito, P., Boscá, J. E. y Ferri, J. (2018). Tuition fees and student effort at university. *Economics of Education Review, 64*, 114-128. https://doi.org/10.1016/j.econedurev.2018.03.012

Betts, J. R. y Mcfarland, L. L. (1995). Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments. *The Journal of Human Resources*, *30*(4), 741-765. https://about.jstor.org/terms

Bhuller, M., Mogstad, M. y Salvanes, K. G. (2017). Life-cycle earnings, education premiums, and internal rates of return. *Journal of Labor Economics, 35*(4), 993-1030.

Bleemer, Z. (2022). Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209*. *The Quarterly Journal of Economics*, *137*(1), 115-160. https://doi.org/10.1093/qje/qjab027

Bloom, N. y Van Reenen, J. (2007). Measuring and Explaining Management Practices Across Firms and Countries*. *The Quarterly Journal of Economics*, *122*(4), 1351-1408. https://doi.org/10.1162/qjec.2007.122.4.1351

Boon, C., Den Hartog, D. N. y Lepak, D. P. (2019). A Systematic Review of Human Resource Management Systems and Their Measurement. *Journal of Management*, *45*(6), 2498-2537. https://doi.org/10.1177/0149206318818718

Bordon, P. y Fu, C. (2015). College-Major Choice to College-Then-Major Choice. *Review of Economic Studies*, *82*(4), 1247-1288. https://doi.org/10.1177/0149206318818718

Borck, R. y Wimbersky, M. (2014). Political economics of higher education finance. *Oxford Economic Papers, 66*(1), 115-139.

Bouchard St-Amant, P.-A., Brabant, A.-N. y Germain, É. (2020). University funding formulas: an analysis of the Québec reforms and incentives. *Canadian Journal of Higher Education*, *50*(1), 1-27.

Braga, M., Paccagnella, M. y Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, *41*, 71-88. https://doi.org/10.1016/j.econedurev.2014.04.002

Bresnahan, T. F., Brynjolfsson, E. y Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: firm-level evidence. *Q. J. Econ.*, *117*. https://doi.org/10.1162/003355302753399526

Cabrales, A., Güel, M., Madera, R. y Viola, A. (2019). Income contingent university loans: Policy design and an application to Spain, *Economic Policy, 34*, 99, 479-521.

Cappellari, L. y Lucifora, C. (2009). The "Bologna Proces" and college enrollment decisions. *Labour Economics*, *16*(6), 63-647. https://doi.org/10.1016/J.LABECO.2009.08.009

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5), 1127-1160. https://doi.org/10.1111/1468-0262.00237

Cardoso, A. R., Loviglio, A. y Piemontese, L. (2016). Misperceptions of unemployment and individual labor market outcomes. *IZA Journal of Labor Policy*, *5*(1). https://doi.org/10.1186/s40173-016-0069-6

Carlsson, M., Finseraas, H., Midtbøen, A. H. y Rafnsdóttir, G. L. (2021). Gender bias in academic recruitment? Evidence from a survey experiment in the Nordic region. *European Sociological Review*, *37*(3), 399-410.

Carruthers, C. K., Fox, W. F. y Jepsen, C. (2023). What Knox Achieved: Estimated Effects of Tuition-free Community College on Attainment and Earnings. *Journal of Human Resources*, 11359R2. https://doi.org/10.3368/jhr.1220-11359R2

Chapman, B. y Ryan, C. (2005). The access implications of income-contingent charges for higher education: lessons from Australia. *Economics ofeEducationrReview, 24*(5), 491-512. https://doi.org/10.1016/j.econedurev.2004.08.009

Checchi, D., De Fraja, G. y Verzillo, S. (2021). Incentives and Careers in Academia: Theory and Empirical Analysis. *The Review of Economics and Statistics*, *103*(4), 786-802. https://doi.org/10.1162/rest_a_00916

Chetty, R., Deming, D. y Friedman, J. (2023). Diversifying Society's Leaders? The Causal Effects of Admission to Highly Selective Private Colleges. In *NBER Working Paper Series* (31492). https://doi.org/10.3386/w31492

Coenen, J., Borghans, L. y Diris, R. (2021). Personality traits, preferences and educational choices: A focus on STEM. *Journal of Economic Psychology*, *84*, 167-4870. https://doi.org/10.17632/vt5v7c27kw.1

Dale, S. B. y Krueger, A. B. (2002). Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservable*. *The Quarterly Journal of Economics*, *117*(4), 1491-1527. https://doi.org/10.1162/003355302320935089

Dale, S. B. y Krueger, A. B. (2014). Estimating the effects of college characteristics over the career using administrative earnings data. *Journal of Human Resources*, *49*(2), 323-358. https://doi.org/10.3368/jhr.49.2.323

De Fraja, G., Facchini, G. y Gathergood, J. (2019). Academic salaries and public evaluation of university research: Evidence from the UK Research Excellence Framework. *Economic Policy*, *34*(99), 523-583. https://doi.org/10.1093/epolic/eiz009

De Fraja, G. y Valbonesi, P. (2012). The design of the university system. *Journal of Public Economics*, *96*(3-4), 317-330. https://doi.org/10.1016/j.jpubeco.2011.10.005

Del Rey, E. (2001). Teaching versus research: A model of state university competition. *Journal of Urban Economics*, *49*(2), 356-373. https://doi.org/10.1006/juec.2000.2193

Diris, R. y Ooghe, E. (2018). The economics of financing higher education. *Economic Policy, 33*(94), 265-314. https://doi.org/10.1093/epolic/eiy003

Denning, J. T. (2017). College on the cheap: Consequences of community college tuition reductions. *American Economic Journal: Economic Policy*, *9*(2), 15–-188. https://doi.org/10.1257/pol.20150374

Dynarski, Susan, CJ Libassi, Katherine Michelmord y Stephanie Owe. (2021). Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Student. *American Economic Review, 111*(6): 1721-56

Durazzi, N. (2019). The political economy of high skills: higher education in knowledge-based labour markets. *Journal of European Public Policy*, *26*(12), 1799-1817. https://doi.org/10.1080/13501763.2018.1551415

Fernández, R. y Rogerson, R. (1998). Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform. *The American Economic Review, 88*(4), 813-833. http://www.jstor.org/stable/117007

García Hombrados, J. (2022, December 7). *Limitaciones a la atracción de talento joven en las universidades públicas españolas*. Nadaesgratis. https://nadaesgratis.es/jorge-garcia/limitaciones-a-la-atraccion-de-talento-joven-en-las-universidades-publicas-espanolas

García-Montalvo, J. (2020). El impacto de la introducción de subvenciones progresivas en la educación superior: el caso de Cataluña. *Papeles de Economía Española*, *166*, 123-142.

García-Penalosa, C. y Wälde, K. (2000). Efficiency and equity effects of subsidies to higher education. *Oxford Economic Papers, 52*(4), 702-722. https://doi.org/10.1093/oep/52.4.702

Gerashchenko, D. (2022). Academic leadership and university performance: do Russian universities improve when they are led by top researchers? *Higher Education*, *83*(5), 110-1123. https://doi.org/10.1007/s10734-021-00732-5

Ginther, D. K. y Kahn, S. (2021). Women in academic economics: Have we made progress? *AEA Papers and Proceedings*, *111*, 138-142.

Goodall, A. H. (2009). Highly cited leaders and the performance of research universities. *Research Policy*, *38*(7), 1079-1092. https://doi.org/10.1016/j.respol.2009.04.002

Grossman, J., Tomkins, S., Page, L. C. y Goel, S. (2023). *The Disparate Impacts of College Admissions Policies on Asian American Applicants* (31527; NBER Working Papers). National Bureau of Economic Research. https://doi.org/10.3386/w31527

Hanushek, E. A. (2016). Will more higher education improve economic growth? *Oxford Rev. Econ. Pol.*, *32*. https://doi.org/10.1093/oxrep/grw025

Hastings, J. S., Neilson, C. A. y Zimmerman, S. D. (2013). *Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile* (19241; NBER Working Paper Series). http://www.justinehastings.com/images/downloads/HNZ_Chile_Appendix_2013a.pdf.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *Review of Economics and Statistics*, *91*(4), 717-714. https://doi.org/10.1162/rest.91.4.717

Hoxby, C. M. (2009). The Changing Selectivity of American Colleges. *Journal of Economic Perspectives*, *23*(4), 95-118. https://doi.org/10.1257/jep.23.4.95

Hoxby, C. Minter. (2004). *College choice : the economics of where to go, when to go, and how to pay for it*. University of Chicago Press.

Johnson, G.E. (1984). Subsidies for Higher Educatio. *Journal of Labour Economics, 2,* 303-18.

Kamis, R., Pan, J. y Seah, K. K. (2023). Do college admissions criteria matter? Evidence from discretionary vs. grade-based admission policies. *Economics of Education Review*, *92*. https://doi.org/10.1016/j.econedurev.2022.102347

Kasper, H. T. (2003). The changing role of community college. *Occupational Outlook Quarterly*, 14-21.

Kehm, B. M., Larsen, M. R. y Sommersel, H. B. (2020). Student dropout from universities in Europe: A review of empirical literature. *Hungarian Educational Research Journal*, *9*(2), 147-164. https://doi.org/10.1556/063.9.2019.1.18

Kirkeboen, L. J., Leuven, E. y Mogstad, M. (2016), Field of study, earnings, and self-selection. *The Quarterly Journal of Economics, 131*(3), 1057-1111.

Kurban, H., Gallagher, R. M. y Persky, J. J. (2015). Demographic changes and education expenditures: A reinterpretation. *Economics of Education Review*, *45*, 103-108. https://doi.org/10.1016/j.econedurev.2015.01.001

Lassibille, G., Lucía, M. y Gómez, N. (2008). Why do higher education students drop out? Evidence from Spain. *Education Economics*, *16*(1). https://doi.org/10.1080/09645290701523267ï

Lazear, E. P., Shaw, K. L. y Stanton, C. T. (2015). The Value of Bosses. *Journal of Labor Economics*, *33*(4), 823-861. https://doi.org/10.1086/681097

Ma, L., Li, X., Zhu, Q. y Ye, X. (2023). College-major choice to college-then-major choice: Experimental evidence from Chinese college admissions reforms. *Economics of Education Review*, *94*, 102380. https://doi.org/10.1016/J.ECONEDUREV.2023.102380

Manthei, K., Sliwka, D. y Vogelsang, T. (2023). Talking About Performance or Paying for It? A Field Experiment on Performance Reviews and Incentives. *Management Science*, *69*(4), 2198-2216. https://doi.org/10.1287/mnsc.2022.4431

Marginson, S. (2016). The worldwide trend to high participation higher education: dynamics of social stratification in inclusive systems. *Higher Education*, *72*(4), 413-434. https://doi.org/10.1007/s10734-016-0016-x

Martínez-Morales, I. y Marhuenda-Fluixá, F. (2020). Vocational education and training in Spain: steady improvement and increasing value. *Journal of Vocational Education & Training*, *72*(2), 209–227. https://doi.org/10.1080/13636820.2020.1729840

Mccormack, J., Propper, C. y Smith, S. (2014). Herding Cats? Management and University Performance. *The Economic Journal*, *124*(578), F534-F564. https://doi.org/10.1111/ECOJ.12105

Modena, F., Rettore, E. y Tanzi, G. M. (2018). *The effect of grants on university drop-out rates: evidence on the Italian case* (1193; Temi Di Discussione). https://ssrn.com/abstract=3415672

Montalbán, J. (2023). Countering Moral Hazard in Higher Education: The Role of Performance Incentives in Need-Based Grant*. *Economic Journal*, *133*(649), 355-389. https://doi.org/10.1093/ej/ueac062

Mountjoy, J. (2022). Community colleges and upward mobility. *American Economic Review*, *112*(8), 2580-2630.

Musselin, C. (2009). *The market for academics*. Routledge.

OEC. (2018). The future of education and skills: education 2030. *OECD Education Working Papers*, 1-23. http://www.oecd.org/education/2030/E2030 Position Paper (05.04.2018).pdf

OEC. (2019). *PISA 2018 Results (Volume I)*. OECD. https://doi.org/10.1787/5f07c754-en

OEC. (2022a). *Pathways to Professions: Understanding Higher Vocational and Professional Tertiary Education Systems*. OECD. https://doi.org/10.1787/a81152f4-en

OEC. (2022b). *Education at a Glance 2022* (Education at a Glance). OECD. https://doi.org/10.1787/3197152b-en

OEC. (2023). *Education at a Glance 2023* (Education at a Glance). OECD. https://doi.org/10.1787/e13bef63-en.

Patnaik, A., Wiswall, M. y Zafar, B. (2021). College Majors 1. *The Routledge Handbook of the Economics of Education*, 415-457.

Pietro, G. D. y Cutillo, A. (2006). University Quality and Labour Market Outcomes in Italy. *LABOUR*, *20*(1), 37-62. https://doi.org/10.1111/j.1467-9914.2006.00333.x

Psacharopoulos, G. y A. Patrinos (2004). "Returns to investment in education: a further update". *Education Economics, 12*, 2

Psacharopoulos, G. y A. Patrinos (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics, 26*, 5, 445-458.

Richardson, J. T. E., Mittelmeier, J. y Rienties, B. (2020). The role of gender, social class and ethnicity in participation and academic attainment in UK higher education: an update. *Oxford Review of Education*, *46*(3), 346-362.

Ruiz-Lázaro, J. (2021). Evaluation of the academic performance of students entering university in Spain. *Edunovatic 2021*, 446.

Sandel, M. J. (2020). *The tyranny of merit: What's become of the common good?* Penguin, UK.

Schanzenbach, D. W., Turner, J. A. y Turner, S. (2023). *Raising state minimum wages, lowering Community College enrollment* (31540; NBER Working Paper Series). http://www.nber.org/papers/w31540

SII (2023). *Indicadores Universitarios de Rendimiento Académico. Curso 2021-2022.*

Solis, Alex (2017). Credit Access and College Enrollment. *Journal of Political Economy, 125*(2): 562-622

Tayebi, A., Gomez, J. y Delgado, C. (2021). Analysis on the Lack of Motivation and Dropout in Engineering Students in Spain. *IEEE Access*, *9*, 66253-66265. https://doi.org/10.1109/ACCESS.2021.3076751

Terviö, M. (2011). Divisions within academia: Evidence from faculty hiring and placement. *Review of Economics and Statistics*, *93*(3), 1053-1062.

Triventi, M. (2013). Stratification in Higher Education and Its Relationship with Social Inequality: A Comparative Study of 11 European Countries. *European Sociological Review,(Vol. 29*, Issue 3). https://www.jstor.org/stable/24479990

Valero, A. y Van Reenen, J. (2019). The economic impact of universities: Evidence from across the globe. *Economics of Education Review*, *68*, 53-67. https://doi.org/10.1016/j.econedurev.2018.09.001

Vergolini, L. y Vlach, E. (2017). Family background and educational path of Italian graduates. *Higher Education*, *73*(2), 245-259. https://doi.org/10.1007/s10734-016-0011-2

Wapman, K. H., Zhang, S., Clauset, A. y Larremore, D. B. (2022). Quantifying hierarchy and dynamics in US faculty hiring and retention. *Nature*, *610*(7930), 120-127.

Williams, W. M. y Ceci, S. J. (2015). National hiring experiments reveal 2: 1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, *112*(17), 5360-5365.

Wiswall, M. y Zafar, B. (2015). *Determinants of College Major Choice: Identification using an Information Experiment* (Vol. 82, Issue 2). https://www.jstor.org/stable/43551547

# Capítulo 5a
# LA ELECCIÓN ESCOLAR

Andreu Arenas
Departament d'Economia
Institut d'Economia de Barcelona
Universitat de Barcelona
ORCID: https://orcid.org/0000-0002-0185-6017

Caterina Calsamiglia
Institute of Political Economy and Governance
ICREA
ORCID: https://orcid.org/0000-0003-2079-7179

**Resumen**

Este capítulo aborda el proceso elección escolar, donde las familias buscan seleccionar una escuela para sus hijos en lugar de depender de la asignación basada en la ubicación residencial. El criterio de asignación puede considerar factores como la ubicación, hermanos en la escuela, poder adquisitivo y necesidades pedagógicas. Existen múltiples formas de tener en cuenta las preferencias de las familias, que pueden influir en la demanda de escuelas, la segregación escolar y en la eficiencia. El proceso suele funcionar en áreas geográficas definidas y se ve afectado por la presencia de escuelas privadas y por la calidad de las escuelas y de la información y las preferencias de las familias. La elección escolar conlleva beneficios y costes, y su implementación varía en todo el mundo, lo que requiere un análisis continuo entre teoría y práctica.

## 1. INTRODUCCIÓN

La elección escolar es un término amplio que describe el proceso mediante el cual las familias tratan de escoger una escuela para sus hijos o hijas, en contraposición a una asignación basada únicamente en el barrio de residencia. El objetivo es asignar alumnos a escuelas teniendo en cuenta sus preferencias, sin considerar su disposición a pagar. Una primera forma de hacerlo es mediante cheques escolares. En lugar de proveer educación pública, el gobierno transfiere los recursos a

las familias, para que los gasten en la escuela de su elección. El proceso de admisión a las escuelas, en tal caso, suele ser descentralizado (por ejemplo, en Chile, o en Estados Unidos). Otra forma de elección escolar, más habitual en Europa, consiste en solicitar a las familias un ranking de preferencias por escuelas y tratar de satisfacerlas de la mejor forma posible mediante un proceso centralizado, a partir de un criterio predeterminado.

Por una parte, este criterio suele establecer qué tipos de familias tienen prioridad para distintos tipos de escuela, por ejemplo, en función del lugar de residencia o de trabajo de los padres, de la presencia de hermanos o hermanas en el centro, del poder adquisitivo, o de las necesidades pedagógicas de los alumnos. Dentro de estas posibilidades, también cabe el prescindir de cualquier consideración que tenga en cuenta las circunstancias de la familia y tratar a todos los potenciales alumnos por igual.

Por otra parte, el criterio de asignación también establece la forma como se tratan de satisfacer las preferencias de las familias. Este es un aspecto importante a tener en cuenta. Cuando las familias manifiestan sus preferencias, deben considerar dos cuestiones clave. La primera, acerca de sus preferencias reales, ¿cuál es la escuela que querrían para sus hijos? La segunda, las preferencias que más les conviene manifestar. Por ejemplo, muy a menudo se pone especial énfasis en satisfacer la primera preferencia de las familias. Imaginemos que nuestra escuela favorita queda lejos de nuestro domicilio, y esto dificulta mucho la admisión, porque se priorizan las familias que viven más cerca. Si manifestamos que esta escuela es nuestra escuela favorita, es muy probable que no consigamos plaza. En tal caso, pasamos a *competir* para ser admitidos en nuestra segunda preferencia. Sin embargo, si se prioriza satisfacer la primera preferencia de las familias, es posible que nos encontremos en desventaja frente a otras familias que sí manifestaron que aquella escuela era su opción favorita. El resultado final podría ser que nos dieran nuestra tercera preferencia. Si por el contrario, decidimos no pedir nuestra escuela favorita, y manifestamos que nuestra segunda preferencia real es nuestra primera opción, tal vez logramos conseguirla. Este ejemplo ilustra como la forma de satisfacer las preferencias de las familias puede repercutir –sobre todo, para aquellas familias más sofisticadas– sobre las escuelas que se demandan. En general, es útil distinguir entre criterios de asignación que incentivan revelar nuestras preferencias de forma estratégica, como en el ejemplo anterior, y criterios que incentivan la revelación de nuestras preferencias reales.

El proceso de elección escolar funciona mediante un mecanismo centralizado dentro de un área geográfica concreta, que puede corresponderse con los

límites de un barrio, de un municipio, de un área metropolitana, o de una región de tamaño parecido. Un aspecto determinante es el porcentaje de escuelas cuyo acceso se determina mediante el mecanismo centralizado de elección escolar. Es posible que existan escuelas privadas que asignen sus plazas en función de la disposición a pagar de las familias, de forma descentralizada. La posibilidad de acceder a estas escuelas condiciona el funcionamiento del mecanismo de elección escolar. Si una familia tiene la capacidad y la disposición a pagar por una escuela privada, solamente se matriculará en la escuela resultante del proceso centralizado de elección escolar si le parece mejor que la alternativa privada.

El proceso de elección escolar –en comparación con una asignación puramente residencial– conlleva potenciales beneficios y costes. La teoría económica y de diseño de mercados nos permite anticipar algunos de los principales efectos de la elección escolar de diferentes tipos. Sin embargo, como veremos más adelante, la importancia y la variabilidad de los detalles en la implementación de la elección escolar en varios lugares del mundo ha supuesto un diálogo constante entre la teoría y el análisis empírico.

## 2. DETERMINANTES DE LA DEMANDA DE ESCUELAS

Uno de los mayores beneficios potenciales de la elección escolar es la posibilidad de conocer mejor las preferencias de las familias. Esto es importante porque permite al gobierno identificar a los centros más demandados, aprender de sus aciertos, darles recursos para aumentar su capacidad. Este proceso puede llevar a un círculo virtuoso que mejore el bienestar social y de las familias.

Estos hipotéticos beneficios se basan en varios supuestos. En primer lugar, las familias deben disponer de información suficiente acerca de los centros educativos para juzgar su calidad. Por otra parte, la calidad de la escuela debe ser un factor central en la demanda de las familias. La verosimilitud de estos supuestos es fundamentalmente una cuestión empírica. Por otra parte, que estos supuestos se cumplan depende en parte del contexto institucional. Por ejemplo, puede haber espacio para políticas públicas que proporcionen información acerca de las escuelas, así como de la importancia de su calidad, para lograr un entorno en que estos supuestos sean más plausibles y la elección escolar funcione mejor.

Entender los determinantes de la demanda de escuelas, por tanto, es una cuestión fundamental que condiciona cualquier análisis e interpretación que hagamos de los resultados del proceso. La evidencia empírica muestra que las

familias tienen una preferencia por escuelas con alumnos de alto rendimiento. Por ejemplo, Hastings y Weinstein (2008) realizan dos experimentos en que se proporcionaba, de forma aleatoria, información directa a familias de bajos ingresos acerca de los resultados académicos de varias escuelas públicas en Carolina del Norte (Estados Unidos). Los resultados muestran que, efectivamente, las familias que reciben esa información terminan escogiendo escuelas con mejores resultados, y que esto acaba resultando en mejores notas para los hijos e hijas de esas familias (lo cual no es sorprendente, tal y como vimos en el capítulo 3c acerca del efecto de los compañeros de clase).

En España, Vega-Bayo y Mariel (2018) estudian la probabilidad de que los padres elijan una escuela en particular en función de sus características (y de las características familiares), con datos de PISA (2012). Los resultados indican que las características con mayor poder explicativo en la elección de escuela son su titularidad (pública o privada), su ubicación y su reputación (medida a partir de la nota media de la escuela). Otras características visibles que podrían ser importantes al elegir una escuela, como el tamaño, la infraestructura, la oferta de extraescolares o la composición socioeconómica tienen un menor poder explicativo, y su importancia depende de la Comunidad Autónoma.

Otra forma de cuantificar hasta qué punto las familias valoran la calidad escolar es analizando los precios de la vivienda en distritos escolares con mejores o peores escuelas. La intuición es sencilla: si vivir en un barrio en concreto proporciona un acceso preferente a escuelas de calidad, y las familias valoran esa calidad, deberían estar dispuestas a pagar un mayor precio por vivir en ese barrio. Si observamos una mayor demanda (y, por tanto, precios más altos) en esos barrios, podemos cuantificar el precio que las familias están dispuestas a pagar por una mayor probabilidad de acceder a una buena escuela, lo cual es una medida de su importancia. El principal reto para realizar este cálculo es que los barrios con mejores o peores escuelas suelen ser diferentes no solo en la dimensión escolar, sino en muchas otras, como por ejemplo su composición demográfica, capital social, o nivel de renta. Una forma creativa de superar este reto consiste en comparar los precios de la vivienda entre dos aceras de la misma calle, que pertenezcan al mismo *barrio* y donde los vecinos experimenten el mismo contexto socioeconómico, pero sin embargo tengan acceso a distritos escolares distintos.

Fack y Grenet (2010) realizan este ejercicio con datos de París, clasificando la calidad de las escuelas a partir de su puntuación promedio en un examen

nacional realizado al final de la escuela Secundaria, la fracción de estudiantes de Secundaria que son admitidos en el plan de estudios general de la escuela preparatoria en lugar de estudios de FP, y la proporción de estudiantes de entornos socioeconómicos privilegiados. Los resultados muestran que los precios de las viviendas que proporcionan un mayor acceso a escuelas de calidad (con un rendimiento de una desviación estándar más) aumentan alrededor de un 2%. Además, este efecto es especialmente pronunciado en aquellas zonas donde existe poca disponibilidad de escuelas privadas. La intuición es la siguiente. En aquellas zonas sin presencia de escuelas privadas, la única forma de acceder a una escuela de calidad es comprando una vivienda en el distrito escolar con mejores escuelas, y por eso aumenta su precio. Sin embargo, en aquellas zonas con presencia de escuelas privadas, existe otra forma de acceder a una buena escuela: pagando directamente la matrícula de una escuela privada de calidad. Por este motivo, el efecto de la calidad de las escuelas es más pequeño en estas zonas.

En resumen, existen estudios con datos de países de alrededor del mundo que confirman que existe una preferencia por escuelas con un mayor rendimiento académico.

### 2.1. Rendimiento absoluto *vs.* valor añadido

Las escuelas con mejores resultados, sin embargo, no son necesariamente las mejores escuelas: tal vez simplemente son escuelas que logran atraer a mejores alumnos, o de entornos más favorecidos económicamente. En otras palabras, que una escuela tenga alumnos con mejores notas no significa que los alumnos tengan mejores notas como consecuencia de estar matriculados en esa escuela. En tal caso, acceder a una escuela con mejores notas medias podría tener un efecto nulo sobre el rendimiento académico individual. Un estudio de Gibbons et al. (2013) con datos de Inglaterra replica el mismo tipo de análisis empírico de Fack y Grenet (2010), pero estudiando la relación entre los precios de la vivienda y el valor añadido de las escuelas, un valor añadido que se mide como el incremento medio en resultados académicos entre los 7 y los 11 años. El valor añadido es una medida importante porque separa la productividad de la escuela de la composición de sus alumnos, que debería verse ya reflejada en los resultados a los 7 años. Los resultados del estudio muestran como los precios de la vivienda responden tanto al valor añadido de la escuela como a las características iniciales de sus estudiantes, y que no es posible encontrar ninguna variable alternativa que explique el efecto del valor añadido. Este resultado es importante porque, de ser extrapo-

lable, significaría que las preferencias de las familias responden a la calidad de una forma que da incentivos a que las escuelas mejoren, y no solamente a que atraigan mejores alumnos (lo cual es un juego de suma cero, basado en intentar mover buenos alumnos de una escuela a otra, en lugar de mejorar a los alumnos).

Ejercicio 1. *a) Considera una familia que duda entre comprar una casa de 300.000 euros u otra de 306.000 euros. La casa de 306.000 euros está en un barrio con una escuela con un valor añadido de una desviación estándar mayor. ¿Cuál es la rentabilidad implícita (es decir, el incremento en el salario esperado por una educación de mejor calidad) que justifica pagar un 2% más en vivienda?; b) Considera que también existe la posibilidad de acceder a una escuela privada, independientemente del barrio de residencia. Asistir a la escuela privada supone un salario esperado igual al de la escuela de la casa de 306.000 euros. La matrícula de la escuela privada (la suma a lo largo de los años de escolarización) es de 10.000 euros. ¿Qué decisión tomará la familia?*

Otro resultado interesante, y con más matices, lo encontramos en la ciudad de Nueva York. Abdulkadiroğlu et al. (2020) estudian la relación entre las preferencias de las familias (a partir de sus solicitudes), la calidad de los compañeros de clase, y medidas de valor añadido de las escuelas. Los resultados muestran, de nuevo, que las familias prefieren escuelas que admiten alumnos de alto rendimiento. A su vez, resulta que estas escuelas son de alto valor añadido: generan mayores mejoras en los resultados académicos a corto y largo plazo de los estudiantes. Sin embargo, los datos muestran que una vez se descuenta la preferencia por la calidad de los alumnos, las preferencias no están relacionadas con la eficacia escolar y la calidad académica de las escuelas. En este caso, esto podría ser compatible con que las familias estén utilizando la calidad de los alumnos como indicador de valor añadido. Pero, en cualquier caso, esto sugiere que los programas de elección de escuela pueden generar incentivos fuertes para una selección más estricta de alumnos, más que para mejorar la calidad académica.

Existen más estudios que muestran que, a menudo, la demanda de las familias responde muy poco a medidas de valor añadido de las escuelas. Mizala y Urquiola (2013) estudian el caso de Chile y su Sistema Nacional de Evaluación de Desempeño (SNED). El SNED es un programa que identifica las escuelas con mayor valor añadido, a partir de una comparación de la evolución de notas de la misma cohorte, comparada con la misma evolución en otras escuelas con características parecidas. De esta forma, el SNED identifica las escuelas con un mayor valor añadido, para cada grupo de comparación, que habitualmente tie-

nen niveles de rendimiento y de nivel socioeconómico distintos. Los resultados del SNED son ampliamente publicitados en medios de comunicación y en las propias escuelas. Los profesores en los centros *ganadores*, además, reciben un bonus salarial. Este es el escenario ideal para entender mejor la pregunta de interés, porque el SNED se acerca muy razonablemente al entorno ideal: se trata de una medida rigurosa de valor añadido, en comparación con escuelas parecidas, e incluye la diseminación de información con el objetivo de mejorar la elección y el esfuerzo de las escuelas.

Para entender la reacción de las familias a la información que una escuela es de alto valor añadido, Mizala y Urquiola (2013) aprovechan que ganarse esa etiqueta por parte del SNED depende de una puntuación (que a su vez es la media de seis sub-puntuaciones de valor añadido). El SNED premia al top 25% de las escuelas dentro de cada grupo, sin revelar la puntuación de cada una de ellas. Como consecuencia, hay un cierto número de escuelas (por ejemplo, entre el top 24% y el top 26%), cuyo valor añadido es muy parecido, pero que reciben un reconocimiento muy distinto, por una cuestión prácticamente de azar (tal y como explicamos en el capítulo 2, se trata de un diseño de regresión discontinua). El estudio se basa en comparar la demanda por estas escuelas al año siguiente, para ver los efectos de proveer información acerca del valor añadido. Los resultados no muestran ninguna reacción; a los padres parece no importarles el reconocimiento del SNED. Estos resultados, en resumen, sugieren que la preferencia por escuelas con alto valor añadido es menos intensa que la preferencia por escuelas con alumnos que suelen tener buenos resultados académicos (aunque estos sean independientes de la escuela).

¿Cuáles son las implicaciones de estos resultados para los beneficios de la elección escolar? MacLeod y Urquiola (2019) plantean un modelo teórico para arrojar luz sobre esta cuestión. El punto de partida es que uno de los mayores motivos para adquirir una buena educación es su posterior rendimiento económico. Una consideración central en el modelo es acerca de la existencia de fricciones, por ejemplo, información imperfecta, en el mercado laboral. La intuición es la siguiente: si las empresas no observan directamente la productividad de los trabajadores, pero sí observan la productividad media de la escuela donde se gradúan, esto puede explicar que sea preferible ir a una escuela de alto rendimiento medio (y, por tanto, con mejores perspectivas laborales), aunque no suponga ningún valor añadido, que no a otra escuela con un menor rendimiento medio (y peores perspectivas laborales), pero un mayor valor académico añadido. En tal caso, la elección escolar no implica una mejora de la calidad vía

competencia entre las escuelas, porque las escuelas compiten por buenos alumnos, no para mejorarlos.

Ejercicio 2. *Explica si el modelo de MacLeod y Urquiola (2019) es más relevante para admisiones en Primaria, Secundaria, o Universitarias, y razona tu respuesta.*

### 2.2. Multidimensionalidad

Existen otros mecanismos que podrían llevarnos a preferir escuelas que no son necesariamente aquellas con mayor valor añadido académico. Basta darse cuenta de que la escuela que atendemos tiene efectos sobre nuestras vidas más allá de nuestros exámenes. Por ejemplo, la escuela determina en buena parte la red social que tendremos a lo largo de nuestras vidas, incluyendo las perspectivas de encontrar una pareja que nos satisfaga, o amigos con valores e intereses parecidos. Los padres a menudo nos socializan y transmiten sus preferencias, motivados por una forma de altruismo paternalista, evaluando nuestras acciones a partir de sus propias preferencias (Bisin y Verdier, 2001), tal y como explicamos en el capítulo 3b acerca de la familia. La multidimensionalidad de la escuela, en definitiva, aparece como una cuestión potencialmente importante, donde el valor añadido es una dimensión más. En tal caso, simplemente hay que ser cautos al interpretar las demandas de las familias por determinadas escuelas como demandas por calidad educativa que debe ser recompensada, replicada, o emerger de forma natural en un entorno competitivo.

Un estudio de Beuermann et al. (2023) con datos de Trinidad y Tobago trata de caracterizar la importancia de la multidimensionalidad en la demanda de escuelas, así como sus efectos, y la relación entre ellos. En primer lugar, examinan el efecto causal de acceder a una escuela en concreto sobre una batería de resultados: resultados académicos de todo tipo, probabilidad de graduarse, de un embarazo adolescente, de ser arrestado, y de participar en el mercado de trabajo. Para cada escuela, obtienen el efecto de matricularse sobre cada uno de los impactos. Muchas políticas educativas, como el *No Child Left Behind* en Estados Unidos o las *League Tables* en Reino Unido tienden a asumir que las escuelas que mejoran los resultados académicos de los alumnos son las mejores escuelas. Un resultado muy interesante es que la correlación entre el valor añadido académico de una escuela y su valor añadido en otras dimensiones es bastante baja. Por ejemplo, la correlación entre el valor añadido académico (en términos de notas en exámenes estandarizados) y el valor añadido en la probabilidad de graduación es solamen-

te de 0,12. Esto significa que la mayoría de las escuelas que mejoran las notas de sus alumnos afectan muy poco su probabilidad de graduarse, y viceversa. La correlación entre el valor añadido académico y el valor añadido en la probabilidad de tener un empleo a posteriori es de 0,12. La correlación con el valor añadido en la probabilidad de arresto es de 0,28, y no hay ninguna correlación para los embarazos adolescentes. Esta cuantificación, aunque se refiera a un caso muy concreto –escuelas públicas en Trinidad y Tobago– ilustra perfectamente la importancia de la multidimensionalidad de las escuelas.

El segundo paso en el estudio de Beuermann et al. (2023) es ver hasta qué punto las familias entienden el valor añadido de cada escuela, y cómo se refleja en sus preferencias. Aunque las familias no explicitan las razones de su elección de centro escolar, sus preferencias por atributos se pueden deducir a partir de la lista de sus escuelas preferidas, que adjuntan con la solicitud de admisión. Los resultados muestran que las familias suelen escoger escuelas con un mayor valor añadido académico, especialmente las familias de alumnos con buenos resultados académicos, y especialmente cuando se trata de alumnas. Por otra parte, las familias también suelen escoger aquellas escuelas que reducen la probabilidad de un embarazo adolescente, sobre todo las familias de alumnos con peores resultados académicos. Igualmente, existe una preferencia por aquellas escuelas que reducen la probabilidad de un arresto y aumentan la participación en el mercado de trabajo.

La interpretación de estos resultados es interesante. Inicialmente, vimos la observación (en estudios anteriores) que las familias reaccionan poco al valor añadido académico como un problema importante para la materialización de los potenciales beneficios de la elección escolar, porque las escuelas no tratarán de mejorar en esa dimensión. Sin embargo, los resultados de Beuermann et al. (2023) matizan un poco esta conclusión, porque si cuando las familias no prestan atención al valor añadido académico es porque están prestando atención al valor añadido en otras dimensiones con valor social, entonces la elección puede llevar a las escuelas a tratar de mejorar en todas esas dimensiones.

También existe evidencia acerca de la multidimensionalidad de la calidad escolar en España. El trabajo de Loviglio (2023) estudia el valor añadido de los institutos públicos de Secundaria en Barcelona, mediante un modelo estructural dinámico de acumulación de habilidades, y distinguiendo entre su valor añadido en rendimiento y en probabilidad de graduación. Esto es particularmente

importante en este contexto, dado que la tasa de abandono escolar en España es del 17%, y un 35% no se matricula en Bachillerato. Los resultados sugieren que las medidas tradicionales de valor añadido deben complementarse con los efectos sobre la probabilidad de graduación, especialmente para los alumnos menos favorecidos, subrayando de nuevo la multidimensionalidad de lo que entendemos como calidad escolar.

## 3. MECANISMOS DE ASIGNACIÓN

Independientemente de los determinantes de las preferencias por las escuelas, la competencia no es el único potencial beneficio de la elección escolar. En un mundo de preferencias heterogéneas, asignar a los alumnos a sus escuelas más preferidas puede suponer importantes mejoras de bienestar, incluso si el comportamiento de las escuelas no cambia. Para ello, es fundamental concretar la forma en que las preferencias se traducen en la asignación de alumnos a escuelas. La elección escolar se materializa a través de mecanismos de asignación, algoritmos que asignan a los alumnos a las escuelas: a) en función de sus preferencias; y b) de los criterios o prioridades que especifiquen las escuelas o el sistema.

Las preferencias de las escuelas o del sistema suelen materializarse en forma de puntos (por ejemplo, por residencia en el barrio) y definen la prioridad que tiene cada alumno a la hora de obtener una plaza en cada una de las escuelas: en caso de exceso de demanda, solo aquellos con más puntos serán admitidos en el centro. En la práctica, existen dos grandes tipos de prioridades, no mutuamente excluyentes. En primer lugar, prioridades que buscan un objetivo social. Por ejemplo, los puntos por tener hermanos en el centro, de proximidad, o por nivel socioeconómico. En segundo lugar, prioridades que buscan un objetivo académico –habitualmente–, tienen en cuenta las notas en etapas anteriores, o necesidades especiales. Este tipo de prioridades son más frecuentes en las últimas etapas del sistema educativo.

Un mecanismo de elección escolar resuelve un problema de elección escolar. Formalmente, un problema de elección escolar consiste en un conjunto finito de estudiantes y un conjunto finito de escuelas con un número finito de plazas disponibles para matrícula. A cada alumno se le debe asignar una plaza en una (y no más de una) escuela. Suponemos que cada alumno tiene preferencias estrictas acerca de todas las escuelas y de la posibilidad de quedarse sin plaza. Asimismo, cada escuela tiene preferencias estrictas (prioridades diferentes) sobre los estudiantes. Un emparejamiento es un resultado de un problema de elección escolar, una asignación de alumnos a plazas de escuelas, de forma que:

- Cada alumno queda emparejado con una escuela o con su opción de reserva.
- El número de alumnos emparejados con una escuela no excede la capacidad de la escuela.
- Un alumno $a$ está emparejado con la escuela $e$ sí y solo sí la escuela $e$ está emparejada con el alumno $a$.

Existen varios mecanismos que asignan a los alumnos con criterios diferentes. Por ejemplo, hay mecanismos que dan más o menos importancia a satisfacer la primera preferencia de las familias. Para compararlos, nos fijaremos en las propiedades de los emparejamientos generados por cada mecanismo. Las principales propiedades deseables de un emparejamiento son la estabilidad, la eficiencia de Pareto, y la no-manipulabilidad.

**Estabilidad.** Un mecanismo es estable si selecciona un emparejamiento estable para cada problema de elección escolar. Un emparejamiento es estable si satisface las condiciones de racionalidad individual, no desperdicio, y ausencia de envidia justificada.

- Racionalidad individual: cada alumno prefiere su emparejamiento a su opción de reserva.
- No desperdicio: si un alumno prefiere otra escuela a su emparejamiento, esa escuela debe estar llena.
- Ausencia de envidia justificada: si un alumno $i$ prefiere la escuela donde un alumno $j$ está emparejado a su escuela de emparejamiento, debe cumplirse que $i$ tiene menor prioridad en la escuela de emparejamiento de $j$ que $j$.

**Eficiencia.** Un mecanismo es eficiente en sentido de Pareto si selecciona un emparejamiento eficiente de Pareto para cada problema de elección escolar. Un emparejamiento $\mu'$ domina en sentido de Pareto a otro emparejamiento $\mu$ si todos los alumnos prefieren $\mu'$ a $\mu$, y al menos uno de ellos lo prefiere de forma estricta, sin indiferencia. Es decir, en el marco de la elección escolar, solo tenemos en cuenta el bienestar del alumno. La noción de eficiencia en sentido de Pareto no se refiere al bienestar total, sino a la imposibilidad de mejorar el bienestar de un alumno sin tener que necesariamente empeorar el de algún otro alumno.

**No manipulabilidad.** Un mecanismo es no-manipulable si ningún estudiante puede en ningún caso beneficiarse de tergiversar sus preferencias. Declarar las

preferencias verdaderas debe ser una estrategia débilmente dominante para cada alumno en el juego de revelación de preferencias inducido por el mecanismo. Una estrategia débilmente dominante es una estrategia para la que la utilidad es mayor o igual que la de otra estrategia del jugador sea cual sea la estrategia de los otros jugadores. Por tanto, es un requisito muy restrictivo, pero a la vez muy interesante porque jugar óptimamente requiere muy poca información sobre lo que hacen los demás.

Los mecanismos más populares de elección escolar son Boston, Gale-Shapley, y Top Trading Cycles.

**Boston.** En el mecanismo de Boston (también conocido como mecanismo de aceptación inmediata), en cada ronda los alumnos solicitan la escuela preferida especificada en el ranking, de la que no hayan sido rechazados en rondas anteriores. Las escuelas aceptan las solicitudes por orden de prioridad. En el momento en que se llenan las plazas disponibles, se rechazan las siguientes solicitudes de la misma ronda, y también de la ronda futura. Por ejemplo, si una escuela recibe más solicitudes que plazas en la primera ronda, se rechazan todas las solicitudes en las rondas sucesivas. El algoritmo finaliza cuando todos los alumnos tienen una plaza asignada en una escuela.

**Gale-Shapley.** En el mecanismo Gale-Shapley (también conocido como mecanismo de aceptación diferida), en cada ronda los alumnos solicitan la escuela preferida especificada en el ranking, de la que no hayan sido rechazados en rondas anteriores. Cada escuela a su vez acepta, provisionalmente (a diferencia de Boston, donde la aceptación es definitiva), las solicitudes por orden de prioridad, hasta llegar a la capacidad de la escuela, rechazando las demás. En cada ronda se consideran las solicitudes pre-aceptadas en la ronda anterior, así como las nuevas solicitudes (que provienen de alumnos que fueron rechazados en la ronda anterior por escuelas más preferidas). Por ejemplo, si una escuela recibe más solicitudes que plazas en la primera ronda, se siguen considerando todas las solicitudes en las rondas sucesivas. El algoritmo finaliza en la ronda en la que ninguna escuela rechaza una solicitud. La asignación de esta ronda es la asignación definitiva.

**Top Trading Cycles.** El mecanismo Top Trading Cycles parte de una asignación inicial donde los alumnos se asignan a la escuela donde tienen más prioridad. A partir de esta asignación, se buscan ciclos de intercambio entre solicitudes que sean mutuamente beneficiosas. La búsqueda de estos ciclos empieza por un

alumno seleccionado aleatoriamente. Supongamos que este alumno se encuentra en la escuela A. Examinamos las escuelas preferidas por todos los alumnos según los ranking entregados. Si existe un ciclo, por ejemplo, si el alumno en la escuela A prefiere la escuela B, y existe un alumno en B que prefiere A, se reasignan y su asignación se considera final. El ciclo no tiene por qué ser bilateral: si un alumno en A prefiere B, un alumno en B prefiere C, y un alumno en C prefiere A, cada uno se mueve a su escuela favorita, y su asignación se considera final. Cuando no existen ciclos de intercambio, la asigación es definitiva.

Ejercicio 3. *A partir de 5 alumnos y 3 escuelas (y sus preferencias y prioridades, ordenades de mayor a menor), calcula la asignación resultante de los 3 mecanismos. La escuela 1 dispone de una plaza, las demás, de dos plazas.*

- $P(i_1) = s_1, s_2, s_3$
- $P(i_2) = s_3, s_1, s_2$
- $P(i_3) = s_1, s_3, s_2$
- $P(i_4) = s_2, s_3, s_1$
- $P(i_5) = s_3, s_2, s_1$
- $P(s_1) = i_5, i_3, i_2, i_4, i_1$
- $P(s_2) = i_2, i_5, i_3, i_4, i_1$
- $P(s_3) = i_3, i_1, i_2, i_4, i_5$

### 3.1. Propiedades

¿Qué diferencias hay entre las propiedades de las asignaciones producidas por los distintos mecanismos? Es fácil darse cuenta de que el mecanismo de Boston es manipulable. Esto se debe a que las asignaciones en primera ronda son definitivas. Por tanto, si la primera preferencia de un alumno en primera ronda es rechazada, puede darse el caso de que su segunda o tercera preferencia, donde hubiera sido aceptado, esté ya llena en la segunda ronda. En tal caso, se quedaría, por ejemplo, con la cuarta preferencia. Sin embargo, si el alumno (o la familia) es capaz de anticipar la baja probabilidad de entrar en su primera preferencia, puede declarar su segunda preferencia como opción más preferida, y en tal caso, lograría matricularse en ella. Una preocupación central a la hora de implementar un mecanismo Boston, en consecuencia, es que todas las familias entiendan por

igual las probabilidades de matricularse con éxito, para competir en igualdad de condiciones.

Por otra parte, en el mecanismo Gale-Shapley, puesto que las asignaciones son provisionales, si un alumno es rechazado por su primera preferencia, compite con los demás por su escuela de segunda preferencia sin ser penalizado, y tiene las mismas probabilidades de ser aceptado que si la hubiera puesto en primera preferencia. Es importante darse cuenta de que esto solamente se cumple si los alumnos pueden declarar tantas preferencias como quieran. Si un alumno o familia solo pudiera listar 3 escuelas (y se quedara sin asignar si no obtuviera plaza en ninguna de las 3), esto introduciría un comportamiento estratégico y Gale-Shapley sería manipulable.

En Top Trading Cycles, la mejor estrategia también es revelar las preferencias verdaderas. Si existe un intercambio posible para la segunda preferencia, este seguirá siendo viable si el alumno declara una escuela más preferida. Si también hay un intercambio posible para la escuela más preferida, este tendrá lugar. En caso contrario, el mismo intercambio que podía existir declarando la segunda como primera sigue existiendo.

También podemos ver fácilmente que la asignación del mecanismo de Boston no es estable. Si la primera preferencia de un alumno en primera ronda es rechazada, puede darse el caso que su segunda preferencia esté llena en la segunda ronda, y ocupada por alumnos con menor prioridad en esa escuela, pero que obtuvieron la plaza al declararla en primera preferencia. Por tanto, esa escuela deja fuera alumnos con mayor prioridad y que la prefieren (en comparación con alumnos asignados).

Por contra, la asignación Gale-Shapley sí es estable. Sin embargo, no es Pareto-eficiente. En el caso de Top Trading Cycles, es al revés. En resumen, no hay ningún mecanismo que sea no manipulable, Pareto-eficiente y estable.

Finalmente, además de las propiedades de la asignación, hay otros detalles de implementación que muy probablemente influyen en la calidad de las asignaciones. Cantillon (2017) propone fijarse, por ejemplo, en el tiempo que pasa entre la solicitud y la decisión o la facilidad de uso del sistema de solicitudes por parte de los padres y las escuelas, entre otras.

Ejercicio 4. *A partir de las asignaciones del ejercicio anterior, explica si son estables, eficientes, y manipulables.*

Ejercicio 5. *a) ¿Cómo interpretas que no haya exceso de demanda en ninguna escuela si la asignación es Gale-Shapley? ¿Y si es Boston? b) ¿Cómo interpretas una asignación donde todo el mundo recibe la primera preferencia si la asignación es Gale-Shapley? ¿Y si es Boston?*

### 3.2. Estudios empíricos

¿Se corresponden las predicciones teóricas con los datos? Una primera aproximación útil a esta pregunta es mediante experimentos. El experimento de Chen y Sönmez (2006) simula un problema de elección de escuela con varios alumnos y escuelas de diferentes características, y compara los resultados cuando de forma aleatoria se presenta a los participantes con los diferentes mecanismos de asignación. Cualitativamente, los resultados experimentales son efectivamente consistentes con las predicciones teóricas. El porcentaje que declara sus preferencias reales bajo Boston es muy reducido, mientras que en Gale-Shapley y TTC es mayoritario. Gale-Shapley es el más estable, seguido de Boston y TTC. TTC es el más eficiente, seguido de Gale-Shapley y Boston. Los costes de eficiencia de Gale-Shapley, sin embargo, son cuantitativamente pequeños. Esto ha hecho que en los últimos años se haya producido una tendencia a sustituir Boston por Gale-Shapley (por ejemplo, en Nueva York, en Boston, o en Inglaterra), tal y como reflejan varios de los casos de estudio que veremos a continuación. Una de las limitaciones de los experimentos es que se basan en grupos pequeños de estudiantes y se simulan preferencias que no tienen por qué ser representativas de las que existen en la realidad. Por ello ha sido muy importante complementar la evidencia experimental con evidencia obtenida de mercados educativos reales, es decir, evidencia a partir de datos administrativos. Por ello hay un elevado número de estudios que usan datos administrativos, aprovechando cambios de sistema para explorar y refinar la conexión entre la teoría y la práctica. Los mecanismos más populares alrededor del mundo son Boston y Gale-Shapley. En el caso de Boston, una cuestión clave es cuantificar la relevancia del comportamiento estratégico y su heterogeneidad entre familias de distintos tipos. Esto es importante porque si algunas familias son menos capaces de ser estratégicas, Boston puede producir una asignación menos favorable para ellas.

Calsamiglia y Güell (2018) estudian esta pregunta con datos administrativos de la ciudad de Barcelona. Desde mediados de la década de los 1990, la asignación escolar en la ciudad funciona con el mecanismo de Boston. Los padres

envían la lista con las preferencias (hasta 10 escuelas). Como explicamos en el apartado anterior, en una primera etapa, se considera la escuela de primera preferencia. Si no hay exceso de demanda, se asigna esa escuela a todos los alumnos. Si hay exceso de demanda, se deciden las plazas en función de los puntos de los alumnos.

La tabla 1 muestra la importancia del barrio y otros criterios para las asignaciones. Un alumno que vive en el barrio de la escuela obtiene 30 puntos. Si no vive en el barrio, pero sus padres o tutores trabajan allí, obtiene 20 puntos. Si no cumple ninguna de las anteriores, pero vive en el distrito (fuera del barrio), 15 puntos. Si simplemente vive en la misma ciudad, 10 puntos. Tener un hermano/a en la escuela preferida suma 40 puntos. Por otra parte, existen puntos extras para familias con características socioeconómicas más desaventajadas, que suman independientemente de la escuela, y que solo reciben un 10% de los solicitantes. En resumen: a igualdad de condiciones, es mucho más probable ser aceptado en la escuela del barrio. Las escuelas fuera del barrio donde un alumno puede ser aceptado serán seguramente escuelas con poca demanda.

**Tabla 1. Puntos de prioridad en caso de sobredemanda para una escuela, Barcelona.**

| Categoría | Puntos |
|---|---|
| **Puntos básicos** | |
| *Hermanos/as:* | |
| El solicitante tiene un hermano/a en la escuela preferida | 40 |
| *Barrio*: | |
| Residencia en el barrio de la escuela preferida | 30 |
| Uno de los padres/tutores trabaja en el barrio de la escuela preferida | 20 |
| Residencia en el distrito administrativo (no barrio) de la escuela preferida | 15 |
| Residencia en la ciudad de la escuela preferida | 10 |
| *Características socioeconómicas:* | |
| Familia de bajos ingresos | 10 |
| Solicitante, padre/tutor o hermano discapacitado | 10 |
| **Puntos extra** | |
| *Características socioeconómicas:* | |
| Familia numerosa (tres o más hermanos) o padres solteros | 15 |
| Solicitante con enfermedad digestiva crónica | 10 |

Nota: Los criterios de barrio son mutuamente excluyentes.
Fuente: Calsamiglia y Güell (2018).

¿Hasta qué punto esto condiciona las demandas de las familias? Es posible que las familias quieran ir a la escuela del barrio independientemente de los puntos. ¿Cuántas familias están renunciando a su opción realmente preferida por miedo a perder su mejor opción en su barrio? Para responder a esta pregunta, aprovechan un cambio en la definición de los barrios. Antes de 2007, la ciudad estaba dividida en barrios correspondientes a divisiones administrativas. Un problema de esta división es que no capturaba bien la distancia a las escuelas. Por ejemplo, las familias viviendo cerca de la frontera de un distrito no tenían puntos de prioridad para aquellas escuelas, muy cercanas, al otro lado de la frontera. Además, en función de la densidad del distrito, el número de escuelas donde cada familia tenía puntos de prioridad era distinto. La reforma de 2007 cambió esta definición. El barrio pasó a definirse en función de la distancia del bloque de casas de la familia, de forma que incluyera las 6 escuelas más cercanas: 3 públicas y 3 concertadas (de gestión privada, pero matrícula subvencionada). Como resultado de la reforma, para cada familia, podemos distinguir 4 tipos de escuela: 1) escuelas para las que tenía puntos de prioridad y los mantiene; 2) escuelas para las que tenía puntos de prioridad y los pierde; 3) escuelas para las que no tenía puntos de prioridad y pasa a tenerlos; y 4) escuelas para las que no tiene puntos de prioridad en ningún caso.

Los datos muestran como la probabilidad de pedir una escuela para la cual ya no se obtienen puntos de proximidad cae un 50%. Por otra parte, la probabilidad de pedir una escuela para la cual ahora sí se obtienen puntos de proximidad se duplica. Esto sugiere que el comportamiento estratégico juega un papel muy importante a la hora de pedir escuelas en el mecanismo de Boston, en línea con la teoría. Las familias suelen pedir escuelas para las cuales la admisión está prácticamente garantizada en la mayoría de los casos (58%), incurren en riesgos en un 40% de los casos (y su matrícula dependerá de una lotería), y solo en un 2% de los casos piden escuelas donde es prácticamente imposible que sean admitidas. Además, las autoras identifican un 20% de familias con un patrón de demanda naif: asumen riesgo en la primera ronda, y su segunda preferencia siempre se llena en la primera ronda (y por tanto, no será posible acceder a ella).

No todas las familias que solicitan escuela terminan matriculándose. Los datos de Calsamiglia y Güell (2018) permiten identificar las familias con una opción de reserva fuera del sistema (es decir, una escuela privada). Dado que las familias que incurren riesgos obtienen su plaza en función de una lotería, las que la consiguen y las que no son prácticamente idénticas. Los resultados muestran

que alrededor de un 13% de los que no la consiguen terminan matriculándose en una escuela privada, en comparación con las que sí la consiguen gracias a la lotería. Este resultado es importante porque muestra cómo la capacidad de tener una opción de reserva permite incurrir mayores riesgos, y también explica una parte del comportamiento que a priori parece naif o exageradamente arriesgado.

Los datos también muestran como aquellos con una opción de reserva fuera del sistema (y que incurren en riesgos) tienen un nivel socioeconómico más alto que aquellos que escogen la opción segura. Por otra parte, aquellos que incurren en riesgos sin tener una opción de reserva tienen un nivel socioeconómico parecido al de aquellos que escogen la opción segura. Estos resultados son especialmente interesantes, porque muestran que, a igualdad de opción de reserva, el comportamiento estratégico no está relacionado con el estatus socioeconómico, lo cual es una preocupación central en mecanismos manipulables como Boston. En otras palabras, el diferencial de comportamiento estratégico por nivel socioeconómico en este caso se explica por la capacidad de evitar una mala asignación, no por diferencias en sofisticación.

¿Habrían sido los resultados distintos con Gale-Shapley o Top Trading Cycles? Usan los datos de Barcelona para estimar la distribución conjunta de preferencias y sofisticación estratégica y hacer un análisis contrafáctico de los resultados y del bienestar que tendría lugar en cada caso. En general, encuentran que cambiar a Gale-Shapley daría peores resultados, pero Top Trading Cycles mejoraría la situación. Estas estimaciones se refieren a la media: siempre habría ganadores y perdedores. En el caso de Gale-Shapley, el bienestar decrece sobre todo para las familias no estratégicas, las que ponían sus preferencias verdaderas porque o bien no tenían nada que perder por tener una escuela del barrio muy mala o porque tenían la opción de ir a una escuela privada. Aquellos que ganan suelen estar en mejores barrios. Bajo TTC, las familias menos estratégicas son las que ganan más, y están más repartidas.

En este caso, Gale-Shapley habría asignado las familias a escuelas de mayor calidad, menos distancia, y más caras. Sin embargo, para las familias menos educadas la calidad habría aumentado menos, la distancia también se reduciría también menos, pero el precio habría aumentado más. TTC habría asignado las familias a escuelas de mayor calidad, más distancia y más caras. En conjunto, los resultados muestran como Boston crea ineficiencias al dar incentivos a las familias a pedir escuelas cercanas para las cuales tienen mayor prioridad, en detrimento de escuelas de mayor calidad fuera del barrio.

Finalmente, cabe destacar que el desarrollo de las nuevas tecnologías plantea oportunidades para proveer información durante el mismo proceso de solicitud, para mejorar el resultado de los mecanismos de elección. Por ejemplo, un experimento reciente de Arteaga et al. (2022) proporciona información en directo (durante el proceso de solicitud) en los sistemas de elección de Chile y New Haven. El 22% de los solicitantes que presentan solicitudes con un alto riesgo de no ser asignados a una escuela responden a las advertencias añadiendo escuelas a sus listas, lo que reduce el riesgo de no ser asignados en un 58% y aumenta el valor agregado de las calificaciones en las escuelas donde se inscriben en 0,10 desviaciones estándar.

### 3.3. Elección escolar y segregación

Los sistemas de elección escolar suelen dar prioridad a escuelas cercanas al lugar de residencia o de trabajo de las familias. Esto sigue una lógica evidente, consistente en minimizar la distancia de desplazamiento y maximizar la involucración de los padres en la escuela. Sin embargo, la otra cara de la moneda es que la segregación residencial –es decir, las diferencias socioeconómicas entre barrios– se convierte también en segregación escolar.

La segregación residencial y escolar suponen un límite importante a la igualdad de oportunidades. Por ejemplo, Chetty et al. (2016) estudian los resultados del experimento "Moving to Opportunity" (MTO) en EE. UU., que ofreció cheques de vivienda a familias seleccionadas al azar para que se mudaran de proyectos de viviendas de alta pobreza a vecindarios de menor pobreza. Mudarse a un vecindario de menor pobreza cuando son los niños jóvenes (antes de los 13 años) aumentó la asistencia a la Universidad y los ingresos, y redujo las tasas de maternidad o paternidad monoparentales. Por otra parte, Goux y Maurin (2007) encuentran, con datos de Francia, que el rendimiento de un adolescente al final de la Secundaria está fuertemente influenciado por el rendimiento de otros adolescentes en el vecindario, en línea con lo que explicamos en el capítulo 3c acerca del efecto de los compañeros de clase. Arenas y Hindriks (2021) estiman que las políticas de de-segregación tienen impactos significativos sobre la movilidad intergeneracional a un coste de eficiencia muy reducido. Una pregunta natural, por tanto, es si es posible diseñar los mecanismos de elección escolar de forma que minimicen la segregación. Como veremos, la existencia de puntos por barrio y de opciones de reserva en forma de escuelas privadas juegan un papel central en la respuesta a esta pregunta.

Calsamiglia et al. (2021) estudian la posibilidad de segregación cardinal. Esta surge cuando hogares con preferencias ordinales idénticas reaccionan al mecanismo de elección con estrategias diversas y terminan con asignaciones escolares diferentes. Resulta que, bajo Boston, la elección de escuela puede caracterizarse como una elección entre una lotería de escuelas y una escuela segura. La variación en el grado de aversión al riesgo hacia las escuelas puede, por lo tanto, resultar en diferencias en las solicitudes óptimas a enviar. Intuitivamente, esta condición requiere que las familias privilegiadas estén dispuestas a correr mayores riesgos. La existencia de escuelas privadas exagera estos efectos y dificulta aún más la igualdad de oportunidades educativas al reducir la calidad de los compañeros de las peores escuelas públicas. Los que pueden pagar o valoran las escuelas privadas acaban con mayores posibilidades de llevar a su hijo a las mejores escuelas públicas. La implicación notable y preocupante es el hecho de que la asignación a escuelas públicas gratuitas se ve afectada por las diferencias en la disposición de los hogares para pagar por una escuela privada.

Por otra parte, Calsamiglia y Miralles (2023) estudian un caso caracterizado por la existencia de una escuela que todo el mundo quiere evitar, y una asignación Gale-Shapley. Sus resultados muestran que, si se aplica prioridad para las escuelas del vecindario, y el mercado es lo suficientemente grande, todas las familias acaban siendo asignadas a su escuela local del vecindario, independientemente de sus preferencias. La intuición es la siguiente: todas las familias fuera del área de la escuela *indeseable* piden antes la del barrio, y la consiguen. En cambio, todas las familias en el área de la escuela *indeseable*, aunque pidan la última, terminan siendo rechazadas por las demás, ya llenas por las familias del barrio, y acaban en la escuela *indeseable*. Este mecanismo intuitivo ilustra la importancia de una política pública que trate de evitar a toda costa la existencia de una escuela de este tipo.

Ejercicio 6. *Explica, a partir de un ejemplo sencillo, cómo la existencia de una escuela indeseable para todos combinada con puntos por residencia y una asignación Gale-Shapley, hace que todas las familias acaben siendo asignadas a su escuela del vecindario. Compáralo con un caso donde las preferencias son heterogéneas y no existe una escuela indeseable.*

En España, Gortázar et al. (2023) analizan el caso de Madrid, donde el mecanismo de asignación es Boston, con un criterio de puntos con una estructura similar a la de Barcelona. Los autores aprovechan que, en 2013, se eliminaron los

puntos de prioridad por proximidad dentro de la ciudad para analizar la cuestión empíricamente. Los resultados muestran que eliminar las prioridades basadas en la residencia aumenta la asignación a escuelas fuera del distrito y a escuelas ubicadas más lejos del domicilio familiar. Inicialmente, la respuesta se da por parte de familias con hijos o hijas nacidos en España, aunque poco a poco las demás familias convergen. La mayoría de las familias acceden a escuelas de algo más de calidad, especialmente los hijos de padres con menor nivel educativo. Por tanto, la eliminación de las prioridades basadas en la residencia parece ayudar a las familias de entornos con menor nivel educativo a acceder a escuelas de mejor rendimiento, lo que potencialmente lleva a mejores resultados a largo plazo. Aunque una evaluación de bienestar completa es difícil de hacer cuando no podemos interpretar las preferencias como verdaderas, al tratarse del mecanismo de Boston, los autores también muestran que las familias ejercen más elección y participan más en el proceso de elección de una escuela.

Kutscher et al. (2023) estudian la implementación de un mecanismo Gale-Shapley en 2016 por parte del gobierno chileno, como componente central de una importante reforma educativa, que tenía como objetivo promover la inclusión social y reducir la segregación escolar. Los autores aprovechan que la reforma comenzó a aplicarse en distintas zonas del país en distintos cursos, para comparar los cambios en la segregación que tuvieron lugar antes y después de la reforma.

Los resultados muestran que la reforma apenas cambió la segregación escolar media. Sin embargo, este resultado oculta una heterogeneidad considerable entre distritos escolares, en función de su segregación residencial y oferta escolar previa. Por un lado, el sistema centralizado amplía la elección de escuelas, y es probable que algunas familias de bajos recursos socioeconómicos utilicen el nuevo sistema para enviar a sus hijos a escuelas públicas o con cheques de mejor calidad que se encuentran más lejos de sus domicilios. Sin embargo, esto no necesariamente reducirá la segregación escolar si los estudiantes de alto nivel socioeconómico lo anticipan, y responden cambiándose a escuelas privadas. Los resultados de Kutscher et al. (2023) muestran que los municipios con una mayor representación de escuelas públicas o con cheques experimentaron una disminución en la segregación escolar, mientras que los municipios con más escuelas privadas experimentaron un aumento en la segregación escolar. Este patrón de cambio de escuela es especialmente intenso para las familias de altos ingresos. Estos resultados son muy consistentes con el mecanismo teórico propuesto, y de

nuevo, el papel de las opciones de reserva aparece como fundamental a la hora de entender los resultados de los mecanismos de elección escolar.

Finalmente, es importante destacar que, aunque los sistemas de elección escolar centralizada transfieren formalmente el poder de elección a las familias (de acuerdo a las prioridades establecidas), las escuelas siguen teniendo cierta influencia, de forma sutil. Una investigación de Lafuente (2021) ilustra muy bien este canal. En un experimento aleatorizado, se enviaron correos electrónicos a más de 2.500 escuelas de la Comunidad de Madrid, en el periodo previo al proceso de inscripción de estudiantes de 2018. En estos mensajes, tres tipos de parejas ficticias, una española y dos rumanas, solicitaron visitar la escuela. Los respectivos matrimonios rumanos señalaban diferentes niveles de asimilación cultural con el origen del nombre de su hijo. Para mejorar la comparabilidad entre los perfiles familiares, se controlaron una serie de características socioeconómicas que influyen en la selección hacia la asimilación y que afectan la discriminación.

Los resultados muestran que los inmigrantes que eligen un nombre extranjero (rumano) para sus hijos tienen una tasa de respuesta 9 puntos porcentuales más baja, en comparación con una tasa de respuesta base del 77% para los nativos. Por otra parte, la decisión de los padres inmigrantes de dar a su hijo un nombre nativo reduce significativamente la discriminación en un 50%. Cualitativamente, estos resultados se mantienen con independencia de las características de las escuelas, aunque el poder estadístico de la muestra para detectar tales diferencias sea menor. Vistos estos resultados, es fácil imaginar que estas diferencias en tasas de respuesta expliquen diferencias en información, y potencialmente, en la demanda por esas escuelas. De esta forma, las escuelas pueden incentivar la demanda por parte de grupos demográficos que les interesen más o menos, poniendo más o menos barreras a obtener información acerca del centro, y generando segregación.

## 4. LA ELECCIÓN UNIVERSITARIA

El problema de elección universitaria guarda muchas similitudes con el de elección escolar: se trata de asignar un número finito de alumnos a un número finito de plazas universitarias. En multitud de países alrededor del mundo, la asignación a la Universidad se produce de forma centralizada, utilizando los mismos mecanismos –sobretodo, Gale-Shapley–, con algunas modificaciones, y con preocupaciones distintas.

En la elección universitaria, la prioridad de admisión se centra en los méritos académicos. La proximidad suele tener un papel formalmente inexistente (aunque evidentemente el coste de desplazamiento entre como un factor de importancia en la demanda). En algunos países se establecen cuotas para grupos demográficos o socioeconómicos cuya participación en el sistema ha sido tradicionalmente menor.

La principal decisión a la hora de diseñar el sistema de elección universitaria es la definición de las prioridades. En prácticamente todo el mundo, la prioridad viene definida por el rendimiento en pruebas estandarizadas de acceso, a menudo en conjunción con el rendimiento académico en Bachillerato. En el caso de España, la nota media de Bachillerato supone 6 puntos, y la evaluación del Bachillerato para el acceso a la Universidad (EBAU) hasta 8 puntos más, llegando a un máximo de 14 puntos. Anteriormente, la EBAU había tenido menos peso relativo (hasta 2010, la nota media de Bachillerato suponía 6 puntos y la EBAU 4). El uso de pruebas estandarizadas para determinar la prioridad de matrícula tiene ventajas e inconvenientes. La principal ventaja, obvia, es que se miden las habilidades con una métrica común (dentro de cada Comunidad Autónoma, con cierta autonomía en el currículum de Bachillerato). La principal limitación es que se miden las habilidades en un día y en un formato de examen concreto. Por ejemplo, el rendimiento en ese tipo de pruebas puede ser bastante desigual entre grupos demográficos, sin que necesariamente esas diferencias reflejen un potencial de rendimiento universitario diferente. Arenas y Calsamiglia (2022) muestran, con datos de Cataluña, como la reforma de 2010 afectó negativamente las notas de admisión de las mujeres (cuyo rendimiento es muy superior al de los hombres en Bachillerato, y parecido en la EBAU), sin que esa diferencia esté relacionada con diferencias en el rendimiento esperado en la Universidad (en el capítulo 5c veremos muchas más diferencias de género en educación).

Otra decisión importante acerca de las prioridades concierne al peso relativo que se le da a las distintas asignaturas. Por ejemplo, en caso de empate, se puede priorizar a los alumnos que hayan cursado asignaturas más relevantes para el Grado solicitado. En la práctica, en España esto se materializa a través de los pesos que cada Grado otorga a la calificación de las asignaturas de modalidad.

La prioridad de los alumnos suele resumirse en una nota de admisión. Esto tiene la ventaja de facilitar la cuantificación de notas de corte en cada Grado

universitario (es decir, la nota de admisión del último alumno que logró obtener plaza), y por tanto, la evaluación de las probabilidades de acceso del alumno en cada Grado. Las notas de corte suelen estar fuertemente correlacionadas en el tiempo, pero incluso en caso de que no lo estén, la información es pública y es fácil de tener en cuenta.

En España, la admisión a la Universidad está gestionada por las Comunidades Autónomas. El proceso de asignación es el mismo en todas partes, pero los solicitantes deben rellenar una solicitud por cada Comunidad Autónoma donde consideren matricularse. El número de preferencias que los alumnos pueden declarar suele ser, desafortunadamente, limitado, lo que puede incentivar a cierto comportamiento estratégico y a ineficiencias. Sin embargo, la información pública acerca de las notas de corte muy probablemente minimiza la probabilidad de errores de estrategia o de diferencias de sofisticación en la práctica.

## 5. CONCLUSIONES

En este capítulo se ha examinado el impacto que la elección de escuela puede tener en la eficacia y equidad del sistema educativo. Se ha destacado que permitir a las familias elegir la escuela puede ser una estrategia interesante para motivar la mejora de las instituciones educativas. Sin embargo, para que esto funcione es crucial que la demanda esté guiada por las preferencias de las familias en relación a los aspectos que la escuela puede influir, lo que a su vez las motivaría a seguir perfeccionando. No obstante, hemos constatado que, por un lado, las familias no siempre cuentan con la información necesaria sobre las diversas dimensiones que determinan la idoneidad de la escuela para su hijo o hija y que, aun así, su capacidad de elegir colegio se ve muy afectada por el sistema de prioridades impuesto. Por otro lado, las escuelas no siempre disponen de la capacidad para adaptarse a las necesidades de la demanda, ya que el nivel de autonomía del centro suele ser limitado. Asimismo, aspectos como la composición de los compañeros en la escuela, que dependen de las familias que el centro atrae, están fuertemente influenciados por los sistemas de priorización de la demanda. Por último, también hemos observado que la elección de escuela tiene externalidades que pueden afectar la equidad del sistema, influyendo en los niveles de segregación observados. Además, hemos visto cómo la existencia de escuelas que no son deseadas por la mayoría de las familias puede desactivar el potencial transformador de la elección de escuela, al mismo tiempo que perpetúa la presencia de escuelas con dificultades.

## BIBLIOGRAFÍA


Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J. y Walters, C. R. (2020). Do parents value school effectiveness? *American Economic Review, 110*(5),1502-1539.

Arenas, A. y Calsamiglia, C. (2022). Gender differences in high-stakes performance and college admission policies. Technical report, Institute of Labor Economics (IZA).

Arenas, A. y Hindriks, J. (2021). Intergenerational mobility and unequal school opportunity. *The Economic Journal, 131*(635),1027-1050.

Arteaga, F., Kapor, A. J., Neilson, C. A. y Zimmerman, S. D. (2022). Smart matching platforms and heterogeneous beliefs in centralized school choice. *The Quarterly Journal of Economics, 137*(3),1791-1848.

Beuermann, D. W., Jackson, C. K., Navarro-Sola, L. y Pardo, F. (2023). What is a good school, and can parents tell? evidence on the multidimensionality of school output. *The Review of Economic Studies, 90*(1), 65-101.

Bisin, A. y Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory, 97*(2), 298-319.

Calsamiglia, C., Fu, C. y Güell, M. (2020). Structural estimation of a model of school choices: The boston mechanism versus its alternatives. *Journal of Political Economy, 128*(2),642-680.

Calsamiglia, C. y Güell, M. (2018). Priorities in school choice: The case of the Boston mechanism in Barcelona. *Journal of Public Economics, 163*, 20-36.

Calsamiglia, C., Martínez-Mora, F. y Miralles, A. (2021). School choice design, risk aversion and cardinal segregation. *The Economic Journal, 131*(635), 1081-1104.

Calsamiglia, C. y Miralles, A. (2023). Catchment areas, stratification and access to better schools. *International Economic Review*.

Cantillon, E. (2017). Broadening the market design approach to school choice. *Oxford Review of Economic Policy, 33*(4), 613-634.

Chen, Y. y Sönmez, T. (2006). School choice: an experimental study. *Journal of Economic Theory, 127*(1), 202-231.

Chetty, R., Hendren, N. y Katz, L. F. (2016). The effects of exposure to better neighbor hoods on children: New evidence from the moving to opportunity experiment. *American Economic Review, 106*(4), 855-902.

Fack, G. y Grenet, J. (2010). When do better schools raise housing prices? evidence from paris public and private schools. *Journal of Public Economics, 94*(1-2), 59-77.

Gibbons, S., Machin, S. y Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics, 75*, 15-28.

Gortázar, L., Mayor, D. y Montalbán, J. (2023). Residence-based priorities and school choice. *Economics of Education Review, 95*, 102384.

Goux, D. y Maurin, E. (2007). Close Neighbours Matter: Neighbourhood Effects on Early Performance at School. *The Economic Journal, 117*(523), 1193-1215.

Hastings, J. S. y Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics, 123*(4), 1373-1414.

Kutscher, M., Nath, S. y Urzúa, S. (2023). Centralized admission systems and school segregation: Evidence from a national reform. *Journal of Public Economics, 221*, 104863.

Lafuente, D. M., de (2021). Cultural assimilation and ethnic discrimination: an audit study with schools. *Labour Economics, 72*, 102058.

Loviglio, A. (2023). School quality beyond test scores: the role of schools in shaping educational outcomes. Technical report, Dipartimento Scienze Economiche, Universita'di Bologna.

MacLeod, W. B. y Urquiola, M. (2019). Is education consumption or investment? implications for school competition. *Annual Review of Economics, 11*, 563-589.

Mizala, A. y Urquiola, M. (2013). School markets: The impact of information approximating schools' effectiveness. *Journal of Development Economics, 103*, 313-335.

Vega-Bayo, A. y Mariel, P. (2018). School choice across different regions of Spain. *Hacienda Pública Española*, (227), 11-36.

# Capítulo 5b
# DESIGUALDAD EDUCATIVA

Marisa Hidalgo-Hidalgo
Departamento de Economía, Métodos Cuantitativos e Historia Económica
Universidad Pablo de Olavide
ORCID: https://orcid.org/0000-0001-9813-7766

Iñigo Iturbe-Ormaetxe
Departamento de Fundamentos del Análisis Económico
Universidad de Alicante
ORCID: https://orcid.org/0000-0002-2385-2471

**Resumen**

En los últimos años, los niveles educativos han aumentado globalmente, pero persisten desigualdades significativas. Aunque España ha mejorado, sigue teniendo una proporción considerable de jóvenes que no completan la Educación Secundaria Superior. A pesar de la moderada desigualdad de ingresos y riqueza en comparación con otros países de la UE, en España la tasa de abandono depende mucho de las características familiares. Este capítulo examina las disparidades educativas en España, analizando tasas de abandono escolar, educación universitaria y puntuaciones en pruebas internacionales. A pesar de la estabilidad en la desigualdad de ingresos y riqueza, las desigualdades educativas persisten, lo que plantea interrogantes sobre las medidas de equidad y las posibles soluciones.

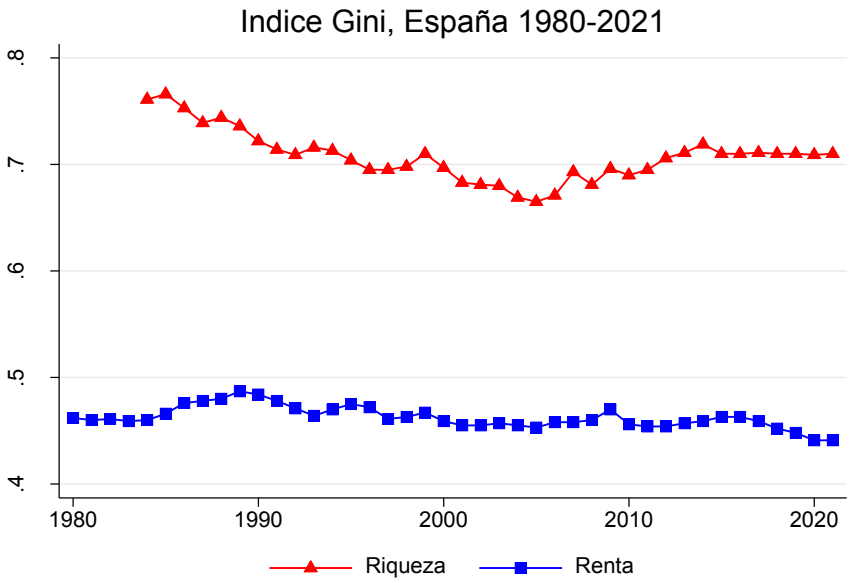
## 1. INTRODUCCIÓN

Los niveles educativos de la población han aumentado en los últimos años en todos los países. En los países de la OCDE, la proporción media de personas entre 25 y 34 años con estudios superiores (Universitarios y de Formación Profesional Superior) ha aumentado de un 27% en el año 2000 a un 48% en el año 2021. En el caso de España el aumento ha sido del 34% al 49% (OECD, 2022). En ese mismo grupo de edad la proporción que no

completó sus estudios secundarios superiores ha bajado del 19% en 2011 al 14% en 2021. En España la bajada ha sido del 35% al 28%. A pesar de la evidente mejoría, España se sitúa entre los países de la OCDE con una mayor proporción de jóvenes que no completan los estudios secundarios superiores (Bachillerato o FP Media), solo superada por Turquía (36%), México (44%) y Costa Rica (45%) (OECD, 2022) (vease el capítulo 1 de indicadores). En este capítulo vamos a ver que el abandono escolar está estrechamente relacionado con las características socioeconómicas del individuo. En particular, los individuos con padres de bajo nivel educativo tienen mayores probabilidades de abandonar los estudios.

La desigualdad educativa es un tema muy relevante desde el punto de vista de las políticas públicas entre otros motivos por las consecuencias futuras de la educación, tanto en resultados laborales (mayor probabilidad de empleo, mayores salarios) como también en la salud o la participación social. También hay un amplio consenso en qué parte de esta asociación refleja una relación causal (ver Lochner, 2011). Dados estos efectos positivos de la educación, una sociedad en la que algunos grupos quedan rezagados de forma sistemática, condena a estos grupos a una posición de desventaja en el futuro.

**Gráfico1. Desigualdad en renta y riqueza en España.**




Fuente: World Inequality Database.

La equidad en la educación es, de hecho, un tema central en la mayoría de países. Los Objetivos de Desarrollo Sostenible (ODS) adoptados por las Naciones Unidas en 2015 con el fin último de "poner fin a la pobreza, proteger el planeta y garantizar que para el 2030 todas las personas disfruten de paz y prosperidad" abogan por "garantizar una educación inclusiva y equitativa de calidad y promover oportunidades de aprendizaje permanente para todos".

Por supuesto, la desigualdad educativa puede ser un reflejo de la desigualdad de renta o de riqueza en un país. ¿Cuál es la situación de España en cuanto a desigualdad de renta y de riqueza? La base de datos más completa sobre desigualdad es la World Inequality Database coordinada por Facundo Alvaredo, Lucas Chancel, Thomas Piketty, Emmanuel Saez y Gabriel Zucman (ver wid.world). En el gráfico siguiente representamos el índice de Gini de la distribución de la renta (antes de impuestos) y de la riqueza en España en los últimos cuarenta años.

El índice de Gini es una medida clásica de desigualdad que toma valores entre 0 y 1, donde 0 significa igualdad máxima (todos los individuos tienen la misma renta) y 1 significa desigualdad máxima (un único individuo posee toda la renta). Por tanto, valores mayores del índice indican una mayor desigualdad. Según la World Inequality Database, entre los países de la UE el índice de Gini para la renta antes de impuestos en el año 2021 oscila entre 0,38 (Eslovaquia) y 0,62 (Lituania). La cifra más reciente para España es 0,44 en el año 2021, lo que significa que la desigualdad de renta es relativamente moderada. Además, en el gráfico de arriba vemos que el índice de Gini no ha experimentado grandes fluctuaciones en los últimos 40 años. En cuanto a la riqueza, esta tiende a estar más concentrada que la renta. En el caso de España el último valor del índice en 2021 es de 0,71, por debajo de la media de la UE (0,76). En el gráfico vemos que el índice de Gini de la riqueza fue reduciéndose hasta el 2005, pero que en los últimos años ha vuelto a aumentar, lo que indica una mayor concentración de la riqueza.

En conclusión, comparativamente con otros países de la UE, España no parece un país con desigualdades de renta y riqueza severas. Además, tampoco parece que en los últimos cuarenta años haya habido grandes fluctuaciones en la desigualdad de renta o riqueza, al menos si consideramos el índice de Gini. Según esto, en principio no deberíamos esperar encontrar grandes desigualdades con respecto a los niveles educativos. Sin embargo, como veremos más adelante, esto no es lo que encontramos en los datos, lo que puede resultar paradójico. Esto es

lo que trataremos de ilustrar a lo largo del capítulo, pero para ello tenemos que empezar por discutir cómo medimos la educación y qué medida de desigualdad usamos. Esto es lo que hacemos en la sección 2. En la sección 3 presentamos evidencia sobre desigualdad educativa en España respecto a tasas de abandono escolar, población universitaria y puntuaciones en pruebas internacionales. En la sección 4 hablamos sobre las posibles consecuencias de la desigualdad educativa actual. Para finalizar, en la sección 5 discutimos sobre las causas de la desigualdad educativa así como posibles soluciones.

Ejercicio 1. *Además del índice de Gini que hemos usado para construir el gráfico 1, hay muchas otras medidas de desigualdad en la literatura. Por ejemplo, el porcentaje de renta (o riqueza) que poseen el 10% de los individuos más ricos, o el porcentaje de renta (o riqueza) que poseen el 1% de los individuos más ricos. En la base de datos de wid.world tienes datos de estas dos medidas para España, los presentan como el "top" 10% y como el "top" 1%, respectivamente. Descarga los datos de España correspondientes al periodo 1980-2021 y construye un gráfico similar al gráfico 1. ¿Cómo ha sido la evolución de estos indicadores a lo largo del periodo? ¿Es esta evolución similar a la que vemos usando el índice de Gini?*

## 2. MEDIDAS DE DESIGUALDAD EDUCATIVA

El primer problema a la hora de medir la desigualdad educativa es decidir cómo medir el nivel educativo de un individuo. Algunos autores usan el número de años de educación de la persona. En este caso, una medida de desigualdad podría ser la desviación estándar de los años de estudio de la población adulta. Al tratarse de una medida del grado de dispersión de los datos respecto a la media, una mayor desviación estándar supondrá una mayor desigualdad (ver Ram, 1984, 1990, entre los primeros trabajos en usarla). Otra opción es usar el índice de Gini educativo, como hicieron Checchi (2001) o Castellò y Doménech (2002) a partir de los datos sobre niveles educativos de Barro y Lee (2001). Además del índice de Gini, es posible calcular un conjunto más amplio de índices de dispersión, entre los que se incluyen el índice de Theil (ver Meschi y Cervini, 2014 para un análisis de desigualdad educativa usando estos dos índices). Sin embargo, el uso de los años de educación es problemático por varias razones. En primer lugar, es una medida que puede sobreestimar la educación individual (este sería el caso, por ejemplo, entre aquellos que han repetido curso). Para evitar esta primera crítica se suelen usar medidas del logro educativo del individuo,

esto es, medidas del nivel educativo máximo alcanzado. Dos medidas muy usadas son el haber completado o no estudios de Secundaria Superior (abandono escolar) y el haber completado estudios universitarios o de FP Superior. En segundo lugar, una crítica habitual al uso de número de años de estudio es que dificulta la comparación entre países, en tanto que un año adicional de estudio puede ser muy diferente en un país desarrollado o en otro en desarrollo o incluso dentro de un mismo país a lo largo del tiempo. La creciente disponibilidad de datos comparables sobre rendimiento de los estudiantes en pruebas internacionales ha facilitado el uso de otras medidas, en general confirmando que cantidad y calidad de la educación no siempre van de la mano (Hanushek y Woessmann, 2010).

Nosotros vamos a usar diferentes medidas de la educación de los individuos y usaremos como medida de desigualdad la brecha respecto a las distintas medidas educativas entre individuos de grupos diferentes. En concreto, vamos a considerar dos formas de desigualdad educativa:

a) Desigualdades entre individuos procedentes de distintos grupos sociales. Analizaremos la brecha en diferentes variables educativas: puntuaciones (o notas) y logros académicos (niveles educativos alcanzados). Consideraremos así mismo distintos tipos de grupos sociales en función de, por ejemplo, la educación familiar o el nivel socioeconómico.

b) Desigualdades en resultados a lo largo de la vida, por ejemplo, estatus laboral y de pobreza, entre individuos con distinto nivel educativo. Aquí estudiaremos la brecha en distintos resultados laborales entre estos grupos.

Nos centraremos en el caso de España, poniéndola en perspectiva con los países del entorno. Para ello vamos a usar dos bases de datos, EU-SILC y PISA:

1) EU-SILC es el acrónimo de *European Union Statistics on Income and Living Conditions*. EU-SILC es el instrumento principal para el estudio de la pobreza y la exclusión social en la UE. Nosotros vamos a usar la edición de 2019 ya que incorpora un módulo especial conocido como "Transmisión intergeneracional de desventajas, composición del hogar y evolución de la renta" en el que los participantes aportan información sobre sus familias cuando tenían catorce años. El objetivo es obtener datos sobre el entorno familiar en un momento crucial para el desarrollo del individuo. Entre otras cuestiones, los individuos indican qué nivel educativo y qué ocupación tenían sus padres, cuál era la situación económica del hogar, etc. Para el análisis de este capítulo nos centraremos en

individuos entre 30 y 59 años. De esta forma nos aseguramos de que han finalizado su periodo formativo fundamental.

2) PISA es el acrónimo de *Programme for International Student Assessment* que la OCDE realiza cada tres años desde el 2000. Mide las aptitudes de los estudiantes de quince años en tres áreas: Matemáticas, Lectura y Ciencias. En este capítulo usaremos las ediciones desde el 2000 hasta el 2018. La base de datos PISA proporciona, además de datos sobre resultados de los estudiantes en las pruebas de rendimiento de las tres disciplinas mencionadas, información a nivel individual sobre datos demográficos (por ejemplo, sexo, estatus migratorio, mes de nacimiento), antecedentes socioeconómicos (educación y ocupación de los padres), y variables a nivel de centro escolar (titularidad, ubicación, tamaño).

**World Inequality Database on Education**

Además de las bases de datos usadas en este capítulo existen otras con las que realizar análisis similares. Entre ellas destacamos la World Inequality Database on Education (WIDE) desarrollada por la UNESCO (https://www.education-inequalities.org). Esta base reúne datos de diversas encuestas nacionales de más de 170 países y tiene carácter interactivo de forma que es posible crear mapas, gráficos, infografías y tablas a partir de los datos, y descargarlos, imprimirlos o compartirlos en línea. La WIDE permite comparar resultados educativos entre países, y entre grupos dentro de un mismo país, en función de factores asociados a la desigualdad, como la riqueza, el género, la etnia y la ubicación.

Como se discutió previamente en los capítulos del 3 y 4 sobre la función de producción educativa, el logro educativo es el resultado de la combinación de diferentes variables, algunas innatas del estudiante como el sexo, la herencia genética de los padres, la educación, ocupación y renta de estos, y de otras que dependen del propio estudiante como el esfuerzo y dedicación que pone en los estudios, y finalmente las características de las escuelas (gasto público, compañeros de clase, etc.). En este capítulo mostraremos evidencia sobre desigualdad en resultados educativos en base a ciertas características socioeconómicas. Por tanto, la brecha educativa que mostramos pueden deberse también a estos otros factores que no observamos. Esta medida difiere de la desigualdad de oportunidades educativas, que cuantifica cuánta de esa desigualdad en resultados se

explica exclusivamente por circunstancias que los individuos heredan y escapan del control individual. En este sentido, y si bien algunos consideran aceptable cierta desigualdad educativa en los resultados, en tanto que podría reflejar diferencias en el esfuerzo o incluso diferencias en la capacidad innata, es menos frecuente estar de acuerdo en la existencia de desigualdad de oportunidades entre los estudiantes (ver las referencias clásicas de Roemer, 1998, o Fleurbaey, 2008). Al margen de este argumento de justicia social, la existencia de desigualdad de oportunidades educativas también es criticada en base a argumentos de eficiencia en la medida que puede dar lugar a una asignación ineficiente de los recursos. El trabajo de Ferreira y Gignoux (2014) analiza desigualdad de resultados educativos y de oportunidades educativas usando datos de PISA 2006 para países de la OCDE, y encuentra que la desigualdad de oportunidades explica hasta el 35% de todas las disparidades en el rendimiento educativo. Más recientemente, Marrero et al. (2023) estudian también desigualdad de oportunidades educativas usando PISA 2018 para un conjunto de países europeos y usan medidas más precisas de las circunstancias personales. Así, encuentran que la desigualdad de oportunidades educativas representa, en promedio, el 32% de la desigualdad educativa total en Europa occidental, con variaciones sustanciales entre países.

## 3. EVIDENCIA EN ESPAÑA

A continuación describimos la desigualdad educativa en España entre individuos procedentes de distintos grupos sociales en cuanto a tres resultados educativos, el abandono escolar, la obtención de titulación universitaria y los resultados en las pruebas de PISA.

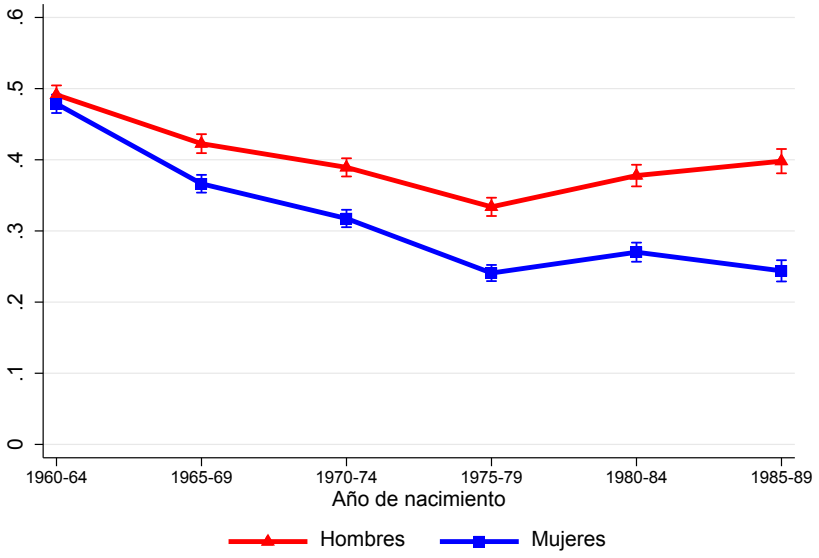
### 3.1. Abandono escolar y características familiares

Según Eurostat, la tasa de abandono temprano de la educación-formación es el porcentaje de personas entre 18 y 24 años que no completan la Educación Secundaria de segunda etapa y tampoco siguen ningún tipo de estudio-formación en las cuatro semanas anteriores a la de la entrevista realizada para la Encuesta de Población Activa (véase también el capítulo 1 de introducción descriptiva de este libro). En España la tasa de abandono para hombres ha sido el 21,7% en 2018, el 21,4% en 2019, el 20,2% en 2020, el 16,7% en 2021 y el 16,5% en 2022. Para las mujeres ha pasado del 14,0% en 2018, al 13,0% en 2019, al 11,6% en

2020, al 9,7% en 2021 y al 11,2 en 2022. En 2022 las cifras para la UE-27 fueron el 11,1% y el 8,0% para hombres y mujeres, respectivamente. Por tanto, aunque las cifras han mejorado en los últimos años, España sigue estando en peor situación que el resto de países de la UE.

En esta sección usamos datos de EU-SILC y proponemos como definición de abandono escolar el porcentaje de individuos que declaran tener un nivel educativo inferior a la Secundaria Superior y que difiere de la definición tradicional de abandono que sí incluye a los que han completado la Secundaria pero no han continuado estudiando. De esta forma podemos ver su evolución a lo largo del tiempo ya que los individuos de la muestra nacieron entre 1960 y 1989. Es importante mencionar que esta definición es distinta de la usada por EUROSTAT ya que no se centra solo en individuos entre 18 y 24 años. Para el conjunto de la muestra, el abandono escolar es un 37%. En el gráfico siguiente vemos la evolución de esta tasa de forma separada para hombres y mujeres nacidos entre 1960 y 1989. Nótese que como la encuesta se realiza en 2019, los más jóvenes en la muestra tienen 30 años y los más mayores 59 años.

**Gráfico 2. Proporción de individuos de cada cohorte que no completan Estudios Secundarios.**




Fuente: EU-SILC.

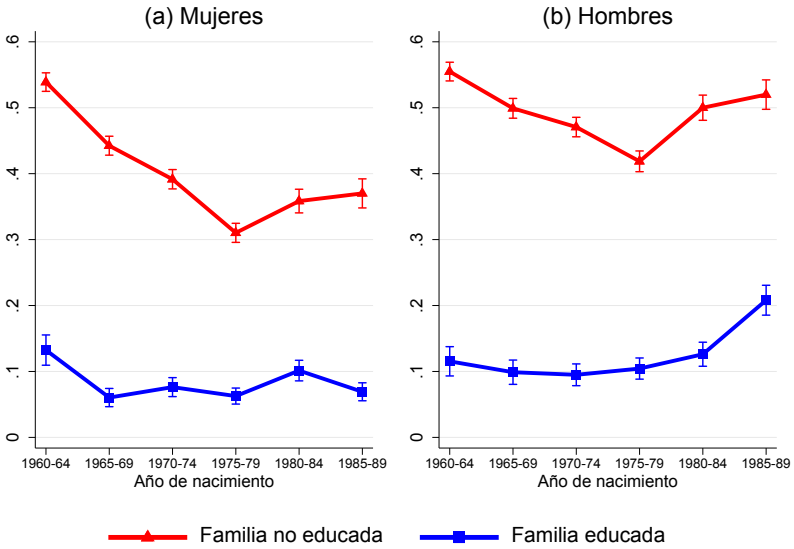
Vemos que la tasa de abandono se fue reduciendo paulatinamente para las generaciones nacidas en los sesenta y setenta, aunque después se estabiliza para los nacidos en los ochenta. También vemos que la divergencia entre hombres y mujeres aumenta paulatinamente, llegando a un máximo de unos quince puntos porcentuales en las cohortes más jóvenes de la encuesta, los nacidos en la segunda mitad de los ochenta (ver capítulo 4b).

Ejercicio 2. *Hemos visto dos medidas de abandono escolar, por un lado, la de Eurostat y por otro lado la que usamos nosotros con los datos de EU-SILC. Discute las diferencias entre ambas medidas. En concreto, ¿crees que la medida usada en el capítulo puede sobre o subestimar la tasa de abandono con respecto a la de Eurostat?*

A continuación analizamos la desigualdad en la tasa de abandono en función de las características familiares. Nuestro criterio principal es el nivel educativo de los padres. Multitud de autores han documentado la existencia de una asociación positiva y fuerte entre los niveles educativos de padres e hijos (ver Holmlund et al., 2011). En nuestro caso vamos a dividir a los individuos en dos grupos de acuerdo a la educación de sus padres. Clasificamos a un individuo como perteneciente a una familia "educada" cuando o bien su madre o su padre (o ambos) completaron la Educación Secundaria Superior. Clasificamos al resto de individuos dentro de familias "no educadas". Es decir, ni el padre ni la madre tienen más que la Secundaria Obligatoria. La proporción de familias educadas en la muestra es ligeramente inferior al 25%, aunque va aumentando con el tiempo. Entre los nacidos en 1960-1964, algo menos del 15% crecen en familias educadas, pero esta proporción supera el 40% para los nacidos en 1985-1989. Si consideramos el conjunto de la muestra, la tasa de abandono para mujeres y hombres de familias educadas es del 8,1% y 12,6%, respectivamente. Por el contrario, para los individuos con padres con nivel educativo bajo, dichas tasas se disparan hasta el 41,4% y 49,3%, respectivamente. En el gráfico 3 representamos la evolución de dichas tasas a lo largo del tiempo.

Lo más destacable del gráfico 3 son las enormes diferencias entre ambos grupos. Aunque las diferencias se reducen paulatinamente para los nacidos en los sesenta y setenta, para los nacidos en los ochenta se mantienen e incluso aumentan para las mujeres. Para las cohortes más recientes, los nacidos entre 1985-1989, la brecha de abandono escolar es de unos treinta puntos porcentuales. Esencialmente, el abandono se concentra en las personas cuyos padres tienen un nivel educativo bajo. Esto tiene implicaciones muy importantes para la movilidad intergeneracional ya que indica que las desventajas educativas de los padres se transmiten a sus hijos.

**Gráfico 3. Proporción de individuos de cada cohorte que no completan Estudios Secundarios por tipo de familia.**

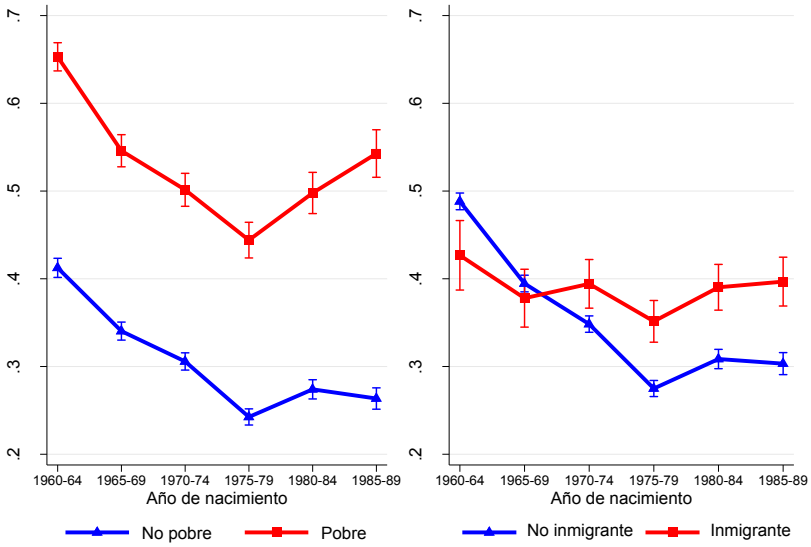



Fuente: EU-SILC.

Para completar el análisis, estimamos un modelo de regresión lineal en el que la variable a explicar es una variable binaria que toma el valor uno cuando el individuo no completa los Estudios Secundarios Superiores y la variable explicativa principal es otra variable binaria que toma el valor uno si la familia no es educada. Además, incluimos como variables explicativas adicionales efectos fijos de Comunidad Autónoma, año de nacimiento, sexo, padre o madre con nacionalidad española, o entorno rural. El coeficiente estimado de la variable explicativa principal es 0,277 (desviación estándar 0,007). Esto quiere decir que la probabilidad de no completar los estudios secundarios es 27,7 puntos porcentuales mayor para un individuo en una familia de educación baja que para otro de una familia con educación alta. El tamaño de este efecto es muy grande, ya que la media de la variable a explicar (probabilidad de no completar la Secundaria) es 36,6%.

Ejercicio 3. *En la regresión lineal que acabamos de ver, también incluimos una variable que toma el valor uno para las mujeres (y cero para los hombres). El coeficiente estimado de esta variable es –0,069 (desviación estándar 0,007). ¿Cómo interpretas este coeficiente? También incluimos otra variable que toma el valor uno si el individuo se crió en una zona rural (cero en otro caso). El coeficiente estimado de esta variable es 0,146 (desviación estándar 0,009). ¿Cómo interpretas este coeficiente? ¿Te parece razonable que tome un valor positivo? ¿Por qué?*

**Gráfico 4. Proporción de individuos de cada cohorte que no completan Estudios Secundarios. Por tipo de familia.**




Fuente: EU-SILC.

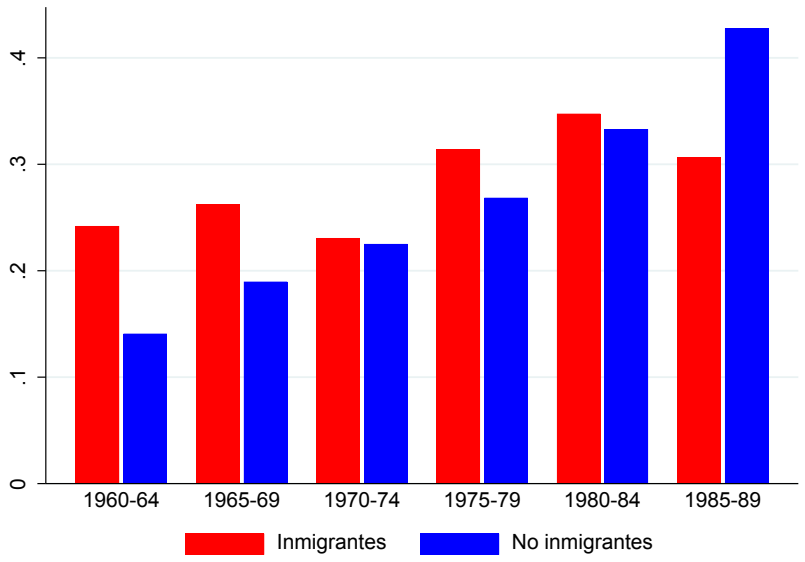
La base de datos EU-SILC también cuenta con otras características interesantes sobre el entorno familiar. Por ejemplo, sobre cómo era la situación financiera en la familia cuando el individuo tenía catorce años o sobre la nacionalidad de los padres, entre otras. Una de las cuestiones más llamativas que contestan los individuos es si cuando tenían 14 años tenían sus necesidades escolares básicas cubiertas. Esto se refiere a los libros y otro equipo para la escuela. A continuación usamos algunas de estas variables para profundizar en los resultados anteriores. En concreto, en el gráfico siguiente representamos en el panel de la izquierda las tasas de abandono dividiendo a los individuos en función de si crecieron en familias pobres (la situación financiera familiar cuando tenían catorce años era muy mala, mala o moderadamente mala) o no. En el panel de la derecha, distinguimos entre nativos e inmigrantes, incluyendo en este segundo grupo a todos los individuos que no han nacido en países de la UE (UE-15)[1].

[1] Consideramos también como nativos a aquellos que declaran que o bien el padre o bien la madre es de nacionalidad española.

En el panel de la izquierda el resultado es similar al gráfico 3. Entre individuos procedentes de familias pobres, la tasa de abandono escolar es muy elevada, en torno al 50%. Para los individuos de familias sin problemas financieros la tasa de abandono ha pasado de estar ligeramente por encima del 40% a bajar del 30%. Como consecuencia, la desigualdad educativa asociada al estatus de pobreza familiar se mantuvo estable para las generaciones entre 1960 y 1979 y se incrementó para las posteriores.

En el panel de la derecha vemos que entre los nacidos en los años 60 las tasas de abandono son similares para nativos e inmigrantes. Sin embargo, a partir de los setenta la situación cambia ya que las tasas de abandono se reducen para los no inmigrantes, mientras que para los inmigrantes apenas hay cambios. Una posible razón de este hecho es que a lo largo del periodo ha cambiado la composición de la población inmigrante. Para ver esto con más detalle, en el gráfico siguiente representamos qué porcentaje de cada grupo (inmigrantes vs no inmigrantes) tenía padres con un nivel educativo superior a la Secundaria Obligatoria. Vemos que para los nativos este porcentaje aumenta paulatinamente a lo largo de todo el periodo. Por el contrario, para los inmigrantes esto no es así. De hecho ha habido un deterioro en el nivel educativo de las madres y padres de jóvenes inmigrantes entre la cohorte de los nacidos en 1980-1984 y los nacidos en 1985-1989.

**Gráfico 5. Inmigración y proporción de individuos que proceden de familias educadas.**



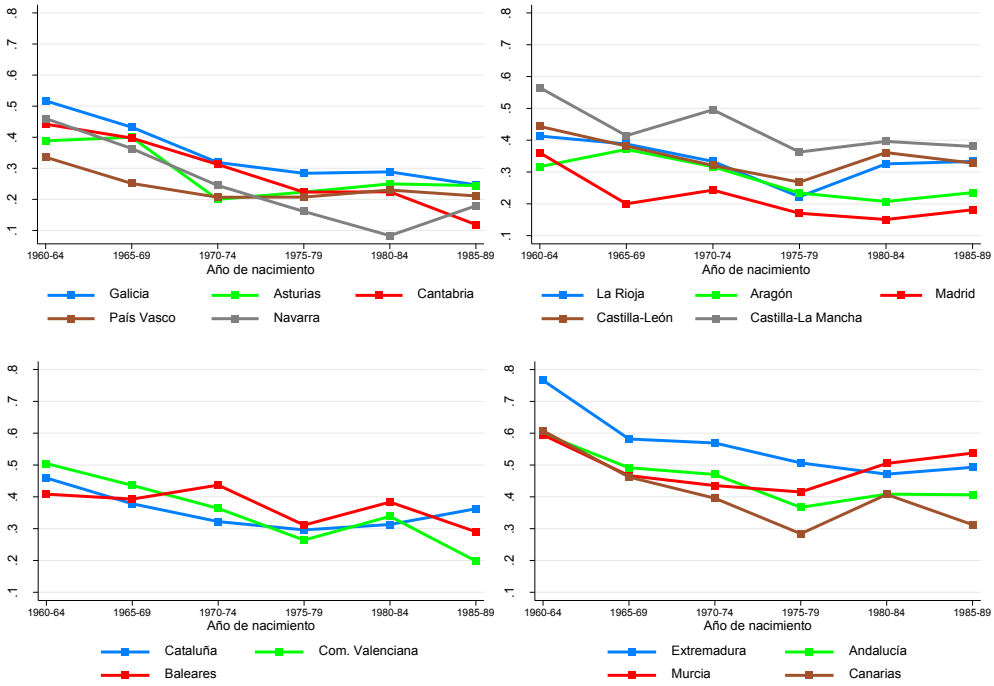

Fuente: EU-SILC.

Finalmente, un breve comentario sobre la cuestión de las necesidades escolares básicas cuando el individuo tenía 14 años. Del total de los encuestados, solo un 4% manifiesta que no tenía cubiertas dichas necesidades básicas, generalmente por problemas económicos. Esto puede verse como un indicador de circunstancias personales especialmente negativas. De hecho, la tasa de abandono escolar para este grupo es de un 68,8%, el doble que para el resto de individuos (35,3%).

### 3.2. Abandono escolar y desigualdades regionales

En España las competencias en materia educativa se encuentran transferidas a las Comunidades Autónomas. Este proceso descentralizador se realizó de forma gradual, comenzando a principios de los 80 y terminando en el año 2000. Esta fuerte descentralización, unida a diferencias de distinta índole entre regiones, puede dar lugar a diferencias tanto en resultados educativos como en

**Gráfico 6. Proporción de individuos de cada cohorte que no completan estudios secundarios por Comunidades Autónomas.**



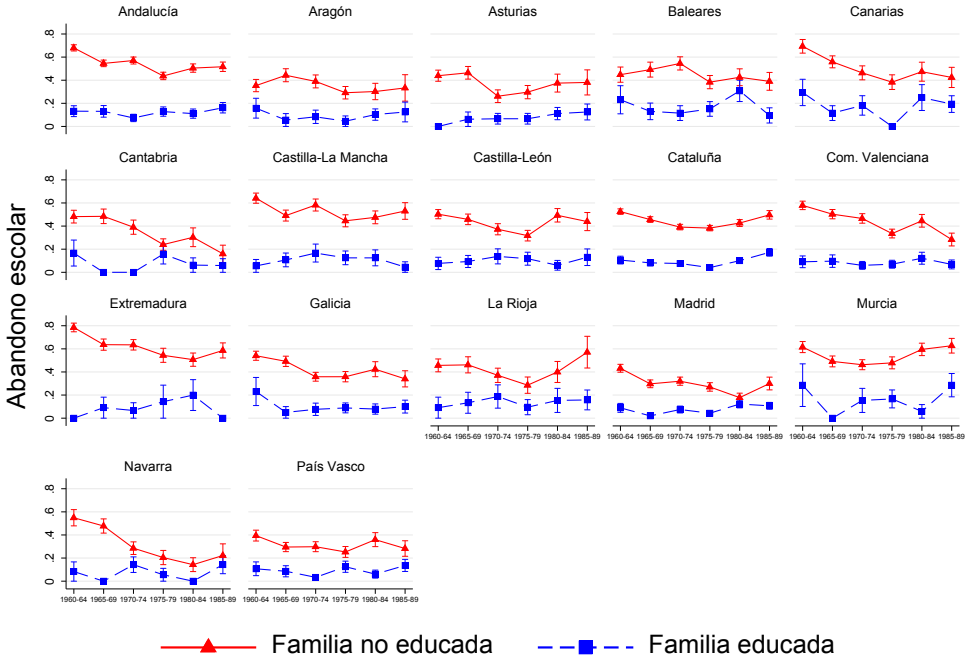

Fuente: EU-SILC.

desigualdad. En cualquier caso, las transferencias de competencias educativas podrían afectar a las dos últimas generaciones de las incluidas en EU-SILC. En esta sección analizamos las diferencias regionales en cuanto al abandono escolar, así como su evolución en los últimos años. Para ello realizamos un análisis similar al de la sección anterior pero desagregando a nivel de comunidades. En el gráfico 6 representamos la evolución de las tasas de abandono escolar en las diecisiete Comunidades Autónomas. Para visualizar mejor los datos las hemos distribuido en cuatro grupos: Norte, Centro, Este y Sur.

En el gráfico 6 vemos una clara división Norte-Sur, con las comunidades del Este en una posición intermedia entre aquellas. Por otro lado, parece que la mayor parte del estancamiento en los nacidos en los años ochenta se concentra en las regiones del Sur y centro de España (exceptuando Madrid). A continuación desagregamos las series en función del nivel educativo de la familia.

En general observamos que aquellas comunidades en las que la tasa de abandono es elevada también lo son las diferencias entre familias educadas y no educadas

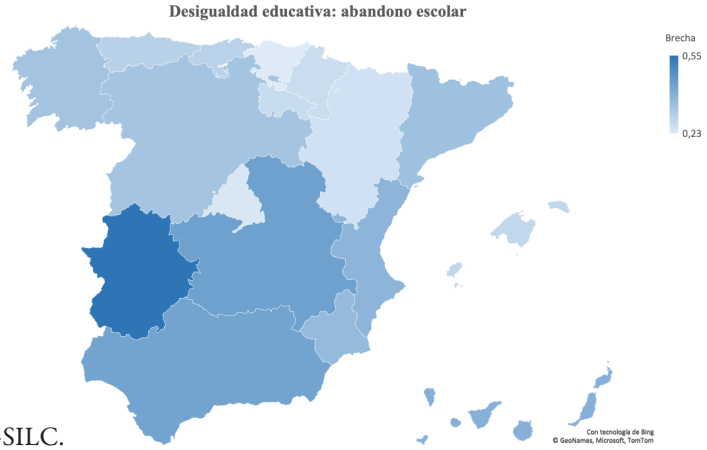
**Gráfico 7. Proporción de individuos de cada cohorte que no completan Estudios Secundarios por Comunidades Autónomas y tipo de familia.**




Fuente: EU-SILC.

(Centro y Sur frente a Madrid y Noroeste). En el mapa siguiente representamos la diferencia media en toda la muestra en la tasa de abandono escolar entre individuos de familias no educadas y educadas, por Comunidad Autónoma. Esta brecha en el abandono escolar por tipo de familia se encuentra entre el mínimo de los 23 puntos porcentuales del País Vasco y el máximo de los 55 de Extremadura.

**Gráfico 8. Brecha en abandono escolar por Comunidad Autónoma y tipos de familia.**

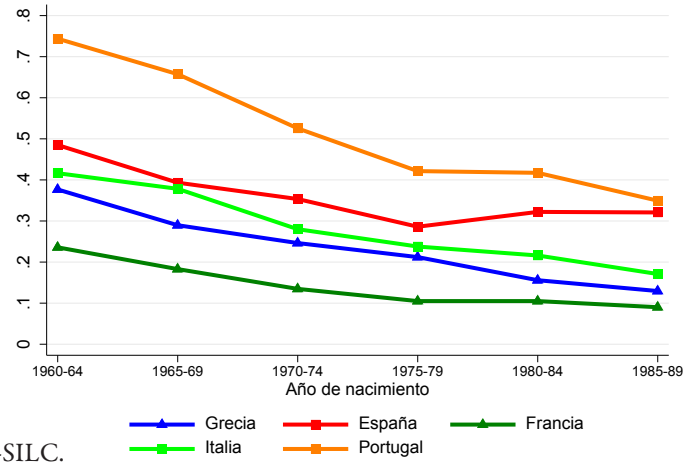



Fuente: EU-SILC.

## 3.3. España en perspectiva con otros países europeos

A continuación, vemos la evolución del abandono escolar, así como la desigualdad en este resultado por nivel educativo de las familias en otros países europeos comparables a España. En concreto nos centramos en Portugal, Francia, Italia y Grecia.

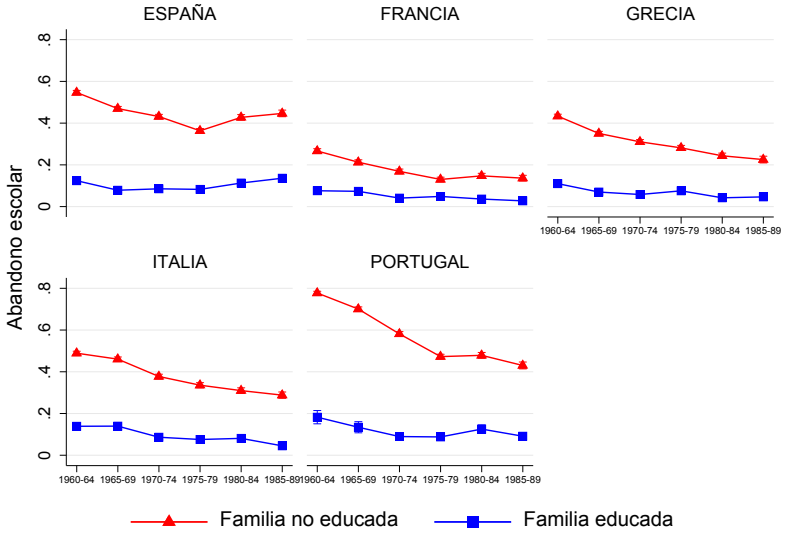
**Gráfico 9. Proporción de individuos de cada cohorte que no completan Estudios Secundarios. Varios países europeos.**




Fuente: EU-SILC.

En este grupo de países solo Portugal cuenta con tasas de abandono superiores a España. Sin embargo, Portugal partía de una situación mucho peor que la de España a comienzos de los sesenta con un abandono cercano al 75% frente al 50% de España. A partir de ahí, en Portugal se ha producido un descenso continuado mientras que en España el abandono escolar se ha mantenido estable en los ochenta, de forma que al final del periodo ambos países cuentan con tasas de abandono similares cercanas al 30%. A continuación, desagregamos la tasa de abandono por tipo de familia para estos mismos países.

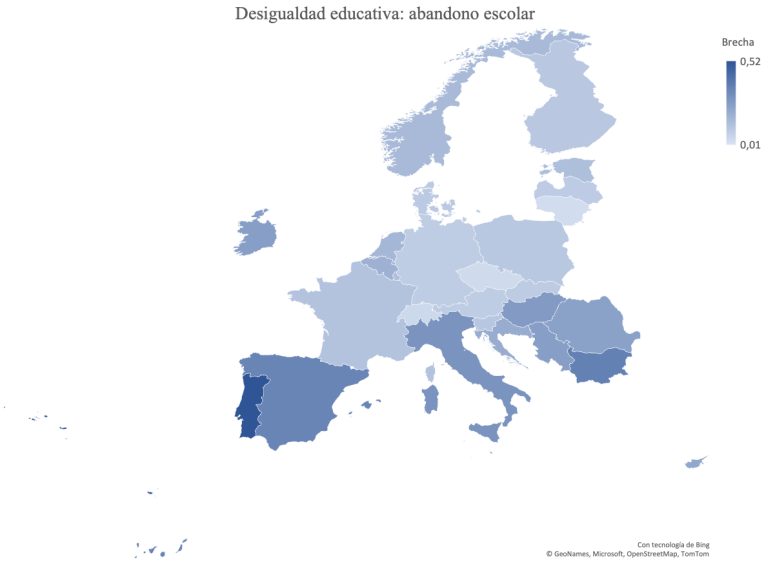
**Gráfico 10. Proporción de individuos de cada cohorte que no completan Estudios Secundarios por tipo de familia. Varios países europeos.**




Fuente: EU-SILC.

Este gráfico muestra que la desigualdad educativa en España está entre las más elevadas del conjunto de países europeos objeto de estudio. Además, a diferencia de otros países con brecha de abandono escolar por tipo de familia similares al inicio del periodo (por ejemplo, Italia o Grecia) en España la desigualdad apenas se ha reducido. En el mapa siguiente hacemos un análisis similar al que hicimos arriba para las Comunidades Autónomas. España vuelve a destacar por la diferencia en tasas de abandono entre ambos tipos de familias. Se observa así mismo una fuerte distinción entre los países del Sur, frente a la Europa continental y los países escandinavos.

**Gráfico 11. Brecha en abandono escolar por país y tipos de familia.**




Fuente: EU-SILC.

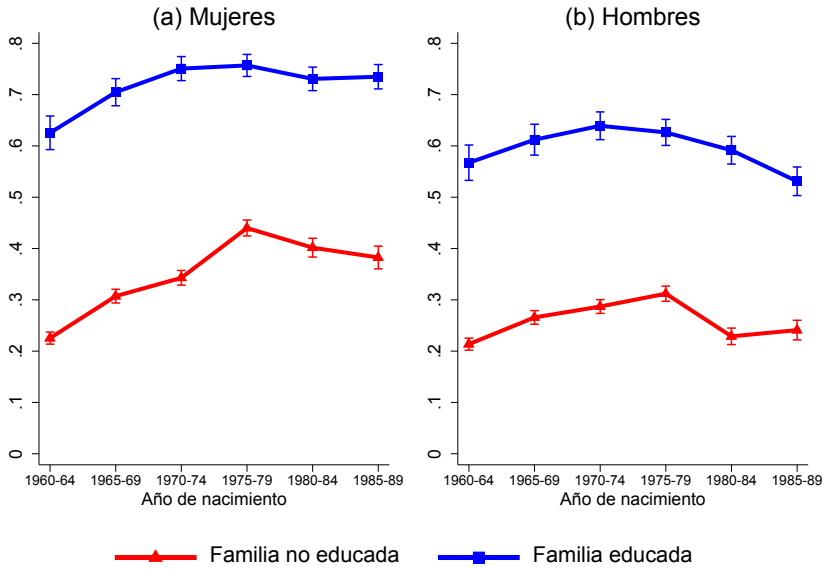
## 3.4. Desigualdad en el acceso a la Educación Universitaria

En la sección 3.1 analizamos la desigualdad en la finalización de Estudios de Secundaria, periodo en el que las decisiones de los padres, y por lo tanto sus características, pueden tener un peso mayor en los resultados educativos individuales. Sin embargo, en los países desarrollados como el nuestro y como indican Blanden et al. (2023), también existe desigualdad en la enseñanza superior. En esta fase, la influencia de los padres tiene un impacto menor si bien sigue estando presente (por ejemplo, en cuanto al campo de estudios, centro universitario a asistir, en qué momento incorporarse al mercado de trabajo, etc.)

A continuación analizamos como indicador de logro educativo el haber completado estudios universitarios. Según la OCDE (2022), en España el porcentaje de individuos de entre 25 y 34 años con formación superior en 2021 es del 49%, lo que muestra un incremento importante pues era tan solo el 34% en 2001 (ver capítulo 4b) en el que se describe con mayor detalle este fenómeno. En el gráfico siguiente representamos qué porcentaje de la población tiene estudios superiores (Universidad o FP Superior) en España, separando a los individuos de acuerdo al nivel educativo de los padres. La proporción de personas con estudios superiores en EU-SILC (nacidos entre 1960 y 1989) es del 39%, si bien va

aumentando con el grupo de edad al que pertenece el individuo. En concreto, pasa del 27,5% entre los nacidos en 1960-1964 a un 44,2% entre los nacidos en 1985-89. Distinguimos entre hombres y mujeres y analizamos la desigualdad en este indicador entre individuos de familias educadas y no educadas, como ya hicimos en el caso del indicador del abandono escolar.

**Gráfico 12. Proporción de individuos de cada cohorte con Estudios Universitarios.**



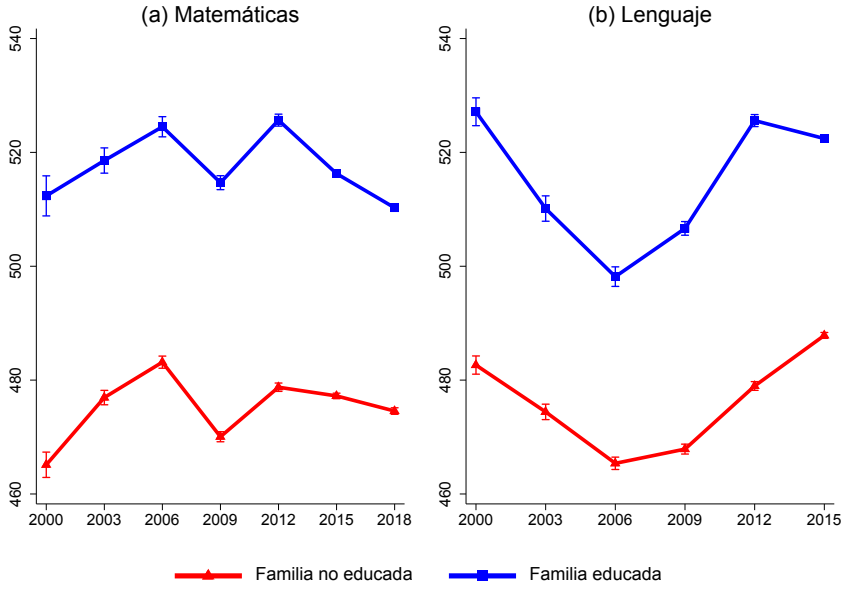

Fuente: EU-SILC.

El gráfico 12 es muy similar al gráfico 3. La proporción de mujeres universitarias es igual o superior a la de los hombres desde la generación de 1960 y tanto entre aquellas que provienen de familias educadas como las que no. El porcentaje de población universitaria aumentó ligeramente entre los hombres (de forma más acusada entre las mujeres) hasta la generación de la primera mitad de los años 70, si bien desciende desde entonces (o se mantiene estable entre las mujeres). La evolución de la población universitaria en este periodo es muy similar entre individuos con familias educadas y no, lo que implica que la desigualdad educativa se mantiene estable en estas generaciones. En concreto, para las cohortes más recientes de hombres la diferencia es de alrededor de 30 puntos porcentuales (22% *vs* 52%) y entre las mujeres algo superior (39% *vs* 72%). La brecha por tanto sigue siendo muy similar a la del abandono escolar (ver gráfico 3) a pesar de que la influencia familiar parecería ser menor.

### 3.5. Desigualdad en resultados escolares

Las diferencias en resultados educativos como el abandono escolar según el origen socio-económico del individuo pueden surgir a edades tempranas (para una revisión de esta evidencia ver Attanasio, 2015; o Currie y Almond, 2011). En esta sección analizamos la desigualdad educativa entre estudiantes de 15 años procedentes de distintos grupos sociales, en concreto el nivel educativo de las familias, su estatus de migrante, así como su estatus socioeconómico. Para ello vamos a usar los resultados de Matemáticas y Lengua en PISA. También usaremos como medida educativa el hecho de repetir curso, ya que esta es una circunstancia que puede considerarse un buen predictor del abandono escolar (ver, entre otros, Cockx et al., 2019). Los resultados escolares en Matemáticas son fuertes predictores de resultados futuros, tales como la probabilidad de finalizar estudios de Secundaria o asistir a la Universidad, o incluso los salarios y el empleo de los individuos en el mercado laboral (Joensen y Nielsen, 2009; Rose y Betts, 2004; Hanushek et al., 2015). Por otra parte, aprender a leer y escribir es una habilidad esencial. De hecho, la evidencia empírica muestra que las competencias básicas en Lectura y Escritura generan importantes ganancias salariales y de empleo en el mercado laboral (Cherry y Vignoles, 2016).

**Gráfico 13. Puntuaciones en Matemáticas y Lengua por tipo de familia.**



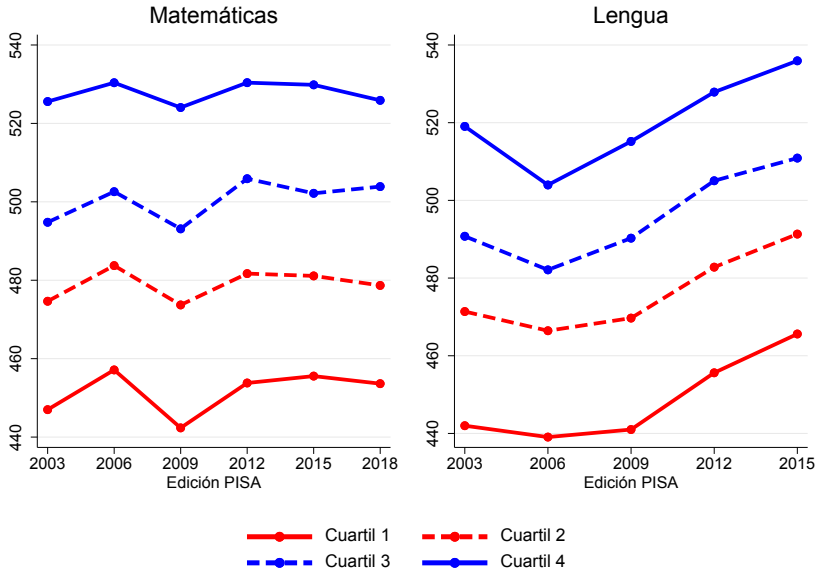

Fuente: PISA.

Para este análisis vamos a usar en concreto los datos de PISA para España desde la edición de 2000, es decir, individuos nacidos en 1985. De esta forma enlazamos con la última generación para la que estudiamos abandono escolar con los datos de EU-SILC. En los gráficos siguientes mostramos la desigualdad en puntuaciones PISA en Matemáticas y Lengua en función del nivel educativo de las familias. En esta ocasión consideraremos familia educada aquella con madre o padre con estudios universitarios. En el total de la muestra un 39,6% de los estudiantes son de familias con Educación Superior.

Observamos diferencias relativamente estables tanto en Matemáticas como en Lengua. En el último año para el que tenemos datos la diferencia en Matemáticas (2018) es de 35,7 puntos y en Lengua (2015) de 34,6 puntos (ver el capítulo 1 donde se explican los motivos por los que no es posible los datos de Lengua de 2018). Estos números representan un 45% y un 43% de la desviación estándar correspondiente, respectivamente.

De forma similar a como hicimos para el estudio de la desigualdad en abandono escolar, dividimos ahora a las familias de acuerdo a su nivel socioeconómico y de acuerdo a si son nativos o inmigrantes. Para medir el estatus socioeco-

**Gráfico 14. Puntuaciones en Matemáticas (izquierda) y Lengua (derecha) y ESCS.**




Fuente: PISA.

nómico de los estudiantes usamos el índice ESCS que proporciona PISA y que se construye a partir de tres variables relacionadas con el entorno familiar: nivel educativo más alto de los padres, estatus ocupacional más alto de los padres y el hogar, así como posesiones, incluidos los libros. En el gráfico 14 dividimos a los participantes en las pruebas de PISA en España desde 2003 (no usamos los datos de la edición del 2000 pues se usó una medida de índice socioeconómico no comparable con el resto de ediciones) en cuartiles de acuerdo con el ESCS y calculamos las puntuaciones medias en cada cuartil. Esto quiere decir que ordenamos a todos los individuos con respecto al valor del ESCS. Los individuos en el cuartil 1 son el 25% de individuos que tienen los valores más bajos del ESCS, los del cuartil 2 son el siguiente 25%, etc.

Observamos una gran distancia entre las puntuaciones de los estudiantes de los cuartiles 1 y 4. Para Matemáticas la distancia es de unos 72 puntos en 2018, equivalente al 90% de una desviación standard y la brecha es similar para Lengua. Además la trayectoria de las cuatro líneas es prácticamente paralela en los últimos 15 años.

Ejercicio 4. *En la base de datos de PISA tienes datos sobre el número de libros, ordenadores y tabletas del estudiante. Descarga los datos de PISA para España de las ediciones 2015 y 2018 y construye un gráfico similar al 14 usando estas variables. ¿Cómo es la desigualdad educativa en Matemáticas y Lengua entre alumnos que tienen menos de 25 libros y los que tienen más de 100? ¿Y entre los que tienen uno o ningún ordenador (o tableta) y los que tienen más de uno? ¿Es similar a la que observamos usando el índice socioeconómico? ¿Hay diferencias entre la desigualdad en Matemáticas y Lengua en función del tipo de posesión educativa?*

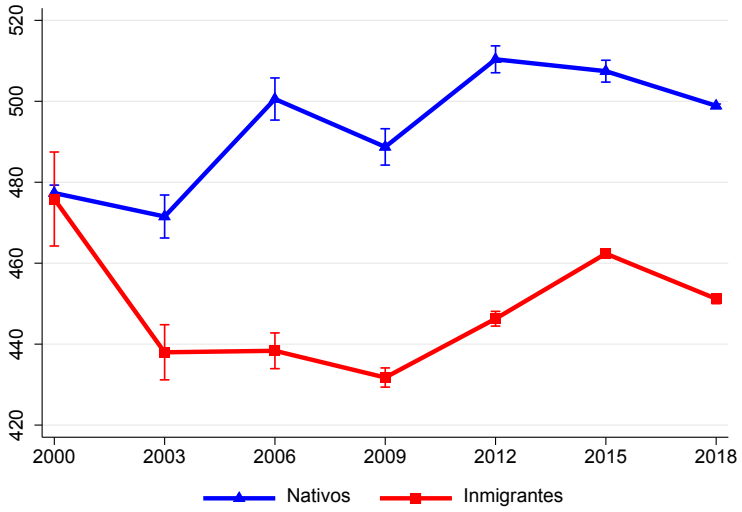
A continuación mostramos la desigualdad en puntuaciones PISA en Matemáticas entre estudiantes de familias inmigrantes y nativos. En el primer grupo incluimos tanto inmigrantes de primera como de segunda generación. Nos centramos en la nota de Matemáticas dado que las dificultades lingüísticas de algunos estudiantes inmigrantes pueden introducir errores de medición en la puntuación en Lengua para este grupo (existe evidencia que muestra que estas dificultades pueden influir en el rendimiento escolar en las clases de Lengua, ver Resnjanskij et al., 2023).

Como observamos en el gráfico 15, a partir de 2003 se abre una brecha entre ambos grupos (en la edición de 2000 la población de estudiantes inmigrantes en España es aún reducida por lo que sus puntuaciones no son representativas) que se ha mantenido relativamente estable desde entonces.

### 3.6. Desigualdad en cuanto a repetición de curso

La repetición de curso es una práctica muy extendida en España, que cuenta con una de las tasas más altas de la OCDE. Entre los participantes en PISA 2018, más del 25% han repetido curso al menos una vez en Primaria o Secundaria, mientras que la media para la OCDE es de solo un 11% y en 34 países es inferior al 5% (OECD, 2020). Existe, además, cierta evidencia de que los

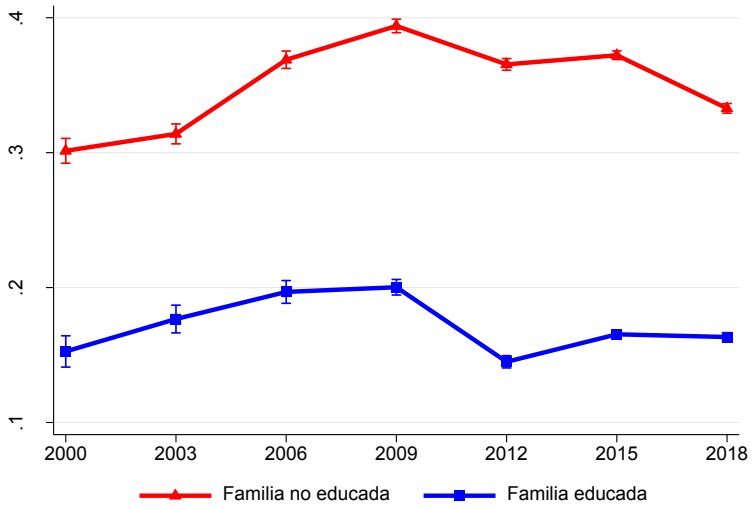
**Gráfico 15. Puntuaciones en Matemáticas y estatus migrante.**




Fuente: PISA.

alumnos de entornos desfavorecidos tienen más probabilidades de repetir curso. Según los datos de PISA 2018 para España, la proporción de repetidores por cuartil del índice ESCS es el 45% en el cuartil 1, el 30% en el cuartil 2, el 17% en el cuartil 3, y del 8% en el cuartil 4. Un alumno del primer cuartil tiene una probabilidad 5,6 veces mayor que un alumno aventajado (cuarto cuartil) de haber repetido curso al menos una vez. Teniendo esto en cuenta, nos preguntamos cómo es la desigualdad educativa en cuanto a la repetición de curso en España, y cómo ha sido su evolución reciente. Para ello seguimos usando la base de datos de PISA desde 2000 hasta 2018. En el gráfico 16 analizamos la tasa de repetidores en función del nivel de educativo de la familia. Observamos que la brecha en repetición de curso es elevada, en torno a 15 puntos porcentuales. Además, se ha mantenido estable o ha aumentado ligeramente a lo largo del periodo.

**Gráfico 16. Proporción de estudiantes que repiten curso por tipo de familia.**




Fuente: PISA.

Ejercicio 5. *Como indican algunos trabajos, el hecho de repetir curso está asociado al abandono escolar. Si bien usando la base de datos de PISA no tenemos información sobre si el estudiante abandona los estudios o no, sí tenemos información sobre sus expectativas educativas. En concreto, existe una variable que recoge el nivel educativo más alto que el estudiante espera completar. Descarga los datos de PISA para España de las ediciones 2000 al 2018. ¿Cómo son las expectativas de abandono escolar para los alumnos que repiten y no repiten? ¿Existe alguna diferencia? ¿Cómo ha evolucionado?*

**Repetición de curso**

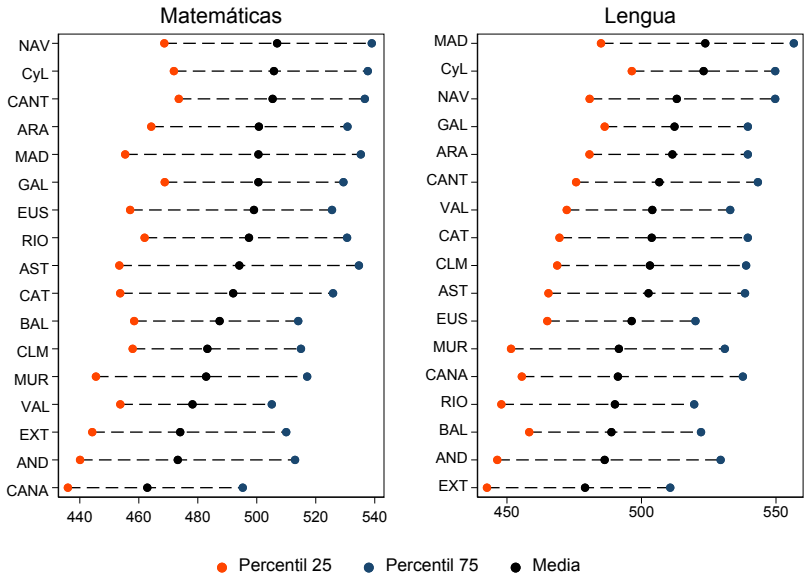
El objetivo de la repetición es dar a los alumnos una "segunda oportunidad" para dominar los conocimientos y las destrezas adecuados a su nivel. Sin embargo, sigue siendo una práctica objeto de debate tanto dentro de la comunidad educativa como científica, pues no es obvia la identificación de sus efectos y la evidencia es mixta (Allen et al., 2009). Entre algunos de los trabajos más recientes, Alet (2014) detecta que las ganancias a corto plazo en las puntuaciones de los exámenes tienden a desaparecer pocos años después de la repetición. García-Pérez et al. (2014), encuentran efectos negativos de repetir curso sobre el rendimiento académico en España. Por último, los alumnos que repiten curso tienen más probabilidades de abandonar la enseñanza Secundaria (Manacorda, 2012; Jacob y Lefgren, 2009; Cockx et al., 2019).

### 3.7. Desigualdad educativa en puntuaciones dentro y fuera de España

En este apartado analizamos brevemente las diferencias regionales en cuanto a los datos de PISA. Nótese que todos los estudiantes incluidos en esta base de datos están afectados por la descentralización educativa. En el gráfico 17 comparamos la desigualdad en puntuaciones entre Comunidades Autónomas usando el criterio de estatus socioeconómico. Para cada Comunidad Autónoma, la gráfica muestra la puntuación media en la evaluación PISA de 2018 en Matemáticas (panel izquierdo) y PISA de 2015 en Lengua (panel derecho), las puntuaciones medias del cuartil inferior y superior del índice PISA de estatus socioeconómico (ESCS), así como la brecha entre ambos.

A partir del gráfico 17 podemos hacer varias observaciones. En primer lugar, la brecha en los resultados de los exámenes entre los cuartiles superior e inferior del estatus socioeconómico oscila entre 51 y 81 puntos en Matemáticas y entre 53 y 83 en Lengua. Estas cifras, aun siendo elevadas, se encuentran por debajo de la brecha media familiar en la OCDE, que es de 86 en Matemáticas y 84 puntos en Lectura (ver Blanden et al., 2023). En segundo lugar, en todas las Comunidades Autónomas, y tanto en Matemáticas como en Lengua, el rendimiento de los alumnos procedentes de entornos desfavorecidos está por debajo de la media de la OCDE de 500 puntos.

**Gráfico 17. Puntuaciones en Matemáticas y Lengua y ESCS por CC.AA.**

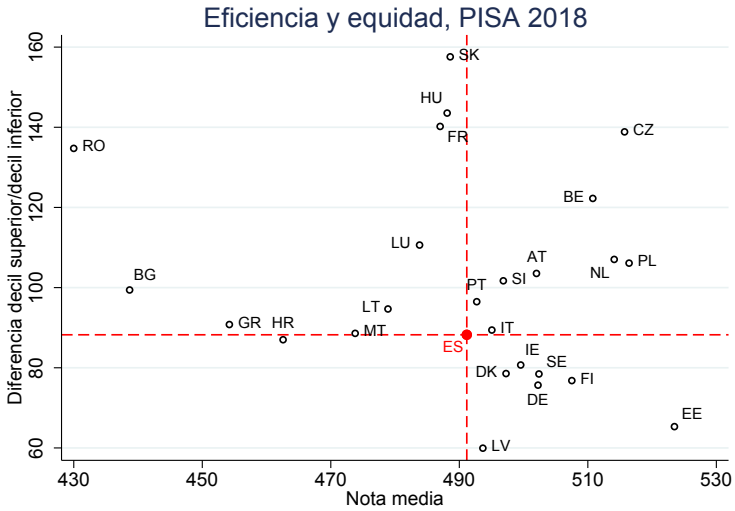



Fuente: PISA 2015 y 2018.

Por último, mientras que la brecha media entre Comunidades Autónomas no llega a los 45 puntos tanto en Matemáticas como en Lengua, la brecha media por entorno familiar es de 66 puntos en Matemáticas y de 68 puntos en Lectura.

Conocida la desigualdad en puntuaciones escolares dentro de España, vamos a intentar ponerla en perspectiva con el resto de países de la UE. Para ello nos centramos en las puntuaciones en Matemáticas en PISA 2018 y el índice ESCS. En el gráfico siguiente vamos a representar la relación entre puntuaciones medias y desigualdad. Para ello, en el eje horizontal representamos la puntuación media en Matemáticas en 2018. En el eje vertical vamos a representar una medida de desigualdad de notas en función del índice ESCS. Para ello calculamos en cada país la nota media que obtienen en Matemáticas los individuos del decil inferior y superior del índice ESCS. Es decir, primero ordenamos a los individuos en cada país de menor a mayor de acuerdo al valor del índice y, a continuación, los dividimos en diez grupos (deciles) de igual tamaño. Los índividuos del decil inferior son los más desfavorecidos, mientras que los del decil superior son los que cuentan con mayores ventajas. Finalmente, la medida de desigualdad que proponemos es la diferencia de notas medias entre el decil superior y el inferior. Cada punto representa un país y marcamos la posición de España en color rojo. Aparentemente España ocupa una posición central en el gráfico. Los países que disfrutan de una mejor

**Gráfico 18. Puntuación media en Matemáticas y diferencias entre el decil superior y el decil inferior.**




Fuente: PISA 2018.

posición que España son los del cuadrante Sureste, ya que tienen una media superior a España y a su vez una diferencia entre deciles menor. En el gráfico hay siete países en ese cuadrante, Alemania, Dinamarca, Estonia, Finlandia, Irlanda, Letonia y Suecia. Los que están en una posición peor son los del cuadrante Noroeste, ya que tienen una media menor y una diferencia de cuartiles mayor. Estos ocho países son Bulgaria, Eslovaquia, Francia, Grecia, Hungría, Lituania, Luxemburgo y Rumanía. Finalmente, el grupo de países del cuadrante Noreste están mejor en una dimensión (media) y peor en la otra (diferencia de deciles).

## 4. CONSECUENCIAS FUTURAS DE LA DESIGUALDAD EDUCATIVA ACTUAL

La existencia de individuos que se quedan atrás significa que estas personas tendrán menos opciones en su vida laboral, lo que reducirá la movilidad social. Por movilidad social nos referimos a la capacidad de que las personas puedan superar a lo largo de su vida las desventajas de haber nacido en una familia con bajo nivel socioeconómico. Si la mayor parte de los que nacieron en familias pobres o con poca educación acaban teniendo un nivel educativo bajo, el resultado será una baja movilidad social. Hablamos de movilidad alta cuando el nivel educativo de una persona es relativamente independiente del nivel educativo que alcanzaron sus padres. La desigualdad y la falta de movilidad suelen ser dos factores que van unidos. La llamada "Curva del Gran Gatsby" ilustra esta relación mostrando que existe una asociación positiva entre la desigualdad económica y la falta de movilidad social intergeneracional a nivel de países.

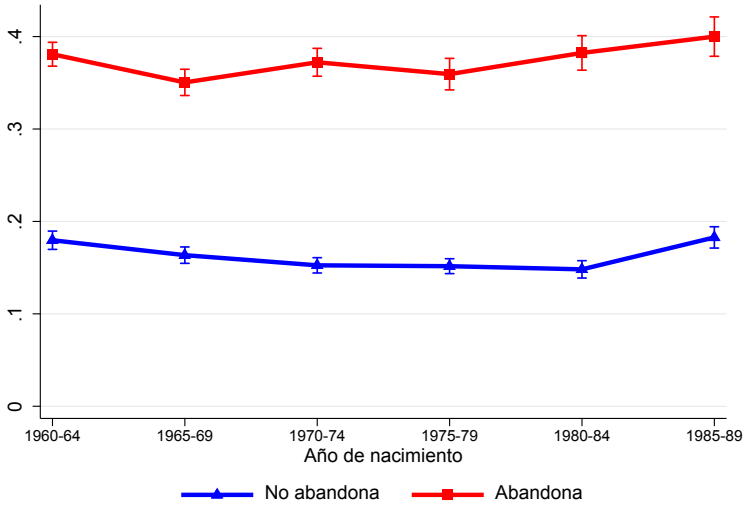
La desigualdad educativa actual puede no solo incrementar la desigualdad económica futura sino también agravar los problemas de pobreza ya existentes. Según las cifras más recientes de Eurostat (2023) España es el cuarto país de la UE con mayor porcentaje de población en riesgo de pobreza y exclusión social. El concepto de riesgo de pobreza o exclusión social (AROPE por sus siglas en inglés) que usa Eurostat se refiere a las personas que viven en una familia que entra en al menos una de las tres categorías siguientes:

1) La renta disponible de la familia está por debajo del umbral de pobreza (60% de la renta mediana del país).
2) La familia sufre privaciones materiales importantes.
3) En el hogar hay adultos que trabajaron menos del 20% de su tiempo potencial durante el año anterior.

La tasa AROPE es la proporción de la población total que se encuentra en dicho riesgo de pobreza o exclusión social. Este es el indicador principal para los estudios de pobreza y exclusión social de la UE. La tasa AROPE en 2022 en España fue del 26%, una cifra similar a la de Letonia, y solo por debajo de las de Rumanía (34,4%), Bulgaria (32,2%) y Grecia (26,3%). La tasa media para los países de la UE en 2022 fue del 21,6%. En el gráfico 19 mostramos la tasa AROPE en función de si el individuo abandonó estudios antes de finalizar la etapa Secundaria o no.

Como vemos en el gráfico 19, los individuos que abandonaron los estudios tienen aproximadamente el doble de probabilidad de padecer riesgo de pobreza y exclusión. Si unimos la información de este gráfico con lo que vimos en los gráficos 3 y 4, nos encontramos con una situación de trampa de pobreza. Las personas que se criaron en familias pobres o con poca educación abandonan antes la educación, lo que le condena a un mayor riesgo de pobreza y exclusión.

**Gráfico 19. Abandono escolar y riesgo de pobreza y exclusión (AROPE).**




Fuente: EU-SILC.

## 4.1. Desigualdad educativa y movilidad social en un contexto de cambio de estructura productiva

Como comenta Breen (2022), entre los años 50 y 80 muchos países europeos registraron elevadas tasas de movilidad social. Entre otros posibles factores, esto podría explicarse por ser también este un periodo de expansión de puestos direc-

tivos y profesionales. Esto favoreció el ascenso de personas de entornos menos favorecidos sin excluir necesariamente a los niños de entornos favorecidos de dichos puestos. Sin embargo, en las últimas décadas del siglo XX, estos cambios estructurales se atenuaron y la situación que prevalece ahora en muchos otros países europeos es una mayor desigualdad económica (Blanden et al., 2023) y de poca movilidad social (Breen y Müller, 2020). Las consecuencias que puede tener esto sobre la desigualdad educativa, y cómo a su vez la educación puede contribuir a reducir desigualdad económica en un contexto de cambio de modelo productivo como consecuencia de la transformación digital, la automatización de tareas, etc., es una cuestión que merece la pena seguir explorando (OECD, 2022).

## 5. CAUSAS DE LA DESIGUALDAD Y POSIBLES SOLUCIONES

Hay múltiples causas de desigualdad entre individuos que provienen de distintos entornos familiares (Piketty, 2000). Una primera vía es sin duda la transmisión genética de habilidades, cuyo efecto puede verse reforzado en presencia de emparejamientos selectivos entre individuos con habilidades similares (ver Clark, 2021). Por otro lado, a pesar de que hay un acceso generalizado a la enseñanza, las familias con más educación habitualmente tienen también más recursos económicos lo que les permite gastar más en bienes y servicios educativos complementarios (por ejemplo, clases de idiomas o clases de apoyo extraescolares). El mayor nivel educativo de los padres también puede traducirse, entre otros, en un mayor interés y preocupación por la educación de los hijos ya que han aprendido a valorar más esta y también en una mayor mayor disponibilidad de tiempo para ayudarles en las tareas escolares. También es obvio que, cuanto mayor es el nivel educativo de los padres, estos son más capaces de ayudar a los hijos en la materias más complicadas (Matemáticas). El estatus laboral de los padres también se ha observado que tiene efectos sobre la educación de los hijos (ver Ruíz-Valenzuela, 2020 para una revisión de la literatura más reciente). Por último, el entorno social incluyendo el vecindario y su impacto en la elección escolar, así como en las amistades de los hijos, ha sido discutido previamente en el capítulo 3c sobre los efectos de los compañeros de clase. Esta influencia en la producción de capital humano también es un factor clave para explicar la desigualdad educativa. Además, esta desigualdad puede intensificarse con el aumento de la segregación escolar, como se analiza en el capítulo 5a sobre elección escolar.

La interdependencia entre todas estas vías que originan la desigualdad educativa dificulta la identificación del factor quizá más relevante o a partir del cual se alimenta la generación de desigualdad educativa a lo largo de todo el proceso de acumulación de capital humano. En este capítulo hemos realizado un análisis descriptivo de la desigualdad educativa, sin pretender establecer una relación de causalidad entre los factores que hemos considerado (educación y estatus socioeconómico familiar) y la desigualdad que observamos. A continuación comentamos algunos trabajos que han tratado de identificar efectos causales entre determinadas medidas de características sociofamiliares y resultados educativos.

La evidencia sobre la relevancia de las características sociofamiliares en la educación individual es abundante. Björklund y Salvanes (2011) resumen esta literatura empírica, y analizan los principales mecanismos detrás de esta relación, con el objeto de aislar el peso de las características de los padres. Estos autores encuentran que las correlaciones de resultados educativos entre padres e hijos pueden de hecho subestimar el papel de las características sociofamiliares. Para ello usan las correlaciones de años de escolarización entre hermanos. Esta correlación puede considerarse una medida global de la importancia de la familia y el entorno que refleja las influencias de los padres, pero también otras influencias compartidas por los hermanos que son ortogonales a las características de los padres. Tras comparar las correlaciones entre hermanos y las intergeneracionales concluyen que los hermanos comparten ventajas educativas en mucha mayor medida de lo que reflejan las correlaciones entre padres e hijos, lo que implica que las correlaciones intergeneracionales representan sólo una pequeña parte de los efectos del entorno familiar. Por otra parte, encuentran que la educación de los padres sólo explica una pequeña parte de los factores que comparten los hermanos.

Entre los trabajos que estudian posibles relaciones causales, cabe destacar la revisión de la literatura de Holmlund et al. (2011) que se centra en cuantificar el efecto causal de la educación de los padres sobre los resultados educativos y laborales de los hijos aislándolo de otros factores anteriormente comentados como la habilidad individual. Así, basándose en una aplicación con datos de Suecia, concluyen que las correlaciones intergeneracionales de educación se deben en gran medida a la selección y no a efectos causales directos, aún siendo la educación de los padres el motor principal. Más reciente, Blingley et al. (2021), usando datos censales de individuos nacidos entre 1965-1985 en Dinamarca, profundizan en la influencia relativa de familia y entorno (vecinos y compañeros de colegio)

tanto en resultados educativos como laborales. Así, encuentran que la familia es el principal factor explicativo de la correlación entre ingresos y nivel educativo de los hermanos, siendo además la prevalencia de este efecto mayor para mujeres que para hombres.

En la medida en que buena parte de la desigualdad educativa surge a edades tempranas (ver Cunha y Heckman, 2007 entre otros), las intervenciones educativas en estas etapas parecen especialmente efectivas. Entre los primeros programas destinados a mejorar las oportunidades educativas de niños de entornos desfavorecidos destacan el Perry Preschool Program, Abecedarian y el CARE Programs, todos desarrollados en los años 60-70 en Estados Unidos. La evidencia sobre el impacto de estos programas es contundente. Un resumen reciente de la misma puede encontrarse en Elango et al. (2015), que comparan además los beneficios de distintos tipos de intervenciones de acuerdo a su carácter universal o no. La evidencia sobre los efectos de los programas de Educación Infantil, que se ha expandido de forma notable desde los 70, es más ambigua. Sin embargo, y a pesar de que sí que se observan impactos positivos entre los niños de entornos socioeconómicos menos favorecidos (ver por ejemplo Cornelissen et al., 2018), es posible que esta expansión no haya llegado de forma completa precisamente a aquellos grupos que más podrían beneficiarse de ellas (ver capítulo 4a para un análisis pormenorizado del papel de la Educación Infantil).

Otras intervenciones que pueden contribuir a reducir desigualdades educativas son los programas de educación compensatoria. Estos están diseñados para ayudar a los estudiantes con peores resultados a alcanzar niveles académicos mínimos, normalmente mediante la asistencia a clases en horario extraescolar y en grupos de tamaño reducido. La mayoría de estas intervenciones se realizan en la adolescencia cuando, como hemos visto a lo largo del capítulo, las diferencias en resultados educativos entre estudiantes de diferente entorno socioeconómico son muy evidentes. De hecho, Guryan et al. (2023), centrándose en los adolescentes beneficiarios del programa Saga Education en Chicago, muestran que la relación coste-beneficio estimada es comparable a la de muchos programas de éxito en la primera infancia.

Este tipo de intervenciones suscitan cada vez más interés, especialmente en los países desarrollados (ver, por ejemplo, Carcillo et al., 2015 para una panorámica de estos programas en países de la OCDE). La evidencia sobre su impacto en los resultados educativos sin embargo es aún escasa: Lavy et al. (2022) encuentran efectos positivos en los ingresos a largo plazo de un programa en

estudiantes de Secundaria en Israel; García-Pérez e Hidalgo-Hidalgo (2017) encuentran mayores puntuaciones en PISA entre estudiantes en centros que siguieron el programa PROA en España; Battaglia y Lebedinski (2015) centrándose en una intervención dirigida a estudiantes de etnia gitana en Serbia no encuentran efectos significativos sobre sus calificaciones o en abandono escolar. Este tipo de intervenciones no son sin embargo exclusivas de etapas educativas secundarias. Alan et al. (2019) encuentran efectos positivos sobre las habilidades no cognitivas de una intervención sobre niños de Primaria en Turquía; Battistin y Meroni (2016), que analizan un programa de educación compensatoria en escuelas de Primaria en Italia, encuentran efectos positivos sobre las notas de Matemáticas. El debate en torno al momento más adecuado para realizar este tipo de intervenciones está abierto (Battaglia e Hidalgo-Hidalgo, 2023).

## 6. OTRAS DESIGUALDADES Y COMENTARIOS FINALES

A lo largo del capítulo hemos ilustrado la desigualdad educativa existente entre individuos de distintos grupos sociales y nos hemos centrado en el nivel de estudios o estatus socioeconómico familiar para identificar esos grupos. Sin embargo, hay otras dimensiones que también generan desigualdad educativa, entre ellas la pertenencia a minorías étnicas. La comunidad gitana es la minoría de mayor tamaño en la UE: entre 6-8 millones de personas. La cifra es muy superior si nos referimos a toda Europa, alcanzando los 10-12 millones, es decir si añadimos los Balcanes y otros países más al Este, como Moldavia, etc. (Comunicación de la Comisión Europea, 2011). En España hay en torno a 700.000 gitanos, que viven en condiciones socioeconómicas precarias. La desigualdad educativa entre esta minoría y el resto de la población es enorme. Por citar un resultado educativo del que hemos hablado en este capítulo, la tasa de abandono escolar entre la población gitana de 18-24 en 2022 fue del 86,3%, frente al 13,3% del total de la población española en 2021 (para profundizar en este y otros indicadores ver el reciente Informe Ejecutivo realizado por la Fundación Secretariado Gitano, 2023). Esta desigualdad impulsó la recomendación del Consejo de Europa para mejorar el acceso a la educación al alumnado gitano y combatir todo tipo de desigualdad en el ámbito educativo[2]. Entre las posibles medidas para alcanzar

[2] Recomendación del Consejo de 12 de marzo de 2021 sobre la igualdad, la inclusión y la participación de la Población Gitana (2021/C93/01).

estos objetivos y acerca de las que existe evidencia empírica, destacamos de nuevo los programas de educación compensatoria. Entre esta evidencia, por ejemplo, Battaglia y Lebedinski (2022) y Battaglia y Lebedinski (2015) encuentran efectos positivos tanto en el corto como en el medio plazo del Roma Teaching Assistant Program en Serbia.

Otra fuente de desigualdad educativa es la existencia de necesidades educativas especiales. Según el Ministerio de Educación (2023) más de 800.000 alumnos recibieron en el curso 2021-2022 atención educativa diferente a la ordinaria por presentar necesidades educativas especiales asociadas a discapacidad o trastorno grave (en torno a 245.000) u otras necesidades de apoyo (aproximadamente 555.000 alumnos)[3]. La mayor parte de estos alumnos (84%) está escolarizado en enseñanzas ordinarias, es decir, en las mismas aulas que el resto de los niños. No existen aún estadísticas en España sobre el nivel educativo de las personas con discapacidad en relación con el resto de la población. Sí sabemos que, en todos los países del mundo, las personas con discapacidad tienen niveles educativos inferiores que las personas sin discapacidad: por término medio, las primeras tienen menos probabilidades de asistir alguna vez a la escuela, tienen más probabilidades de estar sin escolarizar, tienen menos probabilidades de completar la Enseñanza Primaria o Secundaria, tienen menos años de escolarización y tienen menos probabilidades de poseer conocimientos básicos de Lectura y Escritura (Naciones Unidas, 2018). Así, entre las medidas que recomienda la UNESCO (2021) podemos destacar la necesidad de evaluar la educación que reciben estos alumnos, mejora de la formación de docentes o la inversión en tecnología como herramienta para reducir barreras.

En los últimos años, España ha experimentado un notable avance en sus resultados educativos, reduciendo las disparidades existentes. No obstante, persisten desigualdades significativas, especialmente relacionadas tanto con el nivel socioeconómico familiar como con las dimensiones mencionadas en esta sección. Es crucial identificar con precisión las causas de estas disparidades para diseñar políticas efectivas que las puedan mitigar. Además, se hace imprescindible llevar a cabo evaluaciones meticulosas de las intervenciones educativas implementadas. Solo comprendiendo a fondo estas desigualdades y sus causas, podremos

[3] Para más información sobre los conceptos y precisiones asociadas consultar la Metodología de esta estadística.

desarrollar un sistema educativo realmente equitativo, ofreciendo igualdad de oportunidades para todos los estudiantes, independientemente de su origen o situación económica.

## BIBLIOGRAFÍA


Alan, S., T. Boneva y S. Ertac (2019). Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit. *The Quarterly Journal of Economics, 134*(3), 1121-1162. https://doi.org/10.1093/qje/qjz006

Alet, E. (2014). Repetition: Medicine for a Short-Run Remission. *Annales d'Économie et de Statistique, 111-112*, 227-250. https://doi.org/10.2307/23646332.

Allen, C., Chen, Q., Willson, V. y Hughes, J. (2009). Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multilevel Analysis. *Educational Evaluation and Policy Analysis, 31*(4), 480-499. https://doi.org/10.3102/0162373709352239.

Attanasio, O. (2015). The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement, and Polices. *Journal of the European Economic Association, 13*(6), 949-97. https://doi.org/10.1111/jeea.12159.

Barro, R. y Lee, J. (2001). International data on educational attainment: updates and implications. *Oxford Economic Papers, 53*, 541-63. https://doi.org/10.1093/oep/53.3.541.

Battaglia, M. y Lebedinski, L. (2015). Equal Access to Education: An Evaluation of the Roma Teaching Assistant Program in Serbia. *World Development 76*, 62-81. https://doi.org/10.1016/j.worlddev.2015.06.009

Battaglia, M. y Lebedinski, L. (2022). With a little help from my friends: Medium-Term effects of a remedial education program targeting Roma minority. *Economics of Education Review, 86,* 102196. https://doi.org/10.1016/j.econedurev.2021.102196.

Battaglia, M. e Hidalgo-Hidalgo, M. (2023). *Better Early than Late? Primary versus Secondary School Remedial Interventions*, Mimeo, Universidad Pablo de Olavide.

Battistin, E. y Meroni, E. C. (2016). Should We increase Instruction Time in Low Achieving Schools? Evidence from Southern Italy. *Economics of Education Review, 55,* 39-56. https://doi.org/10.1016/j.econedurev.2016.08.003

Bingley, P., Cappellari, L. y Tatsiramos, K. (2021). Family, Community and Long-Term Socio-Economic Inequality: Evidence From Siblings and Youth Peers. *The Economic Journal, 131*(636), 1515-1554. https://doi.org/10.1093/ej/ueaa121

Björklund, A. y Salvanes, K. (2011). Education and Family Background: Mechanisms and Policie's, en Hanushek, E., Machin, S. y Woessmann, L. (eds.), *Handbook of the Economics of Education, Vol. 3*. Elsevier, 201-247. https://doi.org/10.1016/B978-0-444-53429-3.00003-X.

Blanden, J., Doepke, M. y Stuhler, J. (2023). Educational Inequality, en Hanushek, E., Machin, S. y Woessmann, L. (eds.), *Handbook of the Economics of Education, Vol. 6.* Elsevier, 405-497. https://doi.org/10.1016/bs.hesedu.2022.11.003.

Breen, R. (2022). *The stubborn persistence of educational inequality.* IFS Deaton Review of Inequalities, Institute for Fiscal Studies.

Breen, R. y Müller, E. (eds.) (2020*). Education and Intergenerational Social Mobility in Europe and the United States*. Redwood City, CA: Stanford University Press.

Carcillo, S., Fernández, R., Königs, S. y Minea, A. (2015). NEET Youth in the Aftermath of the Crisis: Challenges and Policies. OECD Social, *Employment and Migration Working Papers*, No. 164, OECD Publishing, Paris. https://doi.org/10.1787/5js6363503f6-en

Castelló, A. y Doménech, R. (2002). Human Capital Inequality and Economic Growth: Some New Evidence. *Economic Journal, 112*, C187-200. https://doi.org/10.1111/1468-0297.00024.

Checchi, D. (2001). *Education, inequality and income inequality*. LSE STICERD Research paper DARP 52, LSE, London.

Cherry, G. y Vignoles, A. (2016). *What is the economic value of literacy and numeracy? Basic skills in literacy and numeracy are essential for success in the labor market*. IZA World of Labor, 229. https://doi.org/10.15185/izawol.229.v2.

Clark, G. (2021). *For Whom the Bell Curve Tolls: A Lineage of 400,000 Individuals 1750-2020 shows Genetics Determines most Social Outcomes*. Unpublished Manuscript, UC Davis.

Cockx, B., Picchio, M. y Baert, S. (2019). Modeling the effects of grade retention in high school. *Journal of Applied Econometrics, 34(3)*, 403-424. https://doi.org/10.1002/jae.2670

Comunicación de la Comisión Europea (2011). *Un marco europeo de estrategias nacionales de inclusión de los gitanos hasta 2020*. Comisión Europea, Bruselas.

Cornelissen, T., Dustmann, C., Raute, A. y Schönberg, U. (2018). Who Benefits From Universal Child Care? Estimating Marginal Returns To Early Child Care Attendance. *Journal of Political Economy, 126*(6), 2356-2409. https://doi.org/10.1086/699979

Cunha, F. y Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review, 97*(2), 31-47 DOI: 10.1257/aer.97.2.31

Currie, J. y Almond, D. (2011). Human Capital Development Before Age Five, en Card, D. y Ashenfelter, O. (eds.), *Handbook of Labor Economics, Vol. 4*, Elsevier, 1315-1486.

Elango, S., García, J.L., Heckman, J.J. y Hojman, A. (2015). *Early Childhood Education*. NBER Working Paper 21766.

Eurostat (2023). Living conditions in Europe - poverty and social exclusion. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Living_conditions_in_Europe_poverty_and_social_exclusion&oldid=584082#Poverty_and_social_exclusion

Ferreira, F. y Gignoux, J. (2014). The Measurement of Educational Inequality: Achievement and Opportunity. *The World Bank Economic Review, 28*, 210-246.

Fleurbaey, M. (2008). *Fairness, Responsibility and Welfare*. Oxford University Press.

Fundación Secretariado Gitano (2023). *La situación educativa del alumnado gitano en España. Informe ejecutivo.* Madrid, Ministerio de Derechos Sociales y Agenda 2030.

García-Pérez, J. y M. Hidalgo-Hidalgo (2017). No Student Left Behind? Evidence from the Programme for School Guidance in Spain. *Economics of Education Review 60*, 97- 111. https://doi.org/10.1016/j.econedurev.2017.08.006

García-Pérez, I., Hidalgo-Hidalgo, M. y Robles-Zurita, A. (2014). Does grade retention affect students' achievement? Some evidence from Spain. *Applied Economics, 46,* 12, 1373-1392. https://doi.org/10.1080/00036846.2013.872761.

Guryan, J., Ludwig, J., Bhatt, M. P., Cook, P. J., Davis, J. M. V., Dodge, K., Farkas, G., Fryer, Roland G., J., Mayer, S., Pollack, H., Steinberg, L. y Stoddard, G. (2023). Not Too Late: Improving Academic Outcomes among Adolescents. *American Economic Review, 113*(3), 738-65. DOI: 10.1257/aer.20210434

Hanushek, E., Schwerdt, G., Wiederhold, S. y Woessmann, L. (2015). Returns to Skills around the World: Evidence from PIAAC. *European Economic Review, 73*, 103-130. https://doi.org/10.1016/j.euroecorev.2014.10.006

Hanushek, E., y Woessmann, L. (2010). The economics of international differences in educational achievement. NBER Working Paper, 15949.

Holmlund,H., Lindahl, M. y Plug, E. (2011). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature, 49*, 3, 615-651. https://doi.org/ 10.1257/jel.49.3.615

Jacob, B. y Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics, 1*, 33-58. DOI: 10.1257/app.1.3.33.

Joensen, J. y Nielsen, H. (2009). Is there a Causal Effect of High School Math on Labor Market Outcomes? *Journal of Human Resources, 44*(1), 171-198. https://www.jstor.org/stable/20648891

Lavy, V., A. Kott y G. Rachkovski (2022). Does Remedial Education at Late Childhood pay off after all? Long-run Consequences for University Schooling, Labor Market Outcomes and Inter-generational Mobility. *Journal of Labor Economics, 40* (1), 239-282. DOI: 10.1086/713742

Lochner, L. (2011). Nonproduction Benefits of Education: Crime, Health, and Good Citizenship en Hanushek, E., Machin, S. y Woessmann, L. (eds.), *Handbook of the Economics of Education, Vol. 4*. Elsevier, 183-282. https://doi.org/10.1016/B978-0-444-53444-6.00002-X.

Manacorda, M. (2012). The cost of grade retention. *Review of Economics and Statistics 94,* 596-606. https://www.jstor.org/stable/23262090

Marrero, G., Palomino, J. y Sicilia, G. (2023). Inequality of Opportunity in Educational Achievement in Western Europe: contributors and channels. ECINEQ WP 612.

Meschi, E. y Scervini, F. (2014). Expansion of schooling and educational inequality in Europe: the educational Kuznets curve revisited. *Oxford Economic Papers, 66*, 660-680.

Ministerio de Educación y Formación Profesional (2022). Estadística de las enseñanzas no universitarias. Alumnado con necesidad específica de apoyo educativo. Curso 2020-2021.

Naciones Unidas (2018). Education and Disability: Analysis of Data from 49 Countries. Information Paper N. 49.

OECD (2019). *PISA 2018 Results (Volume II): Where All Students Can Succeed.* PISA, OECD Publishing, París, https://doi.org/10.1787/b5fd1b8f-en.

OECD (2020). *PISA 2018 Results (Volume V): Effective Policies, Successful Schools.* PISA, OECD Publishing, Paris, https://doi.org/10.1787/ca768d40-en

OECD (2022). *Education at a Glance 2022: Educational attainment and labor-force status*, https://data.oecd.org/eduatt/adult-education-level.htm

Piketty, T. (2000). Theories of Persistent Inequality and Intergenerational Mobility, en Atkinson, A. y Bourguignon, F. (eds.), *Handbook of Income Distribution, Vol. 1,* Elsevier, 429-476.

Ram, R. (1984). Population Increase, Economic Growth, Educational Inequality, and Income Distribution: Some Recent Evidence. *Journal of Development Economics, 14*, 419-428.

Ram, R. (1990). Educational Expansion and Schooling Inequality: International Evidence and Some Implications. *Review of Economics and Statistics, 72*(2), 266-274.

Resnjanskij, S., Ruhose, J., Wiederhold, S., Woessmann, L. y Wedel, K. (2023). Can Mentoring Alleviate Family Disadvantage in Adolescence? A Field Experiment to Improve Labor-Market Prospects. *Journal of Political Economy*, en prensa.

Roemer, J. (1998). *Equality of Opportunity*. Cambridge, MA: Harvard University Press.

Rose, H. y Betts, J. (2004). The Effect of High School Courses on Earnings. *The Review of Economics and Statistics, 86*(2):497-513.

Ruiz-Valenzuela, J. (2020). Intergenerational effects of employment protection reforms. *Labour Economics, 62*, 101774.

UNESCO (2021). *Educación y aprendizaje inclusivos para personas con discapacidad*. Instituto Internacional de Planeamiento de la Educación de la UNESCO.

# Capítulo 5c
# BRECHAS DE GÉNERO EN EDUCACIÓN

Claudia Hupkau
Departamento de Economía
CUNEF Universidad
Centre for Economic Performance
London School of Economics
ORCID: https://orcid.org/0000-0002-7545-3835

Jenifer Ruiz-Valenzuela
Departament d'Economia
Universitat de Barcelona
Institut d'Economia de Barcelona (IEB)
Centre for Economic Performance
London School of Economics
ORCID: https://orcid.org/0000-0002-7238-2074

**Resumen**

Los datos internacionales muestran obstinadamente que los chicos tienen un menor desempeño educativo y que las mujeres siguen estando infrarrepresentadas en las carreras universitarias y profesionales relacionadas con las disciplinas STEM (Ciencia, Tecnología, Ingeniería y Matemáticas, por sus siglas en inglés). En este capítulo presentamos al lector diferentes motivos por los que dichas brechas deberían preocuparnos y trataremos de entender las razones que se esconden detrás de estas dos principales brechas de género. Los lectores que hayan ido avanzando progresivamente en la Lectura de este manual podrán intuir que las razones están estrechamente vinculadas con diferentes *inputs* de la función de producción educativa, introducida en los capítulos 3a, b, c y d, respectivamente. En el caso del menor desempeño de los chicos, analizamos el papel de las habilidades no cognitivas, de las familias y de los factores socioeconómicos, así como el impacto de la escuela y del profesorado, y de las condiciones del mercado laboral. En cuanto a la infrarrepresentación de las mujeres en STEM, examinamos el papel de la cultura, normas sociales y estereotipos de género, la ausencia de referentes femeninos entre los profesores de disciplinas STEM, y, por último, el rol de los contextos competitivos.

## 1. INTRODUCCIÓN

En la mayoría de las economías desarrolladas el desempeño educativo de las mujeres ha mejorado de forma sustancial en las últimas décadas. De hecho, el porcentaje de mujeres que se matricula en la Universidad es más alto que el de los hombres en muchos países. En España, por ejemplo, un 56% de los matriculados en la Universidad en el curso 2021-2022 eran mujeres. Sin embargo, el porcentaje de mujeres que decide matricularse en disciplinas del ámbito STEM (Ciencia, Tecnología, Ingeniería y Matemáticas, por sus siglas en inglés) sigue siendo bajo y ha permanecido relativamente estable durante los últimos años. De estas observaciones se pueden intuir dos de las principales brechas de género que más preocupan en el ámbito de la Economía de la Educación.

En primer lugar, en la mayoría de las economías avanzadas observamos un menor nivel de desempeño educativo de los chicos. Uno de los indicadores más usados para captar este fenómeno es el porcentaje de abandono escolar temprano por sexo. Tal y como se mencionó en el capítulo 1, el abandono escolar temprano se define como el porcentaje de individuos con edades comprendidas entre los 18 y 24 años que reúnen las siguientes características:

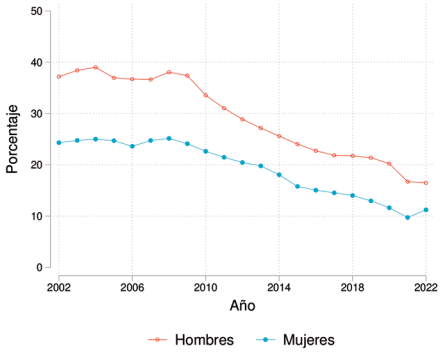
1) Su máximo nivel de educación es la primera etapa de la Educación Secundaria Obligatoria. En España, esto corresponde a la Educación Secundaria Obligatoria (ESO).

2) No están inscritos en ningún curso académico o formativo.

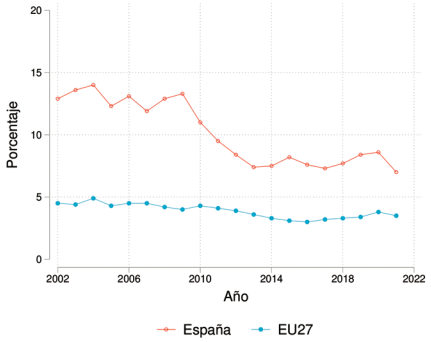
En el gráfico 1 (panel A) podemos ver cómo durante las dos últimas décadas, el porcentaje de abandono escolar temprano de los chicos en España siempre ha estado por encima del de las chicas. Este fenómeno, aunque en menor medida, también se da para la media de los países de la Unión Europea de los 27 (panel B). Los últimos datos para España muestran una brecha de género por encima de 5 puntos porcentuales, con un porcentaje de abandono escolar entre los chicos del 16,5%. Otros indicadores que capturan el desempeño educativo de los alumnos apuntan en la misma dirección. Por ejemplo, los chicos muestran una tasa de repetición de curso más alta que las chicas, y una fracción menor de los chicos termina la segunda etapa de Educación Secundaria en España. Así, con datos del curso 2020-2021, el porcentaje de chicas que termina la segunda etapa de la Educación Secundaria supera la de los chicos en casi 10 puntos porcentuales.

**Gráfico 1. Abandono escolar temprano y brecha de género.**

Panel A: Abandono temprano en ESP por sexo



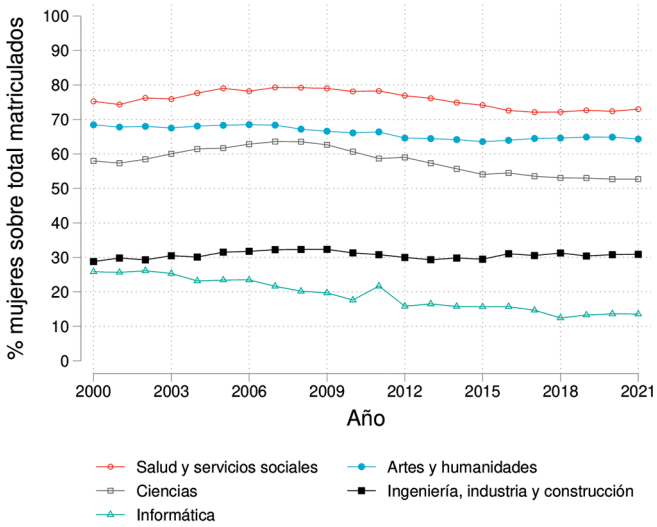

Panel B: Brecha de género en ESP y EU-27




Nota: Cálculos realizados a partir de la información en Eurostat (Early leavers from education and training by sex).

En segundo lugar, y atendiendo a los datos de matriculados en universidades españolas por ramas de enseñanza, vemos que dentro de las disciplinas STEM, el bajo porcentaje de mujeres matriculadas no viene dado por la rama de Ciencias sino por las ramas de Ingeniería, Industria y Construcción e Informática (gráfico 2). Además, este hecho no parece haber cambiado demasiado durante los últimos años, e incluso ha empeorado ligeramente en las ramas de Ciencias e Informática. Si hacemos un análisis más detallado de las ramas de estudio podemos entender de forma más precisa qué ámbitos de estudio concentran un menor porcentaje de mujeres en el curso académico 2021-2022. Los datos del Sistema Integrado de Información Universitaria de la Secretaría General de Universidades que se muestran en la tabla 1 revelan que, durante el curso 2021-2022, las mujeres estaban infrarrepresentadas en Informática e Ingenierías (15% y 26,6% de los matriculados, respectivamente), y dentro del total de Ciencias, también están infrarrepresentadas en Ciencias Físicas, Químicas y Geológicas (42,5%) y Matemáticas y Estadística (36,8%). Esta brecha no se reduce a la elección de carrera universitaria. Las mujeres también están infrarrepresentadas en los ciclos formativos de Formación Profesional en ámbitos STEM. Por ejemplo, según los datos recogidos por el Ministerio de Educación y Formación Profesional, en su publicación *Sistema estatal de indicadores de la educación 2022*, el 94% del alumnado de ciclos superiores de Electricidad y Electrónica está formado por hombres. Para Informática y Comunicaciones, este porcentaje es de casi el 88%. Al igual que la primera brecha, este fenómeno no se da únicamente en España, sino que es una característica compartida por la mayoría de las economías desarrolladas (Encinas-Martín y Cherian, 2023).

**Gráfico 2. Porcentaje de matriculados mujeres en universidades españolas, por ramas de enseñanza.**




Nota: Cálculos realizados a partir de la información del Sistema Integrado de Información Universitaria (SIIU). Secretaría General de Universidades. Estudiantes matriculados en Grado y ciclo.

En este capítulo trataremos de entender las razones que se esconden detrás de estas dos principales brechas de género. Los lectores que hayan ido avanzando progresivamente en la Lectura de este manual podrán intuir que las razones están estrechamente vinculadas con diferentes *inputs* de la función de producción educativa, introducida en los capítulos 3a, b, c y d, respectivamente.

***¿Por qué estas brechas de género en la educación deberían preocuparnos?***

En ambos casos, estas brechas de género deberían preocuparnos tanto por sus implicaciones negativas para el propio individuo como por sus efectos para la sociedad en general.

Empecemos con la primera de las brechas descritas, esto es, el menor nivel de desempeño educativo de los hombres. En *The Price We Pay: Economic and Social Consequences of Inadequate Education*, Belfield y Levin (2007) reúnen a un grupo de académicos expertos en Economía de la Educación, que, utilizando datos de Estados Unidos muestran como una peor o mala educación está asociada a una mayor probabilidad de sufrir desempleo y problemas de salud, obtener menores ingresos, o cometer actos criminales. Por ejemplo, en uno de los capítulos, Enrico Moretti documenta como los *high-school dropouts* (aquellos estudiantes

**Tabla 1. Porcentaje de mujeres matriculadas en universidades españolas, por ámbito de estudio. Curso académico 2021-2022.**

| Total mujeres matriculadas: 56,3 % | | | |
|---|---|---|---|
| **Total Educación** | **77,70%** | **Total Ciencias** | **49,30%** |
| Formación de docentes de enseñanza infantil | 91,30% | Ciencias de la vida | 60,70% |
| Formación de docentes de enseñanza primaria | 67,40% | Ciencias Físicas, Químicas, Geológicas | 42,50% |
| Otra Formación de personal docente y Ciencias de la educación | 82,70% | Matemáticas y Estadística | 36,80% |
| **Total Artes y humanidades** | **62,10%** | **Total Ingeniería, industria y construcción** | **30,30%** |
| Técnicas audiovisuales y medios de comunicación | 55,50% | Ingenierías | 26,60% |
| Artes | 72,10% | Arquitectura y construcción | 46,30% |
| Lenguas | 75,30% | **Total Agricultura, ganadería, silvicultura, pesca y veterinaria** | **53,90%** |
| Humanidades | 38,20% | Agricultura, ganadería y pesca | 31,00% |
| **Total Ciencias sociales, periodismo y documentación** | **63,20%** | Veterinaria | 76,90% |
| Psicología | 76,40% | **Total Salud y servicios sociales** | **72,60%** |
| Economía | 38,20% | Medicina | 70,20% |
| Otras Ciencias sociales y del comportamiento | 58,10% | Enfermería y atención a enfermos | 81,20% |
| Periodismo e información | 57,60% | Otras Ciencias de la Salud | 65,10% |
| **Total Negocios, administración y derecho** | **54,50%** | Trabajo social y orientación | 84,00% |
| Administración y gestión de empresas | 49,10% | **Total Servicios** | **40,80%** |
| Otra Educación comercial y empresarial | 58,90% | Deportes | 21,10% |
| Derecho | 59,20% | Turismo y Hostelería | 65,20% |
| **Total Informática** | **15,00%** | Otros Servicios | 26,50% |

Fuente: Cálculos realizados a partir de la información del Sistema Integrado de Información Universitaria (SIIU). Secretaría General de Universidades. Estudiantes matriculados en Grado y ciclo.

que no completan la segunda etapa de la Educación Secundaria, i.e., Bachillerato o Formación Profesional de Grado Medio en España), tienen el doble de probabilidad de acabar en la cárcel que aquellos individuos que sí acaban la Educación

Secundaria post-obligatoria. Además de estos costes individuales, Belfield y Levin (2007) explican que existen costes bien documentados para el contribuyente y la sociedad en general:

- Un menor nivel de ingresos a nivel individual ralentiza el crecimiento económico y los ingresos impositivos, y pone presión sobre las partidas económicas destinadas a sufragar las políticas asistenciales a familias de bajos ingresos.
- Peores niveles de salud incrementan el coste de la asistencia sanitaria pública. En uno de los capítulos, Cecilia Rouse estima que por cada individuo que acaba la Educación Secundaria post-obligatoria en Estados Unidos, el gobierno ahorraría unos 39.000 dólares debido a un menor uso de los programas públicos de salud.
- Una de las externalidades negativas de bajos niveles de educación, y que se da en mayor medida entre los hombres, es una mayor probabilidad de cometer crímenes. Enrico Moretti estima que un año adicional de escolarización haría disminuir los asesinatos y asaltos en casi un 30%, el robo de vehículos en un 20%, o los robos y hurtos en 6%. Así, calcula que si un 1% adicional de hombres con edades comprendidas entre los 20 y los 60 años se graduaran de la Educación Secundaria post-obligatoria, el gobierno de Estados Unidos se ahorraría más de un billón de dólares al año en costes relacionados con el sistema de justicia (jueces, fiscales, prisiones, etc.).

Estas consecuencias negativas de un bajo nivel de desempeño educativo de los chicos se ven agravadas por la creciente polarización del mercado de trabajo. Así, Autor (2010) explica que las oportunidades laborales están cada vez más concentradas en los extremos de la distribución de puestos de trabajos. Es decir, observamos una mayor demanda de puestos de trabajo para trabajadores altamente cualificados (y que pagan mayores salarios), una menor demanda de empleos para trabajadores con un nivel medio de cualificación, y una mayor demanda de puestos de trabajo de menor cualificación (y que normalmente pagan unos salarios muy bajos). Un menor nivel de desempeño educativo entre los chicos los empuja a encontrarse infrarepresentados entre los trabajadores altamente cualificados. Además de las consecuencias negativas en términos laborales o crimen, David Autor aporta evidencia sobre la asociación entre nivel educativo y estado civil: los hombres que no acaban la Educación Secundaria post-obligatoria tienen una menor probabilidad de contraer matrimonio.

Ejercicio 1. *Acabamos de describir las consecuencias negativas que tiene para el individuo un bajo nivel de educación, así como las externalidades negativas para la sociedad en general. Os invitamos a discutir en clase si dichas externalidades negativas son peores en el caso de que la infraeducación se dé mayoritariamente entre los hombres. Un buen artículo para apoyar la discusión es el escrito por David Autor en 2010* "US Labor Market Challenges over the Longer Term".

La segunda de las brechas descritas, esto es, la menor proporción de mujeres en disciplinas STEM, también acarrea efectos adversos para las mujeres en el mercado de trabajo. A nivel individual, los graduados en estas disciplinas en la educación terciaria (ya sean graduados universitarios o en ciclos de formación profesional) tienden a trabajar en industrias (o sectores económicos) y ocupaciones con unos salarios más altos. La evidencia existente nos dice que tanto el sector económico como la ocupación de los trabajadores es una de las razones de la persistente brecha en salarios entre hombres y mujeres (es decir las mujeres tienen en media menor salario que los hombres) (Blau y Kahn, 2017; Black et al., 2021). Además, debido a la creciente importancia de la inteligencia artificial o el análisis de datos para las empresas (*data analytics*), se espera que la demanda de graduados en disciplinas STEM en el futuro siga aumentando. Esto ejercería presiones al alza en los salarios de muchas disciplinas STEM, contribuyendo, si no cambian los patrones de matriculación en disciplinas STEM por género, a una mayor brecha de género en salarios.

Para el lector interesado en entender la relación entre la rama de estudios y sus retornos en el mercado laboral, el capítulo 4b nos ofrece una primera aproximación. Además, el capítulo de Altonji et al. (2016) en el *Handbook of the Economics of Education* recoge un buen resumen de cuáles son los problemas (y potenciales soluciones) a los que el investigador se enfrenta cuando quiere estimar efectos causales de retornos a diferentes ramas de estudio.

Además de estas ganancias individuales asociadas a ramas de estudio STEM, existe evidencia de que los trabajadores STEM son fundamentales para la innovación, uno de los factores principales para explicar el crecimiento de la productividad de una economía. Por ejemplo, con datos de Estados Unidos, Peri et al. (2015) utilizan la técnica de variables instrumentales explicada en el capítulo 2 de este libro de la siguiente manera. En la primera etapa, como instrumento para predecir el crecimiento de trabajadores extranjeros STEM, utilizan variación en el tiempo de la proporción de visados otorgados a diferentes grupos

de nacionalidades en diferentes ciudades, incluyendo efectos fijos de ciudad en sus estimaciones. En la segunda etapa, los autores obtienen estimaciones causales del efecto del crecimiento de trabajadores extranjeros STEM en los salarios y empleo de trabajadores nativos. Sus resultados muestran que un incremento de un punto porcentual en el crecimiento de trabajadores extranjeros STEM conlleva un incremento salarial de entre siete y ocho puntos porcentuales para trabajadores nativos con título universitario. Con estos resultados, los autores simulan el efecto de los trabajadores STEM en la productividad total agregada, encontrando que los influjos de trabajadores STEM extranjeros explican alrededor de un 30%-50% del crecimiento agregado de la productividad en Estados Unidos entre 1990 y 2010. Por lo tanto, en un contexto internacional en el que existe escasez de trabajadores STEM, la infrarrepresentación de mujeres en estas disciplinas puede estar reduciendo considerablemente la innovación y productividad agregada.

## 2. QUÉ FACTORES EXPLICAN UN MAYOR PORCENTAJE DE FRACASO ESCOLAR ENTRE LOS CHICOS

Para la mayoría de las economías avanzadas, los datos son testarudos y reflejan un menor desempeño educativo de los chicos. En esta sección queremos presentar algunos de los factores que contribuyen a explicar esta brecha de género. La lista de factores no es exhaustiva, sino que pretende describir algunos de los principales factores que se encuentran detrás de estas brechas. Además, para cada factor, no se ofrece una revisión completa de la literatura, sino que se ofrecen al lector algunas de las referencias más destacadas. En la mayoría de los casos, el objetivo es entender no sólo los factores, sino también hacer uso de lo visto en el capítulo 2 para ver cómo los distintos autores han intentado abordar problemas de endogeneidad.

Hemos agrupado las potenciales razones en cuatro grandes grupos. En primer lugar, hacemos un breve repaso de diferencias en habilidades no cognitivas entre chicos y chicas. Luego abordamos el rol que las familias y, de forma general, el entorno socioeconómico, juegan a la hora de explicar esta brecha de género (véase también el capítulo 3b sobre el impacto de las familias en la educación). En tercer lugar, repasamos qué nos dice la evidencia existente sobre el papel de la escuela y los potenciales sesgos del profesorado. Por último, damos algunas pinceladas sobre la relación existente entre las condiciones del mercado

laboral, el coste de oportunidad de seguir estudiando, y el desempeño educativo del alumnado.

### 2.1. Habilidades no cognitivas

En el capítulo 5d se ha descrito qué entendemos por habilidades no cognitivas, cómo las medimos, y cuáles son sus determinantes; así como su importancia para el bienestar de los individuos a lo largo de la vida. Uno de los aspectos que resaltan la importancia de estas habilidades no cognitivas es el papel que juegan a la hora de determinar los resultados académicos de los estudiantes. En este capítulo, lo que nos interesa abordar es si existen diferencias en habilidades no cognitivas entre chicos y chicas que puedan estar explicando diferencias de género en fracaso escolar.

La literatura en psicología lleva tiempo estudiando la existencia de diferencias de género en habilidades no cognitivas. Un buen ejemplo es el estudio de Else-Quest et al. (2006), que recoge los resultados de un meta-análisis que analiza diferencias de género en el temperamento de chicos y chicas (i.e., carácter, manera de ser o de reaccionar de las personas). Los autores de este meta-análisis utilizan datos de un total de 189 estudios, seleccionados meticulosamente (el proceso de selección está descrito de forma exhaustiva en su artículo). La conclusión de este meta-análisis es que los chicos tienen niveles menores de control inhibitorio así como de sensibilidad perceptiva, y esto es consistente con una mayor incidencia de exteriorización de problemas de conducta.

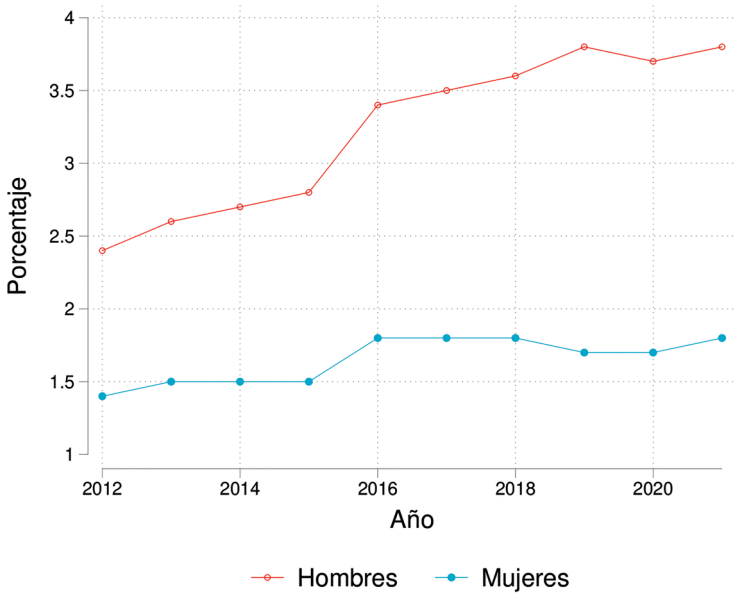
En economía, son varios los estudios que han asociado un menor nivel de habilidades no cognitivas con problemas de rendimiento académico. Uno de los primeros ejemplos en vincular diferencias de género en habilidades no cognitivas con un peor desempeño educativo de los chicos es el artículo del premio Nobel de Economía Gary S. Becker, escrito junto con William H.J. Hubbard y Kevin M. Murphy. En dicho artículo, Becker et al. (2010) intentan explicar por qué las mujeres han superado a los hombres en términos de matriculación universitaria analizando la distribución de costes y beneficios de ir a la Universidad. En general, los autores encuentran que durante las últimas cuatro décadas, los beneficios de ir a la Universidad, a pesar de haber incrementado más para las mujeres que los hombres, no pueden explicar la mayor graduación de las mujeres en la etapa terciaria. De entre los costes, los autores separan aquellos que llaman tradicionales (o monetarios) de los no tradicionales, aclarando que estos últimos

están determinados, especialmente, por diferencias en habilidades no cognitivas entre los individuos. Sobre estos últimos, los autores muestran dos diferencias de género importantes. En primer lugar, las habilidades no cognitivas medias de las mujeres son más altas que las de los hombres. En segundo lugar, las mujeres presentan una menor varianza en su distribución de habilidades no cognitivas. Estos dos hechos implican, de acuerdo con su modelo, que la elasticidad de oferta de las mujeres sea más elevada, para un mismo nivel de beneficios de ir a la Universidad. Así, menores costes de ir la Universidad expresados en términos de mayores (y menos dispares) habilidades no cognitivas entre las mujeres estarían explicando por qué las mujeres han sobrepasado a los hombres en la mayoría de países desarrollados en cuanto a participación en la Educación Terciaria.

Datos descriptivos sobre habilidades no cognitivas para España apoyarían estas asociaciones. Por ejemplo, con datos de matriculados en enseñanzas de Régimen General no universitarias, el gráfico 3 muestra que el porcentaje de alumnado con necesidades educativas especiales (aquel que requiere de determinados apoyos y atenciones educativas específicas derivadas de discapacidad o trastornos graves de conducta) es mayor entre los chicos que las chicas. Además, esta brecha de género se ha ampliado durante la última década.

Existe evidencia también de que los chicos dedican menos horas a hacer deberes o tareas escolares fuera del horario lectivo. Por ejemplo, con microdatos de la Comunidad de Madrid, Montalbán y Ruiz-Valenzuela (2022) muestran que los chicos emplean un menor número de horas en hacer deberes, y que estas diferencias se amplían con el tiempo. En cuarto de la ESO, las chicas hacen entre 1,5 y 2,5 horas de deberes más a la semana que los chicos (la brecha de género se amplía a medida que aumenta el nivel educativo más alto de los progenitores). Estos números están en sintonía con los resultados de Gershenson y Holt (2015) con datos de la encuesta de uso del tiempo de Estados Unidos, que encuentran que las chicas dedican casi 1,7 horas más a deberes que los chicos. Después de tener en cuenta diferencias de género en el uso del tiempo después de la escuela, el grado de involucramiento de los padres, las expectativas educativas, la elección de cursos, el rendimiento académico previo o la capacidad cognitiva, la brecha de género sigue siendo de una hora semanal extra dedicada a tareas escolares por las chicas. Estos mismos autores explican como el tiempo dedicado a hacer deberes está asociado a una importante habilidad no cognitiva: la voluntad de aprendizaje.

**Gráfico 3. Alumnado con necesidades educativas especiales.**




Nota: Los datos muestran el porcentaje de alumnos y alumnas con necesidades educativas especiales. Los porcentajes se refieren al porcentaje de alumnas (alumnos) con necesidades educativas especiales sobre el total de alumnas (alumnos) matriculados en enseñanzas de Régimen General no universitarias. El alumnado con necesidades educativas especiales se define como el alumnado valorado como tal por los equipos o servicios de orientación educativa que requiere por un periodo de su escolarización o a lo largo de toda ella, determinados apoyos y atenciones educativas específicas derivadas de discapacidad o trastornos graves de conducta (véase Estadística de las enseñanzas no universitarias. Alumnado con necesidad específica de apoyo educativo. Metodología).
Fuente: Estadística de las Enseñanzas no universitarias. Subdirección General de Estadística y Estudios del Ministerio de Educación y Formación Profesional.

En suma, diferencias en habilidades no cognitivas entre chicos y chicas, como un mayor porcentaje de problemas de conducta entre los chicos o una menor voluntad de aprendizaje, estarían explicando, en parte, el menor desempeño educativo de los chicos.

Ejercicio 2. *En este capítulo hemos visto que las mujeres han sobrepasado a los hombres en cuanto a la proporción que accede y finaliza la educación universitaria. También hemos visto que los chicos tienen una menor probabilidad de terminar la segunda etapa de la Educación Secundaria Obligatoria. Discute cómo las po-*

*tenciales diferencias en habilidades no cognitivas entre hombres y mujeres podrían explicar esta diferencia. Puedes ayudarte para ello con la Lectura de las referencias citadas en esta sección.*

### 2.2. Familias y factores socioeconómicos

En el capítulo 3b hemos visto que las familias juegan un papel importante en el desempeño educativo de sus hijos. Variables que se deciden en el seno familiar, como los recursos económicos disponibles para la educación u otras actividades extraescolares, el tiempo dedicado a los hijos, o los diferentes estilos parentales, se consideran *inputs* en la función de producción educativa. ¿Podrían estos diferentes *inputs* tener un efecto diferenciado en el desempeño educativo de chicos y chicas? Por ejemplo, ¿afecta de forma diferente a chicos y chicas haber crecido en un peor entorno socioeconómico? ¿Se escogen de manera diferente los *inputs* de la función de producción educativa que deciden los padres, en función del sexo del descendiente?

Respecto a la primera de estas preguntas, una hipótesis que ha sido contrastada con datos de Florida en Estados Unidos por David Autor, David Figlio, Krzysztof Karbownik, Jeffrey Roth y Melanie Wasserman, es que los chicos sufren peores consecuencias académicas cuando crecen en entornos familiares desfavorecidos. Estimar este efecto diferencial no es fácil. Como bien explican los autores, hay dos desafíos importantes. El primero es que el entorno familiar está muy correlacionado con otros *inputs* de la función de producción educativa, como por ejemplo el nivel educativo de los progenitores. Por ejemplo, padres con un mayor nivel educativo normalmente residen en barrios más ricos y con mejores escuelas. Así, es difícil separar el impacto de un entorno económico desfavorecido del impacto de escuelas de peor calidad o peores barrios. El segundo desafío viene dado por la cantidad de datos que se necesitan para poder implementar y justificar una estrategia de estimación causal creíble. El trabajo de Autor y coautores supera estos dos desafíos con muy buena nota. El primer desafío lo superan comparando diferencias en resultados académicos entre hermanos y hermanas. El supuesto que hacen es que hermanos que han sido cuidados por la misma madre están expuestos al mismo entorno familiar. El segundo supuesto que necesitan para poder decir que los efectos encontrados son de naturaleza causal es que no haya diferencias sistemáticas entre hermanos y hermanas por nivel socioeconómico ya al nacer. Este segundo supuesto lo pueden verificar gra-

cias a que suplen con creces el segundo desafío, al disponer de una base de datos con medidas muy precisas de salud de las madres y de los niños al nacer. Con esos datos, demuestran que, al menos en su muestra, no existen diferencias significativas entre hermanos y hermanas (o en el tiempo para una misma madre). A pesar de disponer de muy buenos datos y poder comparar directamente el rendimiento académico de hermanos y hermanas, el efecto que encuentren los autores podría estar determinado por un efecto de género diferencial del barrio donde residen los estudiantes o las escuelas que atienden (más adelante veremos que dichos efectos también existen). Es decir, escuelas de peor calidad, o vivir en peores barrios, podrían afectar de manera desigual al rendimiento académico de chicos y chicas. Para poder obtener estimaciones limpias de esta posibilidad, que capturen únicamente el efecto de crecer en un entorno familiar desfavorecido (y no de crecer en barrios desfavorecidos o atender escuelas de peor calidad), los autores introducen en sus estimaciones variables que miden la calidad de las escuelas que los hermanos y hermanas atienden, e incluyen además efectos fijos de barrio. El ejercicio final indica que, en relación a sus hermanas, los chicos nacidos en familias de entornos desfavorecidos tienen peor rendimiento académico, mayores problemas disciplinarios y un mayor nivel de abandono escolar temprano.

Existen además efectos diferenciales de género de crecer en barrios desfavorecidos. Esto lo muestran Chetty y Hendren (2018), utilizando una estrategia algo similar, pues también se basa en comparar diferentes variables de resultados entre hermanos y hermanos. Estos autores utilizan datos de más de cinco millones de familias en Estados Unidos que se mudan de condado durante los años que se observan en la muestra. Así, pueden comparar los resultados de hermanos y hermanas que tenían diferentes edades al mudarse, y por lo tanto, estuvieron expuestos a diferentes barrios (antes y después de la mudanza) durante un número distinto de años. Los resultados muestran que, para niños que crecieron en el percentil 25 de la distribución de ingresos, crecer (desde el nacimiento) en un condado una desviación estándar mejor aumenta sus ingresos en aproximadamente un 10%. Los niños de familias de bajos ingresos tienen más probabilidades de tener éxito en condados que tienen una menor concentración de pobreza, una menor desigualdad de ingresos, mejores escuelas, una mayor proporción de familias con dos padres y tasas más bajas de delincuencia. Y algo de importancia para explicar las brechas de género en fracaso escolar: los autores muestran que los resultados de los hermanos varían

más entre áreas (son más susceptibles al efecto barrio) que los resultados de sus hermanas y que los niños tienen resultados especialmente pobres en áreas altamente segregadas.

Algo similar sucede con las habilidades no cognitivas, que recordemos tienen la capacidad de influenciar los resultados académicos de los estudiantes, y de las cuales la literatura nos dice que en parte podrían venir dadas por diferencias biológicas. El artículo de Marianne Bertrand y Jessica Pan ofrece evidencia de que, además, crecer en entornos desfavorecidos, en su caso medido por familias que llaman *rotas* (de madres solteras o madres adolescentes) tiene un efecto más pernicioso para los chicos que para las chicas (Bertrand y Pan, 2013).

Las conclusiones de estos tres importantes artículos nos ayudan a responder la primera de las preguntas expuesta al inicio de esta subsección: ¿Podrían diferentes factores socioeconómicos tener un efecto diferenciado en el desempeño educativo de chicos y chicas? La evidencia empírica nos dice que sí: crecer en un entorno familiar o barrio desfavorecido tiene un efecto más perjudicial para los chicos que para las chicas, y por lo tanto, contribuye a explicar la brecha de género en fracaso escolar.

Además de estos efectos del entorno socioeconómico, las familias escogen el nivel de inversión en educación (ya sea mediante recursos monetarios o de tiempo) para cada uno de sus hijos e hijas. La literatura que analiza el impacto de las inversiones parentales en educación es extensa, y no es nuestro objetivo en este capítulo hacer una revisión exhaustiva de esta. Pero sí nos gustaría hacer hincapié en algunos resultados sobre diferencias en el tiempo dedicado por los progenitores a actividades relacionadas con la educación de sus hijos y que emergen desde edades muy tempranas. Estas diferencias podrían estar explicando parte de la brecha de género existente en fracaso escolar. La existencia de datos detallados de encuestas en Canadá, Reino Unido y Estados Unidos han permitido a Michael Baker y Kevin Milligan contrastar algunos de los resultados de la literatura previa, que había llegado a la conclusión de que los padres solían invertir más tiempo en actividades relacionadas con la crianza cuando los descendientes eran niños. Con datos de edades muy tempranas (i.e., desde alrededor de los nueve meses hasta los 5 años), Baker y Milligan (2016) muestran que ese mayor desempeño de los padres con sus hijos aparece con la edad, y no se da a edades más tempranas. De hecho, en los tres países, lo que muestran los autores es que tanto padres como madres invierten más

tiempo en actividades que categorizan como de enseñanza (i.e., leer, cantar, aprender a usar números y letras) con sus hijas, y que estas diferencias se observan tan pronto como a los nueve meses. Con datos de la encuesta de Uso del Tiempo de Estados Unidos, Bertrand y Pan (2013) encuentran que esto se da especialmente en familias de madres solteras.

Así, tal y como nos planteábamos al inicio de esta sección: ¿Se escoge la inversión en tiempo dedicado a tareas de crianza educativa de manera diferente dependiendo del sexo del descendiente? El resumen es que existe cierta evidencia de que la respuesta es afirmativa, al menos en países desarrollados, y que esto podría estar explicando un menor desempeño educativo de los niños en comparación con las niñas.

Ejercicio 3. *Utiliza la evidencia en Bertrand y Pan (2013) y Baker y Milligan (2016) para discutir el rol que las preferencias parentales (por ejemplo, si las madres o padres prefieren tener hijas en lugar de hijos) pueden jugar a la hora de determinar las inversiones de recursos (i.e tiempo y/o dinero) que hacen en la educación de sus hijos e hijas. ¿Qué otros motivos mencionan estos autores para explicar las diferentes inversiones en educación que realizan los progenitores dependiendo del género de los hijos?*

### 2.3. Escuela y profesorado

En el capítulo 3 hemos visto que no son únicamente las familias las que con sus decisiones afectan al desempeño educativo de sus descendientes. Varios de los *inputs* de la función de producción educativa, que varían a nivel de escuela, pueden explicar importantes diferencias de rendimiento entre el alumnado (ver, por ejemplo, diferencias en tamaño de clase en el capítulo 3a; o diferencias relativas a la calidad del profesorado en el capítulo 3d). En este capítulo vamos a centrarnos en un par de aspectos relacionados con la escuela y el profesorado, y que podrían explicar un peor desempeño educativo por parte de los chicos: potenciales sesgos de género del profesorado, y una infrarrepresentación de profesorado masculino en distintas etapas de la Educación Obligatoria.

Empecemos con el rol de los potenciales sesgos que podrían tener los profesores con respecto a los chicos. En la sección 3.3 veremos la existencia de sesgos implícitos (y explícitos) de los profesores, que pueden estar condicionando de forma negativa el resultado de las chicas en asignaturas relacionadas con las Ciencias y las Matemáticas. Sin embargo, existe también evidencia de que los

profesores tienen sesgos que afectan de forma negativa a los chicos. Por ejemplo, la economista Camille Terrier utiliza datos de alumnado en Francia que contienen una característica esencial (Terrier, 2020): poseen información tanto de las notas puestas en las asignaturas de Francés y Matemáticas por los profesores (notas no-ciegas, o *non-blind scores*) como de las notas obtenidas por los alumnos en pruebas de las mismas asignaturas puntuadas por un corrector externo que no conoce al alumno (notas ciegas, o *blind scores*). Algo importante que nos aclara Terrier, es que ambas pruebas (externas y no externas), están diseñadas para medir las mismas habilidades del alumnado que se examina. Con estos datos, Terrier construye una medida del grado de sesgo de género que tienen los profesores, definida como la brecha promedio entre las puntuaciones de las chicas de pruebas no ciegas y ciegas, menos esta misma brecha para los chicos. Es decir, si la diferencia entre la puntuación dada por un profesor y la otorgada por el corrector externo es mayor para las chicas que para los chicos, ese profesor estaría poniendo notas más altas a las chicas y aparecería clasificado en los datos exhibiendo una brecha de género en favor de las chicas. Otro elemento crucial para su estrategia empírica es que los profesores se asignan de forma cuasi-aleatoria. Agregando los datos a nivel de clase, la variación que Terrier explota viene de la comparación entre el progreso de las chicas (relativo al de los chicos) en clases con profesores más sesgados, y el progreso que hacen las chicas (siempre en comparación con los chicos) en clases con profesores menos sesgados. De los resultados de Terrier, cabe destacar que los sesgos del profesorado tienen un impacto negativo en el progreso durante la Secundaria para los chicos, tanto en Matemáticas como en Francés. Otros estudios recientes, con datos de diferentes regiones de España, confirman estos resultados. Por ejemplo, con datos del País Vasco, Gortazar et al. (2022) encuentran que los docentes evalúan peor a aquellos grupos de estudiantes que, en promedio, van más rezagados en la escuela, como es el caso de los chicos.

Pasemos ahora a analizar el potencial rol que puede jugar la infrarrepresentación de profesorado masculino en distintas etapas de la Educación Obligatoria en un peor desempeño educativo de los chicos. Por ejemplo, en España, en el curso 2019-2020, el porcentaje de profesores de sexo masculino es casi inexistente en la etapa de Educación Infantil (el 97,7% de los profesores de esta etapa son mujeres). Las mujeres son también predominantes entre el profesorado de los centros de Educación Primaria y en los centros que imparten Educación Primaria y Secundaria (82,1% y 72,4%, respectivamente). Incluso en los centros de

Secundaria y Formación Profesional, la presencia de profesorado femenino es de casi un 60%. Este fenómeno no es reciente (los datos eran similares hace una década), pero la tendencia de los últimos 10 años refuerza estos patrones (según los datos del informe *Sistema Estatal de Indicadores de la Educación 2022*).

Como sucede para las chicas con la falta de referentes en STEM (ver sección 3.4), en las etapas de Infantil y Primaria hay una ausencia casi total de profesores de género masculino. Varios estudios en la literatura académica han proporcionado evidencia de que tener un profesor de tu misma identidad puede afectar de manera positiva el rendimiento académico del alumnado. Así, una menor proporción de profesores de género masculino en las etapas iniciales de la educación podría estar afectando de manera negativa el desempeño educativo de los chicos. Es el caso del estudio publicado por Thomas Dee en el Journal of Human Resources (Dee, 2007), que utiliza datos de una encuesta en Estados Unidos con una importante ventaja: la posibilidad de observar a un mismo estudiante, y su rendimiento, en distintas asignaturas y con profesores de distinto género. Es decir, Dee puede estudiar qué sucede con el rendimiento de un mismo estudiante cuando se le asigna a un profesor de sexo masculino y cuando se le asigna a una profesora (*within-student comparisons*). Sus resultados indican que la asignación de un profesor del mismo género mejora significativamente el rendimiento tanto de las niñas como de los niños, así como la percepción del maestro sobre el desempeño del estudiante y el compromiso del estudiante con la materia del maestro. Este resultado es relevante para explicar potenciales diferencias de rendimiento –sobre todo en Lectura– entre chicos y chicas, dado que, como hemos visto, la mayoría de los profesores, desde Infantil hasta la Secundaria, son mujeres.

Otro ejemplo muy reciente es el que estudian Ursina Schaede y Ville Mankki (Schaede y Mankki, 2022). Para analizar esta cuestión, Schaede y Mankki utilizan una reforma finlandesa en la que se suprimía la cuota obligatoria de un 40% de plazas reservadas en la Universidad para aspirantes a profesores de Primaria de sexo masculino. Esta reforma les permite estudiar la pregunta en un contexto de diferencias en diferencias (DiD, por *differences in differences* en inglés, ver capítulo 2) usando datos de antes y después de la reforma. Pero estos autores van algo más allá y combinan su contexto DiD con el método de variables instrumentales (DiD-IV). Es decir, estos autores encuentran una variable instrumental (una fuente de variación exógena) para la variable dependiente de interés, que en su caso mide la composición a nivel de municipio del género de los profesores de Primaria.

El instrumento que utilizan viene dado por la interacción entre el antes y después de la reforma con información sobre las jubilaciones obligatorias de profesores que ocurren cada año a nivel municipal cuando los profesores cumplen 60 años.

Lo primero que encuentran estos autores es que la eliminación de la cuota obligatoria del 40% de profesores hombres efectivamente redujo la proporción de hombres admitidos a aspirantes a profesores de Primaria desde un 40% a un 20%. En la segunda etapa examinan efectos en variables que miden el rendimiento educativo de los estudiantes, y, en su caso, encuentran que aquellos estudiantes expuestos a una mayor proporción de profesores de sexo masculino alcanzan mayores niveles educativos y también les va mejor en el mundo laboral. Estos resultados se dan tanto para los chicos como para las chicas, con lo que los autores concluyen que la existencia de la cuota en favor de tener al menos un 40% de candidatos admitidos hombres, mejora el proceso de selección de talento de los profesores.

Hay muchos más estudios en esta área, y veremos algunos en la sección 3.4, pero lo que nos sugieren estos dos estudios es que aumentando el número de profesores de sexo masculino en las etapas de Infantil y Primaria, se podría incrementar el desempeño educativo de los chicos sin que ello parezca perjudicar a las chicas.

### 2.4. Condiciones en el mercado laboral y coste de oportunidad de seguir estudiando

En la introducción a este capítulo vimos en el gráfico 1 la evolución del abandono escolar temprano por sexo. Una característica a resaltar del gráfico del panel A, es que la tasa de abandono temprano ha caído de forma sustancial a partir de la Gran Recesión de 2008. Estos datos sugieren que el mercado laboral, y en particular la disponibilidad de empleos de baja cualificación, pueden jugar un rol importante en la decisión de abandonar los estudios de forma temprana.

Para ver la conexión entre las condiciones laborales y las decisiones educativas, pensemos en un simple modelo de acumulación de capital humano (Becker, 1962) (el lector puede también repasar lo visto en el capítulo 4b). A los 16 años, cuando se acaba la Educación Obligatoria en España, un alumno tiene que decidir si quiere seguir estudiando o si quiere entrar en el mercado laboral. Es posible que él (o su familia) compare el flujo de ingresos futuros esperados si sigue estudiando un año más con el flujo de ingresos esperados si entra inmediatamente en

el mercado laboral. Si una alta tasa de desempleo es un indicador general de las condiciones del mercado laboral, esta puede indicar tanto bajos salarios iniciales para los que abandonan los estudios como una baja probabilidad de encontrar un empleo al abandonar la escuela. En ambos casos, esto se traduciría en una reducción del valor esperado de los flujos de ingresos si el estudiante decide entrar inmediatamente en el mercado laboral, reduciendo así el porcentaje de alumnado que abandona de forma temprana la escuela. Por otro lado, un entorno de crisis económica en el que un miembro de la familia pierde su empleo, podría obligar a otros miembros de la familia a entrar en el mercado laboral. Rees y Mocan (1997) muestran una correlación negativa entre el nivel de desempleo y el abandono temprano, indicando que el primer canal parece ser más relevante.

En el contexto de la crisis económica de 2008, el sector que más empleo destruyó en España fue el de la construcción, un sector que empleaba mayoritariamente a hombres, y con amplia disponibilidad de trabajos de baja cualificación. Esto posiblemente redujo el coste de oportunidad de seguir estudiando de manera desproporcional para los chicos, porque sus oportunidades laborales se vieron afectadas en mayor manera que las oportunidades laborales de las chicas. De hecho, como puede verse en el panel B del gráfico 1, la brecha de género en abandono escolar temprano se redujo significativamente tras la crisis de 2008-2009. Esto es precisamente lo que estudia Ainoa Aparicio-Fenoll de forma causal en su artículo publicado en el *Journal of Human Capital* en 2016. En ese artículo, Aparicio-Fenoll (2016) explica que el auge inmobiliario generó una variación sustancial en los retornos a la educación a lo largo del tiempo y en diferentes regiones, principalmente para los chicos. Además, argumenta que el auge inmobiliario fue exógeno a las diferencias preexistentes en los resultados educativos entre hombres y mujeres. Es por ello que puede utilizar una estrategia de estimación basada en variables instrumentales (ver capítulo 2). Más precisamente, en la primera etapa la autora instrumenta diferencias en retornos educativos entre hombres y mujeres con el nivel de actividad del sector de la construcción a nivel de provincia. En la segunda etapa, estima el efecto que estas variaciones en los retornos a la educación tienen sobre variables educativas. En particular, para jóvenes de 16 a 18 años, muestra que una disminución del 10 por ciento en la ratio de salarios de individuos educados con respecto a los no educados (es decir, aumentan de forma relativa los salarios en ocupaciones de baja cualificación), conduce a una disminución del 2 por ciento en la probabilidad de estar matriculado en la escuela.

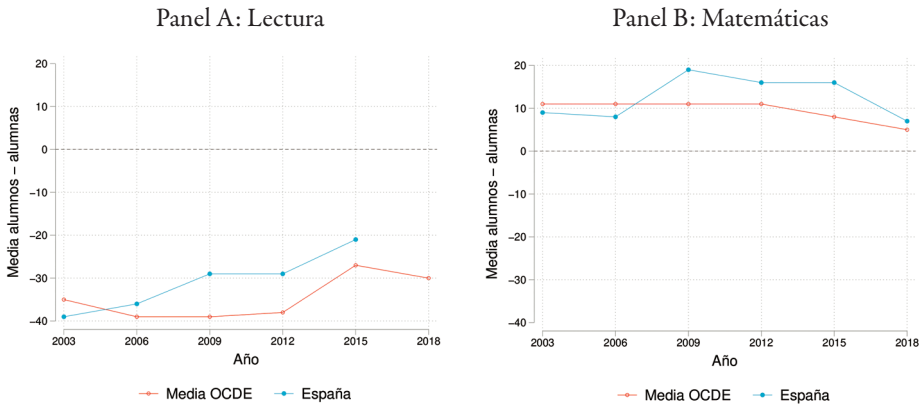
## 3. QUÉ FACTORES EXPLICAN UNA MENOR PARTICIPACIÓN DE LAS MUJERES EN STEM

Como hemos visto en la introducción de este capítulo, las mujeres tienen una probabilidad menor de matricularse en asignaturas relacionadas con las disciplinas STEM, como la Informática o Ingeniería (véase también la tabla 1), y siguen estando infrarrepresentadas en ocupaciones relacionadas con estas áreas. El consenso actual en la investigación sobre esta cuestión en el campo de la Economía de la Educación es que las diferencias en las capacidades de chicos y chicas en Matemáticas no pueden explicar las diferencias de género en las elecciones educativas y profesionales (Breda et al., 2018). Si las habilidades cognitivas de las chicas son muy similares a las de los chicos, ¿qué explica entonces este fenómeno? En esta sección analizaremos algunos de los factores que están detrás de la brecha de género existente a la hora de elegir especializaciones universitarias y profesionales intensivas en Matemáticas.

### 3.1. Ventaja comparativa

Históricamente, se ha argumentado que la falta de mujeres en el ámbito STEM, y más concretamente en las Ciencias intensivas en Matemáticas, se debe a un menor rendimiento de las chicas en esta materia. Sin embargo, en la mayoría de los países desarrollados, esta diferencia actualmente es prácticamente nula. El gráfico 4 muestra la evolución de la diferencia en la puntuación media entre chicos y chicas en la prueba de Lectura y Matemáticas, del Programa para la Evaluación Internacional de Estudiantes (PISA, por sus siglas en inglés) para España y la media de países de la OCDE participantes. Los datos muestran que, a los 15 años, chicos y chicas tienen, en promedio, un rendimiento muy similar en Matemáticas. Además, esta brecha ha ido disminuyendo en la última década. Los últimos datos disponibles para 2018 muestran una brecha en favor de los chicos de tan sólo 5 puntos para la media de la OCDE. Sin embargo, las niñas tienen una ventaja importante sobre los chicos en Lectura, de unos 30 puntos para el conjunto de países de la OCDE que participan en esta prueba (OECD, 2019). Además, cuando las chicas rinden mejor también en Matemáticas, como es el caso del Reino Unido (Cavaglia et al., 2021), la brecha de género en favor de las mujeres suele ser mayor en Lengua.

**Gráfico 4. Diferencia en la puntuación media entre chicas y chicos.**




Nota: Los datos muestran la diferencia en el examen PISA entre alumnos y alumnas para España y la media de los países OCDE. La diferencia de género en rendimiento se calcula para la asignatura de Matemáticas (panel A) y la asignatura de Lengua (panel B).
Fuente: Los gráficos provienen de Montalbán (2021), estadísticas elaboradas a partir de datos obtenidos de PISA Database, OECD.

En un estudio reciente en la revista *Proceedings of the National Academy of Sciences (PNAS)*, los economistas Thomas Breda y Clotilde Napp muestran que una explicación potencial para la diferencia de género en la elección de especializaciones STEM se puede deber a la *ventaja comparativa* de las mujeres en Lectura, y de los hombres en Matemáticas. El concepto de la ventaja comparativa es uno de los más viejos en economía, desarrollado por el economista británico David Ricardo a principios del siglo XIX. Este concepto dice que, aunque una persona sea más eficiente que todas las demás personas en una tarea –por ejemplo, lavar la ropa–, puede que sea preferible pagar a alguien para que lo haga por él o ella. ¿Por qué? Porque cada hora que dedica a estas tareas tiene un coste de oportunidad y es probable que pudiera emplear ese tiempo en una tarea en la que es más productiva como, por ejemplo, preparar la cena. Si es el caso, esa persona tiene una ventaja comparativa en esta otra tarea (preparar la cena). Traducido a nuestro contexto, este concepto sugiere que, aunque una chica sea muy buena en Matemáticas, si sus habilidades en Lengua son superiores, debería dedicarse a una disciplina relacionada con la Lectura, Literatura o Humanidades, en lugar de a una disciplina relacionada con las Matemáticas.

En el estudio mencionado anteriormente (Breda et al., 2018), los autores utilizan datos de PISA para 64 países y 300.000 estudiantes de 15 años y cal-

culan, para cada alumno, la diferencia entre su rendimiento en Matemáticas y Lengua en las pruebas de PISA. Lo primero que muestran es que la diferencia entre el rendimiento de un alumno en Lectura y Matemáticas es del 80% de una desviación estándar (DE) mayor para las chicas que para los chicos, una magnitud considerada muy grande. Dicho de otra manera, la puntuación obtenida en Matemáticas y Lectura es mucho más similar para los chicos que para las chicas, lo que sugiere que las chicas tienen una ventaja comparativa en Lengua mucho más fuerte que la ventaja comparativa de los chicos en Matemáticas. En segundo lugar, estos autores construyen un índice que mide las intenciones de estudiar asignaturas relacionadas con las Matemáticas en la Educación Secundaria post-obligatoria, en la Universidad, así como las intenciones de elegir una ocupación relacionada con las Matemáticas o las Ciencias. La brecha de género para este índice de *intenciones de especialización en STEM* asciende al 22% de una desviación estándar a favor de los chicos en todo el mundo.

En el siguiente paso, los economistas intentan explicar de dónde vienen las diferencias en dicho índice de *intenciones de especialización en STEM* entre chicas y chicos. Se podría argumentar, por ejemplo, que se debe a las mejores habilidades en esta materia de los chicos. Sin embargo, los autores muestran que cuando controlan por la capacidad matemática en una regresión lineal de estas intenciones sobre una variable que indica el género, la estimación de la brecha de género en el índice de *intenciones de especialización en STEM* se reduce en menos de un 10% en todo el mundo y en la mayoría de los países estudiados. En cambio, la diferencia entre chicos y chicas en la intención de cursar carreras matemáticas desaparece casi por completo cuando se controla por las *diferencias* individuales de habilidades entre Matemáticas y Lectura. Es decir, estos autores pueden explicar la brecha de género en el índice de *intenciones de especialización en STEM* cuando controlan en sus regresiones por la ventaja comparativa de los alumnos en Matemáticas o Lectura. En conjunto, el estudio muestra que las diferencias entre el rendimiento de un alumno en Lectura y Matemáticas pueden explicar hasta un 80% de la brecha de género en la intención de elegir cursos relacionados con las Matemáticas en la educación post-obligatoria. En el caso de España, se estima que un 65% de la brecha de género en intenciones de estudiar carreras STEM se debe a esa mayor ventaja comparativa de las mujeres en Lengua (Breda et al., 2018).

### 3.2. El rol de la cultura y las normas sociales

Como muestra la sección 3.1, las chicas tienden a rendir mejor en Lengua que en Matemáticas, y los chicos mejor en Matemáticas que en Lengua. Además, la diferencia en el rendimiento entre estas dos asignaturas suele ser más grande para las chicas que para los chicos. Esto plantea la cuestión de por qué observamos unas habilidades tan dispares entre estas dos asignaturas para las chicas.

Una posible explicación está relacionada con la cultura y las normas sociales. Hay evidencia considerable que sugiere que en países donde los estereotipos de género son más fuertes, la diferencia en el rendimiento en Matemáticas entre chicas y chicos es más elevada. Uno de esos estudios es el de Luigi Guiso, Ferdinando Monte, Paola Sapienza y Luigi Zingales (Guiso et al., 2008), publicado en 2008 en la revista *Science*. Los autores usan datos de PISA de 2003 para 40 países, donde alumnos de 15 años se evaluaron de la misma prueba en Matemáticas, y muestran que la puntuación media de las chicas en Matemáticas fue de 10,5 puntos inferior a la de los chicos (equivalente a un 2% menos que la puntuación media de los chicos). Para explorar el papel de la cultura y de las normas sociales a la hora de explicar estas diferencias, clasifican los países participantes según varias medidas de equidad de género. Por ejemplo, utilizan el *World Economic Forum Gender Gap Index* (GGI), que refleja las oportunidades económicas y políticas, la educación y el bienestar de las mujeres. Al hacer una regresión de la brecha de género en el rendimiento en Matemáticas a nivel de país sobre el índice de desigualdad de género de ese país, descubren que las sociedades más igualitarias tienen una brecha de género mucho menor en el rendimiento en Matemáticas. De hecho, encuentran que si un país como Turquía –con un nivel de equidad de género bajo– tuviera el mismo nivel de igualdad de género que Suecia, la brecha de género en habilidades matemáticas en Turquía desaparecería.

Se podría argumentar que los autores están detectando simplemente una correlación accidental entre los índices de equidad de género y la brecha de género en Matemáticas. Por ejemplo, puede ser que en realidad el factor que hace que se reduzca la brecha de género en Matemáticas sea el desarrollo económico. Como el desarrollo económico también está positivamente relacionado con la equidad de género en la sociedad, corremos el riesgo de que si no introducimos la variable desarrollo económico en la regresión, el coeficiente de la variable equidad de género recoja tanto el efecto del índice de equidad de género como de la variable desarrollo económico. Es decir, si dejamos la variable desarrollo económico

fuera de la ecuación, podríamos obtener estimaciones sesgadas debido a un problema de variables omitidas (el lector puede dirigirse al capítulo 2 para refrescar estos conceptos). Los autores intentan abordar este problema incluyendo una interacción entre el género y el logaritmo del PIB per cápita (que estaría midiendo el nivel de desarrollo económico de cada país). Sus resultados muestran que la relación entre el índice de igualdad de género y la diferencia de puntuación en los exámenes de Matemáticas se mantiene. Por lo tanto, las diferencias en desarrollo económico entre países, que podrían estar explicando, a su vez, diferencias en la cultura o normas sociales de un determinado país, no son las que estarían explicando los resultados de Luigi Guiso y co-autores.

El estudio citado anteriormente sugiere que la cultura y las normas sociales pueden jugar un papel importante a la hora de explicar la brecha de género en el rendimiento en asignaturas STEM y, en consecuencia, en la elección de estudios superiores en áreas STEM u ocupaciones relacionadas con este ámbito. Sin embargo, podría existir una interpretación alternativa de los resultados presentados. Por ejemplo, en aquellos países en los que las chicas obtienen resultados relativamente mejores en Matemáticas y donde la brecha de género en Matemáticas es nula o incluso en favor de las chicas, es posible que las mujeres estén más preparadas, accedan a mejores puestos de trabajo, ganen salarios más altos y tengan más poder político (Nollenberger et al., 2016). Esto resultaría en una mayor igualdad de género en estos países, y los índices que se emplearon en el estudio anterior así lo captarían. Por lo tanto, las estimaciones presentadas con anterioridad podrían sufrir de un problema de causalidad inversa (ver capítulo 2).

Para abordar este problema, en un estudio publicado en 2016, las economistas Natalia Nollenberger, Núria Rodriguez-Planas y Almudena Sevilla (Nollenberger et al., 2016) emplean una estrategia utilizada en el campo de la epidemiología y que se basa en dos aspectos clave (Fernández, 2011). Primero, se concentran en inmigrantes de segunda generación en un *mismo país de acogida* para estudiar el efecto de los valores de género. Esto les permite estudiar el efecto de un único aspecto de la cultura –los valores transmitidos de generación en generación– porque las instituciones y normas sociales a las que están expuestos los individuos en el país de acogida son las mismas para todos. El segundo aspecto es que, aunque los inmigrantes de segunda generación hayan nacido y crecido en el país de acogida, es probable que sus familias les hayan transmitido *diferentes valores culturales* según su país de origen.

Como Guiso et al. (2008), estas economistas también usan datos de PISA, pero para los años 2003, 2006, 2009 y 2012, con una muestra de 11.527 inmigrantes de segunda generación cuyas familias inmigraron desde 35 países diferentes y que vivían, en el momento de la prueba PISA, en nueve países de acogida diferentes. La estrategia empírica consiste en una regresión de la puntuación individual en la prueba de Matemáticas en PISA sobre una variable que indica si el individuo es mujer y la interacción entre la variable que indica si el género es femenino y el índice GGI del país de ascendencia. Lo más importante de la estrategia de identificación es que en sus regresiones incluyen *efectos fijos de país de ascendencia, país de acogida y cohorte*, así como variables de control individuales. También incluyen una interacción entre el indicador del país de acogida y la variable que indica si el género del alumno es femenino para tener en cuenta la variación en las brechas educativas de género que pueden derivarse de las diferencias culturales o institucionales entre los países de acogida. El coeficiente de interés es el de la interacción entre género femenino y el índice GGI. Este coeficiente nos indica las diferencias de género existentes en los resultados de Matemáticas entre inmigrantes de segunda generación que vendrían explicadas por diferencias en los valores culturales transmitidos por sus padres.

Los resultados indican que un incremento en una desviación estándar en el GGI del país de origen de los padres está asociado con una reducción en 7,47 puntos en la prueba PISA en la brecha de género en Matemáticas. Esto representa el 29% de una desviación estándar en la brecha de género, medida entre los países de ascendencia incluidos en la muestra.

Se podría argumentar que otras fuentes de heterogeneidad entre los individuos, aparte de sus creencias culturales sobre el género, podrían estar afectando sus resultados académicos. Un problema podría ser que las chicas de países más igualitarios podrían estar en escuelas que favorecen a las chicas, por ejemplo, porque tienen más profesores femeninos o porque tienen más profesores de su misma etnia. Si ignorásemos estos factores, las estimaciones del efecto de la cultura podrían estar sesgadas por el efecto de variables omitidas (véase capítulo 2). Los autores incluyen una serie de controles adicionales para abordar posibles sesgos y encuentran que sus estimaciones son robustas a esas inclusiones. Concluyen que las políticas que intentan cambiar las creencias culturales sobre el papel de la mujer en la sociedad pueden ser decisivas para reducir la brecha de género en Matemáticas.

Ejercicio 4. *Discute cómo el problema de la causalidad inversa dificulta el estudio del impacto de la cultura o las normas sociales en la brecha de género en Matemáticas. ¿Cómo intentaron Nollenberger et al. (2016) superar este problema en su estudio sobre los inmigrantes de segunda generación?*

Ejercicio 5. *Este ejercicio requiere descargar datos que están disponibles de forma gratuita en internet. Primero, descarga los últimos datos disponibles sobre el Índice de Desigualdad de Género (IDG) para varios países, proporcionado por las Naciones Unidas en su página web. Segundo, descarga los últimos datos agregados de rendimiento en PISA por sexo y país y calcula la brecha de género en rendimiento en Matemáticas. Estos datos agregados están disponibles en la web de la OCDE. Utilizando algún programa estadístico, como por ejemplo Stata o R, une las dos bases de datos. ¿Cuál es la correlación entre la brecha de género en Matemáticas en PISA y el ranking en IDG? Puedes obtener dicha correlación corriendo una regresión donde la brecha de género en Matemáticas sea la variable dependiente, y el ranking en IDG sea la variable independiente. Discute los resultados y compáralos con los de Guiso et al. (2008) y Anghel et al. (2020).*

### 3.3. Estereotipos de género

La sección 3.2 muestra que las creencias culturales o las normas sociales sobre el papel de la mujer en la sociedad son importantes a la hora de explicar la brecha de género en Matemáticas. La literatura en el campo de la Economía de la Educación ha estudiado, de forma más específica, el rol de los estereotipos de género a la hora de explicar brechas de género en elección de disciplinas u ocupaciones STEM. De acuerdo a la definición del Alto Comisionado para los Derechos Humanos de las Naciones Unidas, un estereotipo de género es una visión generalizada o una idea preconcebida sobre los atributos o las características, o los papeles que poseen o deberían poseer o desempeñar las mujeres y los hombres. Dado lo visto en la sección 3.2, cabría preguntarse lo siguiente: ¿Existen estereotipos negativos sobre las mujeres en el ámbito de las Matemáticas, o más generalmente, en el ámbito de las disciplinas STEM?

La economista italiana Michela Carlana responde a esta pregunta en un estudio publicado en el año 2019 en el *The Quarterly Journal of Economics*. En concreto, Carlana (2019) estudia cómo los estereotipos y los sesgos de género de los profesores pueden afectar a la evaluación que éstos hacen del rendimiento de las niñas en Matemáticas. Los estereotipos que transmiten los profesores

pueden ser especialmente perjudiciales para los alumnos, ya que se ha mostrado que afectan al desarrollo del auto-concepto académico. Si una niña percibe de su profesor o profesora que ella no tiene habilidades altas en Matemáticas, esto puede causar un bajo nivel de autoestima en sus competencias matemáticas, que a la vez puede afectar de manera negativa a su rendimiento en Matemáticas.

Para evaluar el efecto de los estereotipos de género en la brecha de género en Matemáticas, Carlana (2019) mide los prejuicios de género de 1.400 profesores y profesoras del norte de Italia mediante el *Test de Asociación Implícita de Género-Ciencia* (IAT por sus siglas en inglés). Este test fue desarrollado por psicólogos sociales en 1998 (Greenwald et al., 1998) y ha sido ampliamente utilizado en estudios de sesgos raciales y de género. El test consiste en dos tareas. Primero, pide a los sujetos que ordenen nombres masculinos y conceptos científicos a un lado de una pantalla y, en el lado opuesto, clasifiquen los nombres femeninos y conceptos relacionados con las humanidades. En la segunda tarea, se pide a los participantes que clasifiquen a un lado de la pantalla nombres femeninos y conceptos científicos y en el lado opuesto de la pantalla los nombres masculinos y los conceptos relacionados con las humanidades. La idea que subyace en este test es que, si los individuos tienen *asociaciones implícitas* entre el sexo masculino y conceptos científicos, realizarán de manera más rápida la primera tarea (que pide categorizar los nombres masculinos en el mismo lado de la pantalla que los conceptos científicos). Además, a los profesores se les preguntó sobre prejuicios o *asociaciones explícitas*, como, por ejemplo, sus creencias sobre diferencias de género en la capacidad innata para las Matemáticas, además de una pregunta explícita que es capaz de medir valores de género más tradicionales.

En el siguiente paso, Carlana emplea dos estrategias para identificar el efecto causal de los estereotipos de género. La primera estrategia consiste en una regresión que mide la brecha de género *dentro* de una misma clase. Es decir, con observaciones de alumnos agrupados en clases, Carlana puede incluir en sus regresiones efectos fijos de clase. Esto le permite hacer una comparación manteniendo constantes todas las características de los compañeros de clase, la escuela y los profesores (ver capítulo 2). En concreto, corre una regresión de la puntuación obtenida en un examen de Matemáticas o de la opción académica elegida por un individuo en un aula determinada, en función de la interacción de una variable que indica el género del alumno con la puntuación estandarizada en el IAT. El coeficiente de esta interacción mide cómo afecta a la brecha de género en la clase la asignación de un profesor de Matemáticas con mayor nivel de estereotipos de

género. En la segunda estrategia, compara alumnos del mismo género dentro de la misma cohorte en el mismo centro educativo. Esto le permite analizar por separado el efecto de los estereotipos en las chicas y en los chicos.

El primer resultado que muestra es que la brecha de género en el rendimiento en Matemáticas aumenta en un tercio (de 0,15 a 0,20 desviaciones estándar) en las clases asignadas a un profesor de Matemáticas que asocia implícitamente a los chicos con las Matemáticas, en comparación con las clases asignadas a un profesor que asocia a las chicas con las Matemáticas. El segundo resultado, utilizando la segunda de sus estrategias, muestra que tener un profesor con una actitud "matemática masculina" tiene un efecto negativo en las alumnas, mientras que una actitud "matemática femenina" tiene un efecto positivo en las alumnas. En cambio, los estereotipos de género de los profesores no afectan a los resultados de los chicos. Además, muestra que los estereotipos son especialmente dañinos para las alumnas con bajo rendimiento en Matemáticas al comenzar la etapa secundaria. Estos resultados son coherentes con la teoría de *stereotype threat* desarollada por los psicólogos sociales Claude Steele y Joshua Aronson (Steele y Aronson, 1995). Según esta teoría, el propio miedo a confirmar un estereotipo negativo (por ejemplo, "las chicas son malas en Matemáticas") provoca estrés cognitivo y reduce el rendimiento, lo que conduce a la confirmación del estereotipo negativo y a un círculo vicioso.

Este estudio muestra como los estereotipos de género pueden contribuir a la brecha de género en Matemáticas en la etapa Secundaria. Cabe destacar que como hemos descrito en la sección 2.3, otros estudios han mostrado que existen también profesores con sesgos negativos sobre los chicos, y que esto puede explicar el peor rendimiento de los chicos en la escuela en general. También es importante recordar que hoy en día, la brecha de género en Matemáticas medida en PISA casi ha desaparecido. Hay muchísimas chicas con un rendimiento alto en Matemáticas, y aun así seguimos viendo que las chicas siguen muy infrarrepresentadas tanto en especializaciones STEM en la educación post-obligatoria como en ocupaciones STEM.

El sesgo de género de los profesores es sólo uno de los posibles canales a través de los cuales los estereotipos de género pueden afectar negativamente al rendimiento de las mujeres en Matemáticas. Otro estudio publicado recientemente por el *Centre for Economic Performance* de la London School of Economics analiza el papel de los estereotipos de género en el seno familiar (Nicoletti et al., 2022).

Nicoletti, Sevilla y Tonei usan datos de encuestas sobre las creencias autodeclaradas de los padres australianos acerca de las aptitudes de sus hijos en Matemáticas y Lectura, y cruzan estos datos con registros administrativos sobre el rendimiento de los niños en las pruebas nacionales estandarizadas de Matemáticas y Lectura a la edad de 8-9 años. Las autoras muestran tres resultados interesantes que confirmarían la existencia de estereotipos negativos respecto a las habilidades matemáticas de las chicas, también en el seno familiar. Primero, estas autoras encuentran que los progenitores tienen una probabilidad de sobrestimar las habilidades de matemáticas de los hijos más alta que de sobrestimar las habilidades Matemáticas de las hijas. Segundo, encuentran que el sesgo de los progenitores respecto a la sobrestimación de las habilidades matemáticas de los chicos desaparece después de recibir información sobre los resultados académicos de sus hijos (es decir, una vez reciben información sobre el rendimiento de sus hijos e hijas medido a través de pruebas estandarizadas). Por último, muestran que el sesgo de género de los progenitores en detrimento de las niñas contribuye a explicar la ampliación de la brecha de género en habilidades matemáticas a lo largo de la infancia. Esto sugiere que la exposición a los estereotipos de género en la familia influye negativamente en la capacidad de las niñas para desarrollar todo su potencial en disciplinas STEM.

### 3.4. Ausencia de referentes en STEM

Una posible explicación de por qué muchas chicas, incluso las que obtienen muy buenos resultados en Matemáticas, no eligen carreras universitarias, formación profesional u ocupaciones STEM, es la falta de referentes femeninos en estas áreas. Dado que las mujeres están tan poco representadas en algunas áreas STEM, la probabilidad de que las niñas encuentren un modelo o referente femenino en esos ámbitos es baja. Por un lado, podría ser la falta de *profesores femeninos* en disciplinas STEM la que afecta de manera negativa la motivación de las chicas de seguir estudiando estas disciplinas. Por otro lado, podría ser la falta de *referentes en el mundo profesional y académico* lo que hace que las niñas subestimen sus posibilidades en este ámbito.

El primer canal, es decir, la falta de profesores femeninos en asignaturas STEM, ha sido estudiado ampliamente. Uno de los primeros artículos es el publicado por Scott Carrell, Marianne Page y James West en *The Quartely Journal of Economics* (Carrell et al., 2010). El reto principal al que se enfrentan estos

autores es que el género del profesor podría no ser aleatorio. Así, las alumnas con mayores habilidades en Matemáticas podrían estar eligiendo un profesor de Matemáticas de género femenino porque sienten que una profesora las trataría de forma menos sesgada, y por tanto, podrían obtener mejores resultados que con un profesor de género masculino. De ser así, esto introduciría un sesgo de selección en las estimaciones. Una regresión que relaciona el género del profesor con los resultados académicos y/o las elecciones de cursos en la Universidad estaría sesgada por el hecho de que las alumnas con mejores habilidades eligen matricularse con profesoras en lugar de profesores, pero estas alumnas hubieran obtenido buenos resultados aún si hubiesen escogido matricularse con profesores de género masculino. El lector puede recordar el concepto de sesgo de selección en el capítulo 2.

Para abordar este reto y poder obtener estimaciones causales, los autores utilizan una característica especifica de una Universidad estadounidense, es decir la asignación aleatoria de los estudiantes a profesores de Matemáticas y Ciencias y el hecho de que todos los alumnos tienen que tomar ciertos cursos obligatorios de introducción y seguimiento de Matemáticas, Física, Química, Ingeniería, Historia e Inglés. Esta aleatorización y la obligatoriedad de cursar ciertos cursos evita una selección sesgada de alumnos y alumnas a: 1) profesores que podrían variar sistemáticamente en sus sesgos de género; y 2) asignaturas relacionadas con las Matemáticas.

Los autores muestran que estar expuesto a un profesor de género femenino tiene un efecto importante en el rendimiento de las alumnas en las clases de Matemáticas y Ciencias, y en la probabilidad de que las alumnas con mejores resultados sigan cursos de Matemáticas y Ciencias en el futuro, y se gradúen con un título STEM. En cambio, el género del profesor no tiene impacto en los resultados o elecciones de los alumnos varones. Sin embargo, los efectos de los profesores no son fáciles de interpretar. Podrían deberse a diferencias de género en las prácticas docentes y el comportamiento en clase, más que por el hecho de que las profesoras actúen como referentes para las chicas.

Otro estudio, analiza directamente el efecto de estar expuesto a un referente femenino. Breda et al. (2023) evalúan un programa del mundo real llamado "Por las niñas y la ciencia" (*Pour les Filles et la Science*) que fue lanzado en 2014 por la Fundación L'Oréal para animar a las niñas a explorar las carreras STEM. El programa consiste en intervenciones de una hora durante

el horario escolar impartidas por jóvenes científicas o jóvenes profesionales empleadas en la división de Investigación e Innovación del grupo L'Oréal. Utilizan datos de 19.451 alumnos de 416 clases de 10º curso y 185 clases de 12º curso (itinerario científico) en institutos públicos y privados de la región de París. La mitad de las clases fueron asignadas aleatoriamente a recibir la visita de uno de los 56 participantes. La otra mitad no recibió ninguna intervención. El impacto de este programa se puede medir de forma causal debido a su diseño como experimento de control aleatorio (ver capítulo 2). Un aspecto crucial del programa es que, en el sistema francés, los estudiantes toman decisiones educativas irreversibles al final de los cursos 10º y 12º. Después del primer año de Bachillerato (10º curso), aproximadamente el 80% de los alumnos siguen en el Bachillerato general en los dos últimos cursos (11º y 12º) y el 20% se orienta hacia el Bachillerato tecnológico. Al final del último año de Bachillerato (12º), los estudiantes eligen los programas que quieren estudiar en la Universidad. Esto les permite estudiar el efecto del programa en la matriculación en Ciencias, Tecnología, Ingeniería y Matemáticas en la Universidad.

En el 10º curso, las visitas a las aulas no tuvieron un impacto detectable en la probabilidad de que los chicos y las chicas se matricularan en la rama de Ciencias en 11º curso. En cambio, el programa indujo un aumento significativo de 2,4 puntos porcentuales en la matrícula universitaria en disciplinas STEM entre las chicas de 12º curso, equivalente a un aumento del 8% sobre la tasa de referencia del 29%. El efecto sobre los chicos fue insignificante. En términos de los mecanismos que pueden explicar este cambio en la elección de las chicas, los autores muestran que la intervención afectó de manera positiva la *percepción de los trabajos relacionados con la ciencia*. En cambio, la intervención no afectó la afinidad autodeclarada por las asignaturas científicas, ni mejoró los resultados académicos en estas asignaturas, y sólo mejoró ligeramente su autoconcepto de las Matemáticas. Esto indica que el problema de la infrarrepresentación de las mujeres en los trabajos relacionados con la ciencia podría venir de una percepción errónea de estos trabajos, y no necesariamente por una falta de autoconfianza o habilidades de las mujeres en estos ámbitos. Los autores mencionan como ejemplo la posible percepción de que los trabajos en ámbitos científicos (como la Ingeniería), tengan una mayor dificultad para conciliar trabajo con la vida familiar.

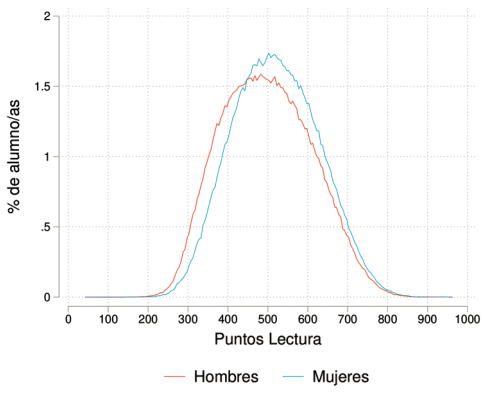
### 3.5. El rol de los contextos competitivos

Como se describe en la sección 3.1, la brecha de género media en Matemáticas ha disminuido en los últimos años y ahora es casi nula en la mayoría de los países desarrollados. Sin embargo, los datos siguen mostrando una falta de chicas entre los alumnos con mejores resultados. Esto se ilustra en el gráfico 5, que muestra histogramas de distribuciones de puntos en Lectura y Matemáticas en la prueba PISA de 2018. Las chicas superan a los chicos de media y están sobrerrepresentadas en los primeros puestos en Lectura. Para Matemáticas, las distribuciones de puntuación casi se superponen, pero en la parte derecha de la distribución (dónde se encuentran los alumnos con mejores puntuaciones) hay más chicos que chicas. ¿Por qué hay menos chicas que chicos entre los mejores en Matemáticas?

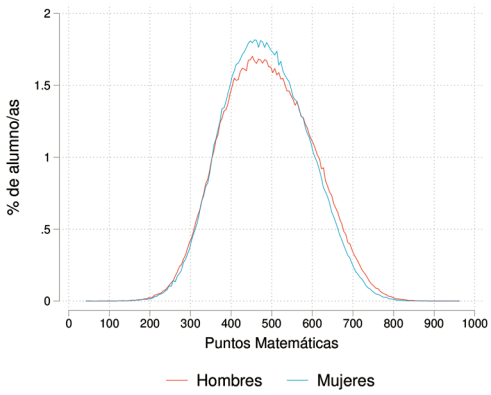
Cada vez más evidencia apunta a una explicación que puede resultar incómoda: las diferencias pueden atribuirse en parte al modo diferencial en que hombres y mujeres responden a entornos competitivos de realización de exámenes. En esta sección analizaremos evidencia de que esto es así. Como señalan Muriel Niederle y Lise Vesterlund en un artículo de referencia publicado en el *Journal of Economic Perspectives* (Niederle y Vesterlund, 2010), las presiones competitivas

**Gráfico 5. Distribución de la competencia en Lectura y Matemáticas, por sexo.**

Panel A: Lectura




Panel B: Matemáticas




Nota: Este gráfico es un histograma del rendimiento utilizando un tamaño de intervalo de cinco puntos de puntuación.
Fuente: Figura II.7.4 de OECD (2019).

asociadas a la realización de exámenes pueden dar lugar a resultados que no reflejan los resultados que se obtendrían en entornos menos competitivos. Dicho de otra manera: el hecho de que los chicos rinden mejor en ciertas pruebas y bajo alta presión competitiva, no necesariamente implica que sus habilidades en Matemáticas sean mejores que las de las chicas. Debido a la forma en que se administran los exámenes y a que muchas oportunidades académicas dependen de los resultados de dichos exámenes, es probable que las mujeres se encuentren en desventaja al no desarrollar todo su potencial en pruebas altamente competitivas (*high-stake exams)*.

Un estudio reciente que analiza los efectos de los entornos competitivos en la brecha de género en Matemáticas es el de los economistas Nagore Iriberri y Pedro Rey-Biel (Iriberri y Rey-Biel, 2019). Utilizando datos del Concurso de Primavera de Matemáticas en España, muestran que, aunque las chicas tienen notas en Matemáticas similares o incluso más altas que los chicos, en el concurso los chicos rinden mejor que las chicas, y que a medida que avanzan de una ronda a la siguiente, esta brecha aumenta. En la primera ronda, que sucede en el colegio del participante, la brecha de género es del 34% de una desviación estándar. En la segunda y final ronda del concurso, que tiene lugar en una universidad pública de Madrid, la brecha aumenta en un 9% de una desviación estándar. Los autores muestran que el aumento en la brecha de género entre las dos rondas se debe a un incremento en la presión competitiva del concurso cuando los participantes pasan a la segunda y final ronda, y no a explicaciones alternativas, como por ejemplo una selección diferencial de chicos y chicas para el concurso, discriminación o reacciones diferentes entre chicos y chicas al incremento de la dificultad de las preguntas entre la primera y segunda ronda. Un aspecto interesante del concurso es que las preguntas no respondidas se premian en comparación con las preguntas no acertadas (los participantes reciben cero puntos si la respuesta es incorrecta, un punto si no responden y cinco puntos cuando la respuesta es correcta). Los autores muestran que las chicas dejan más preguntas sin responder que los chicos, y que esa diferencia aumenta en la segunda ronda. Esto sugiere una mayor aversión al riesgo y una disposición menor a adivinar de las chicas, como posible explicación por la cual emerge una brecha de género en el contexto de una competición comparado con el contexto de clases de Matemáticas normales. Por lo tanto, la manera en la que se premian o penalizan las preguntas no respondidas puede tener un efecto importante en la brecha de género en este tipo de concursos.

En otro estudio, los mismos autores llevan a cabo un experimento de campo natural a gran escala para comprobar y comprender los mecanismos que subyacen a las diferencias de género en la disposición a adivinar (Iriberri y Rey-Biel, 2021). El contexto de este experimento vuelve a ser el del Concurso de Primavera de Matemáticas. Para el experimento, varían la manera en la que se puntúan las preguntas no respondidas. Para las primeras 13 de las 25 preguntas del examen, los alumnos obtienen cero puntos tanto para las preguntas no respondidas como para las preguntas contestadas de forma errónea, y para las últimas 12 preguntas del examen, las respuestas erróneas puntúan 0 y las preguntas omitidas un punto. Esto último implica una recompensa por no responder. Las respuestas correctas siempre puntúan cinco puntos. Al tener observaciones múltiples por cada estudiante (una por pregunta), y dos formas de puntuar las preguntas sin respuesta también para cada individuo, Iriberri y Biel pueden comparar la disposición a adivinar para cada uno de los participantes en cada contexto, y el rendimiento general en ambas partes del examen. Es decir, pueden utilizar una estrategia de investigación que incluya efectos fijos a nivel de estudiante. Los resultados muestran que, en la primera parte del examen, aquella sin recompensa por no responder, las mujeres omiten más preguntas que los hombres (0,17 desviaciones estándar más con respecto a la media), a pesar de que la estrategia dominante sería responder a todas las preguntas cuando no hay recompensa alguna. Y lo que es más importante, esta diferencia aumenta (en 0,14 desviaciones estándar adicionales con respecto a la media) en la segunda parte de la prueba, donde se recompensan las preguntas no contestadas con un punto. La conclusión de este estudio es que una norma de calificación que recompensa, aunque sólo ligeramente, la omisión de preguntas puede actuar en detrimento de las chicas y contribuir a una mayor brecha de género en exámenes competitivos en Matemáticas.

Ejercicio 6. *Discute otros ámbitos de pruebas de alta importancia* ("high-stakes exams") *donde el diseño de las mismas podría afectar de manera negativa el desempeño de las mujeres. ¿Qué recomendaciones de políticas darías para cambiar su diseño para ser más equitativo?*

## 4. CONCLUSIONES

Los datos internacionales muestran obstinadamente que los chicos tienen un menor desempeño educativo y que las mujeres siguen estando infrarrepresenta-

das en las carreras universitarias y profesionales relacionadas con las disciplinas STEM (Ciencia, Tecnología, Ingeniería y Matemáticas, por sus siglas en inglés). En este capítulo hemos presentado algunas razones de por qué nos deben preocupar estas dos principales brechas de género y hemos analizado algunas de las posibles explicaciones de estos fenómenos.

En el caso del menor desempeño educativo de los chicos, analizamos el papel de las habilidades no cognitivas, el rol de las familias y de los factores socioeconómicos, el impacto de la escuela y del profesorado, y, por último, damos algunas pinceladas sobre la relación existente entre las condiciones del mercado laboral, el coste de oportunidad de seguir estudiando, y el desempeño educativo del alumnado.

En cuanto a la infrarrepresentación de las mujeres en STEM, examinamos el papel de la cultura, normas sociales y estereotipos de género, la ausencia de referentes femeninos entre los profesores de Matemáticas y disciplinas científicas, así como en muchas profesiones relacionadas con las Ciencias, y, por último, el rol de los contextos competitivos.

Del análisis presentado se desprende que, aunque todos los factores analizados son relevantes, y el conocimiento en este campo ha avanzado de forma considerable en la última década, estamos lejos de comprender todas las causas subyacentes de las diferencias de género observadas en nuestros sistemas educativos, y que es probable que muchos factores interactúen entre sí.

## BIBLIOGRAFÍA

Altonji, Joseph G, Peter Arcidiacono y Arnaud Maurel (2016). The analysis of field choice in college and graduate school: Determinants and wage effects, en *Handbook of the Economics of Education, Vol. 5*, Elsevier, 305-396.

Anghel, Brindusa, Núria Rodríguez-Planas y Anna Sanz de Galdeano (2020). Gender equality and the math gender gap. *Banco de España Documentos de Trabajo* (2031).

Aparicio-Fenoll, Ainhoa (2016). Returns to education and educational outcomes: The case of the Spanish housing boom. *Journal of Human Capital, 10*(2), 235-265.

Autor, David (2010). US labor market challenges over the longer term. *Federal Reserve Board of Governors*.

— (2019). David Figlio, Krzysztof Karbownik, Jeffrey Roth y Melanie Wasserman. Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes. *American Economic Journal: Applied Economics, 11*(3), 338-81.

Baker, Michael y Kevin Milligan (2016). Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries. *Journal of Human Capital, 10*(4), 399-441.

Becker, Gary S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy, 70*(5), 9-49.

Becker, Gary S, William HJ Hubbard y Kevin M Murphy (2010). Explaining the worldwide boom in higher education of women. *Journal of Human Capital, 4*(3), 203-241.

Belfield, Clive R y Henry M Levin (2007). *The price we pay: Economic and social consequences of inadequate education*. Brookings Institution Press.

Bertrand, Marianne y Jessica Pan (2013). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics, 5*(1), 32-64.

Black, Sandra E, Chandra Muller, Alexandra Spitz-Oener, Ziwei He, Koit Hung y John Robert Warren (2021). The importance of STEM: High school knowledge, skills and occupations in an era of growing inequality. *Research policy, 50*(7), 104249.

Blau, Francine D y Lawrence M Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature, 55*(3), 789-865.

Breda, Thomas, Elyès Jouini y Clotilde Napp (2018). Societal inequalities amplify gender gaps in math. *Science*, *359*(6381), 1219-1220.

— (2023). Julien Grenet, Marion Monnet y Clémentine Van Effenterre. How Effective are Female Role Models in Steering Girls Towards Stem? Evidence from French High Schools. *The Economic Journal*, *133*(653), 1773-1809.

Carlana, Michela (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics, 134*(3), 1163-1224.

Carrell, Scott E, Marianne E Page y James E West (2010). Sex and science: How profesor gender perpetuates the gender gap. *The Quarterly Journal of Economics, 125* (3), 1101-1144.

Cavaglia, Chiara, Stephen Machin, Sandra McNally y Jenifer Ruiz-Valenzuela (2021). Gender, achievement, and subject choice in English education. *Oxford Review of Economic Policy, 36*(4), 816-835.

Chetty, Raj y Nathaniel Hendren (2018). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics, 133*(3), 1163-1228.

Dee, Thomas S. (2005). A Teacher like Me: Does Race, Ethnicity, or Gender Matter? *The American Economic Review, 95*(2), 158-165.

Dee, Thomas S (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3), 528-554.

Else-Quest, Nicole M, Janet Shibley Hyde, H Hill Goldsmith y Carol A Van Hulle (2006). Gender differences in temperament: a meta-analysis. *Psychological Bulletin, 132*(1), 33.

Encinas-Martín, Marta Y Michelle Cherian (2023). *Gender, Education and Skills.* OECD Skills Studies, OECD Publishing, Paris.

Fernández, Raquel (2011). Chapter 11 - Does Culture Matter?, en Jess Benhabib, Alberto Bisin y Matthew O. Jackson (eds.), *Handbook of Social Economics, Vol. 1,* North-Holland, 2011, 481-510.

Gershenson, Seth y Stephen B Holt (2015). Gender gaps in high school students' homework time. *Educational Researcher, 44*(8), 432-441.

Gortazar, Lucas, David Martínez de Lafuente y Ainhoa Vega-Bayo (2022). David Martínez de Lafuente [PDF] de iseak.eu Comparing teacher and external assessments: Are boys, immigrants, and poorer students undergraded? *ISEAK Working Paper*, *115*, 103725.

Greenwald, A. G., D. E. McGhee y J. L. K. Schwartz (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology,* (6), 1464-1480.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza y Luigi Zingales (2008). Culture, gender and math. *Science-New York Then Washington, 320*(5880), 1164.

Iriberri, Nagore y Pedro Rey-Biel (2019). Competitive pressure widens the gender gap in performance: Evidence from a two-stage competition in mathematics. *The Economic Journal, 129*(620), 1863-1893.

— (2021). Brave boys and play-it-safe girls: Gender differences in willingness to guess in a large scale natural field experiment. *European Economic Review, 131*, 103603.

Montalbán, José y Jenifer Ruiz-Valenzuela (2022). ¿Cómo aumentar el acceso de las mujeres a disciplinas científicas y técnicas en la Educación Superior? *Esade EcPol Policy Brief*, 29.

Montalbán, José Castilla (2021). Fracaso escolar en España: ¿Por qué afecta tanto a los chicos y alumnos de bajo nivel socioeconómico? *Esade EcPol Policy Brief, 13*.

Nicoletti, Cheti, Almudena Sevilla y Valentina Tonei (2022). Gender Stereotypes in the Family. *IZA Discussion Paper*, 15773.

Niederle, Muriel y Lise Vesterlund (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of Economic Perspectives, 24*(2), 129-44.

Nollenberger, Natalia, Núria Rodríguez-Planas y Almudena Sevilla (2016). The Math Gender Gap: The Role of Culture. *American Economic Review, 106*(5), 257-61.

OECD (2019). *PISA 2018 Results (Volume II).*

Peri, Giovanni, Kevin Shih y Chad Sparber (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics, 33*(S1), S225-S255.

Rees, Daniel I. y H.Naci Mocan (1997). Labor market conditions and the high school dropout rate: Evidence from New York State. *Economics of Education Review, 16*(2), 103-109.

Schaede, Ursina y Ville Mankki (2022). *Quota vs quality? Long-term gains from an unusual gender quota.*

Steele, Claude M. y Joshua Aronson (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811.

Terrier, Camille (2020). Boys lag behind: How teachers' gender biases affect student achievement. *Economics of Education Review, 77*, 101981.

# Capítulo 5d
# HABILIDADES NO COGNITIVAS

Ildefonso Méndez
Departamento de Economía Aplicada
Universidad de Murcia
ORCID: https://orcid.org/0000-0002-2638-5071

Gema Zamarro
Department of Education Reform
University of Arkansas
ORCID: https://orcid.org/0000-0002-3893-7523

**Resumen**

Las habilidades no cognitivas son un concepto amplio que incluye los rasgos de personalidad, atributos, disposiciones, habilidades sociales, actitudes, capacidades y recursos personales, independientes de la capacidad intelectual. La literatura ha mostrado como estas habilidades y en el bienestar presentan un papel importante en el rendimiento educativo, laboral y bienestar social a lo largo de la vida. Desde una perspectiva multidisciplinar este capítulo discute los conceptos que se incluyen dentro de las habilidades no cognitivas y cómo se relacionan entre sí, así como las dificultades e innovaciones en su medición. A continuación, el capítulo repasa los factores asociados con su desarrollo. Estos incluyen no sólo factores hereditarios, sino también otros factores importantes como la renta familiar, los estilos de crianza, el tamaño de la familia, el entorno escolar y cultural, entre otros. Finalmente, el capítulo discute la literatura que muestra como las intervenciones más costo-eficaces para el desarrollo de las habilidades no cognitivas son aquellas que están dirigidas a los más jóvenes, y que involucran a la familia. Asimismo, los beneficios de las intervenciones eficaces suelen ser mayores para aquellos colectivos socialmente más vulnerables, lo que las hace aún más importantes en la lucha contra la desigualdad.

## 1. ¿QUÉ SON LAS HABILIDADES NO COGNITIVAS?

La investigación en economía ha demostrado que hay un conjunto de habilidades, independientes del cociente intelectual, particularmente moldeables a edades tempranas y con un notable impacto en el bienestar a lo largo de la infancia, la adolescencia y la vida adulta. Este hallazgo ha abierto la puerta al diseño de intervenciones y programas que estimulen el desarrollo de estas habilidades. Estos programas han resultado ser particularmente eficaces, sobre todo cuando se dirigen a la población infanto-juvenil, para reducir la transmisión intergeneracional de la pobreza, las habilidades socioemocionales y, por tanto, la convivencia, el abandono educativo temprano, etc.

El presente capítulo revisa el estado del conocimiento científico sobre las habilidades no cognitivas, su relevancia en resultados educativos, laborales, personales, de salud, etc., sus determinantes y las características de los programas e intervenciones eficaces en su estimulación desde edades tempranas.

El Departamento de Educación de los Estados Unidos definió las habilidades no cognitivas en 2013 como el conjunto de rasgos de carácter, personalidad, atributos, disposiciones, habilidades sociales, actitudes, capacidades y recursos personales, independientes de la capacidad intelectual. Se presentan en contraste con lo que se considera la capacidad intelectual pero junto a ella ayudan a predecir elementos fundamentales para el bienestar en la vida como es el rendimiento académico y laboral, la salud y el bienestar social.

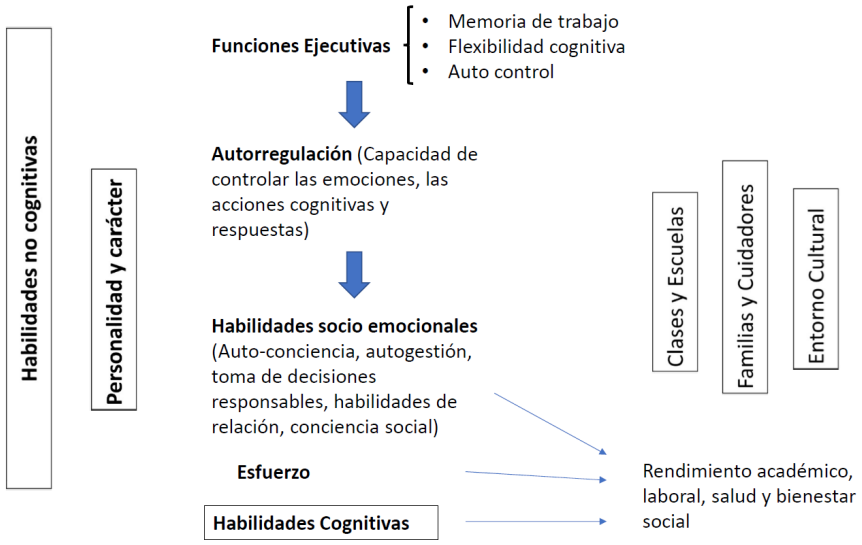
Habilidades no cognitivas es el nombre utilizado por economistas para este conjunto de cualidades individuales y es un concepto amplio que engloba una multitud de atributos interrelacionados, definidos y estudiados en su mayoría por expertos en psicología. Dado su origen interdisciplinar, estas habilidades han recibido distintos nombres en las literaturas económica, psicológica, pedagógica o filosófica, incluyendo rasgos de personalidad, carácter o temperamento, funciones ejecutivas, virtudes, habilidades socioemocionales, y habilidades no cognitivas, entre otros. Estos distintos términos ponen énfasis en distintas características de las habilidades no cognitivas. Conceptos como personalidad, carácter o temperamento, enfatizan la permanencia y carácter hereditario de las habilidades no cognitivas. Mientras que conceptos como las habilidades socioemocionales o habilidades no-cognitivas enfatizan el hecho de que estas habilidades se pueden aprender.

Las *funciones ejecutivas*, por su parte, constituyen una serie de procesos mentales que permiten planificar, mantener la atención, recordar, y manejar múlti-

ples tareas a la vez y están en la base de algunas habilidades no-cognitivas. Las funciones ejecutivas dependen de tres tipos de procesos mentales: la *memoria de trabajo*, la *flexibilidad cognitiva*, y el *control inhibitorio*. La memoria de trabajo nos permite mantener y manipular distintas piezas informativas en un corto periodo de tiempo. La flexibilidad cognitiva nos permite mantener y cambiar el centro de atención según distintas demandas del entorno. Finalmente, el control inhibitorio o auto-control es el componente más directamente relacionado con el concepto económico de habilidades no cognitivas y nos permite fijar prioridades y resistir distracciones y reacciones impulsivas. Las funciones ejecutivas están altamente correlacionadas entre sí y, para su éxito, trabajan en conjunto las unas con las otras.

Las funciones ejecutivas y habilidades no cognitivas ayudan al rendimiento académico, laboral y bienestar social. Esta relación, sin embargo, no ocurre necesariamente en el sentido contrario. Aunque existe un componente hereditario importante de las habilidades no cognitivas, de entre el 30% y el 60% (Bouchard y Loehlin, 2001), tal y como se explica más adelante en este capítulo, estas habilidades se desarrollan a lo largo de la vida con periodos críticos durante la infancia y adolescencia, y su desarrollo se ve influido por factores culturales y el entorno familiar y educativo.

**Gráfico 1. Funciones ejecutivas, habilidades socioemocionales y rendimiento académico, laboral y bienestar social.**



El gráfico 1 muestra un esquema de cómo las distintas habilidades no cognitivas definidas en esta sección se relacionan entre sí, los factores que las influyen, y la influencia que estas habilidades tienen.

## 2. ¿CÓMO SE MIDEN LAS HABILIDADES NO COGNITIVAS?

Obtener medidas de habilidades no cognitivas no es una tarea fácil. La mayoría de los investigadores y evaluadores usan medidas basadas en encuestas donde las habilidades no cognitivas de los encuestados se miden estudiando sus respuestas a una serie de preguntas con formatos de respuesta de elección múltiple. Sin embargo, las medidas basadas en encuestas pueden estar sesgadas si los encuestados responden de manera que sus respuestas sean socialmente aceptables, sin que representen necesariamente la realidad. Este fenómeno se conoce en la literatura como "*social desirability bias*". Otra fuente de posibles sesgos en las medidas de habilidades no cognitivas basadas en encuestas se debe a la posibilidad de que los encuestados difieran en el uso de las escalas de respuesta de las preguntas. Este fenómeno, conocido en la literatura como "*reference group bias*" traducido como sesgos por diferencias en los valores de referencia, puede ser un problema importante al comparar niveles de habilidades no cognitivas entre grupos o áreas geográficas de diferentes culturas (o tradiciones culturales).

Una manera alternativa en la que los investigadores han medido habilidades no cognitivas sin tener que confiar en encuestas, es a través de la medición directa de ciertos comportamientos asociados con distintas habilidades no cognitivas. Por ejemplo, los investigadores han usado medidas de comportamientos académicos como son las notas de curso, el número de faltas de asistencia, o expulsiones escolares, como medidas de baja motivación de los estudiantes.

Otra alternativa consiste en medir ciertas habilidades no cognitivas específicas a través de experimentos, con tareas directas donde se pueda observar el nivel de competencia relacionado con una habilidad no cognitiva determinada de la persona de manera directa, cuando se le analice realizando una tarea específica. Un ejemplo clásico es la prueba de la golosina diseñada por Mishel et al. (1972) para medir los niveles de autocontrol en los niños. Por último, la literatura reciente propone el uso de metadatos, eso es "datos sobre los datos", de exámenes y encuestas para obtener medidas de habilidades no cognitivas.

## 2.1. Medición a través de encuestas

La manera principal en la que los investigadores han intentado medir distintas habilidades no cognitivas es a través de cuestionarios en los que los encuestados responden a una serie de preguntas frecuentemente usando lo que se conoce como una escala de "Likert" donde declaran su nivel de acuerdo o desacuerdo con una serie de declaraciones. Dos ejemplos muy utilizados en la literatura corresponden a los cuestionarios para medir los cinco rasgos de personalidad (Los cinco grandes o "Big 5") y la escala de perseverancia a largo tiempo ("Grit scale").

### *2.1.1. Medidas de personalidad: los cinco grandes o "Big 5"*

Múltiples estudios en psicología y economía han utilizado la taxonomía de los cinco grandes o "Big 5" para medir los siguientes rasgos de personalidad: Apertura, Conciencia, Extraversión, Amabilidad y Neuroticismo. La apertura captura la tendencia de las personas a experimentar y aprender cosas nuevas. Tiene que ver con la imaginación y también la creatividad de las personas. La conciencia refleja los niveles de perseverancia, concentración, control de impulsos y capacidad para alcanzar objetivos. La extraversión mide la capacidad de relacionarse con otros, la expresividad emocional y la capacidad para la interacción social. La amabilidad mide la capacidad de empatía hacia los demás, el altruismo y nivel de afecto. Finalmente, el neuroticismo mide la capacidad física y emocional para enfrentarse al estrés de la vida diaria de una persona. Aquellos con altos niveles de neuroticismo experimentan ansiedad e irritabilidad frecuente como respuesta a los niveles de estrés diarios del trabajo o vida personal.

Siguiendo el trabajo pionero de Fiske (1949), investigadores en psicología han estudiado y desarrollado distintos cuestionarios para capturar rasgos de personalidad y han encontrado que estos cuestionarios recogen consistentemente estas cinco dimensiones de la personalidad. Por ejemplo, John et al. (1991) y John et al. (2008) desarrollaron la serie de preguntas que se conoce como el inventario de los cinco grandes ("The Big Five Inventory"), que consiste en un cuestionario de 44 afirmaciones donde el encuestado reporta su nivel de acuerdo en una escala "Likert" de cinco puntos que va desde 1 ("Totalmente en desacuerdo") a 5 ("Totalmente de acuerdo"). En general, el modelo de personalidad de los cinco grandes y sus cuestionarios es uno de los más utilizados en la investigación en psicología y también más recientemente en economía.

### *2.1.2. Medidas de perseverancia a largo plazo o "Grit"*

Duckworth et al. (2007) definen "Grit" como la tendencia a mantener el interés y esfuerzo a lo largo del tiempo para conseguir objetivos de largo plazo. Junto a sus coautores, la psicóloga Angela Duckworth desarrolló dos posibles cuestionarios para medir el concepto de "Grit". Se puede acceder a estos cuestionarios en la web aquí: https://angeladuckworth.com/grit-scale/.

De manera similar al inventario de los cinco grandes, los cuestionarios para medir perseverancia a largo plazo o "Grit" consisten en una serie de afirmaciones donde los encuestados reportan su nivel de acuerdo en una escala "Likert" de 5 puntos. Estas afirmaciones incluyen, entre otras, frases como "Soy muy trabajador/a", "Soy diligente", o "Las ideas y proyectos nuevos a veces me distraen de ideas y proyectos anteriores" entre otras. Para la versión del cuestionario de perseverancia a largo plazo o "Grit" en español se puede ver el trabajo de Arco-Tirado et al. (2018).

### *2.1.3. Limitaciones de las medidas basadas en encuestas*

A pesar de la evidencia que demuestra la validez del uso de los cuestionarios como el inventario de los cinco grandes o la escala de perseverancia a largo plazo o "Grit", esta manera de medir habilidades no cognitivas tiene limitaciones importantes. En particular, tal y como se mencionó previamente, las medidas basadas en encuestas pueden verse afectadas por dos sesgos importantes "*social desirability bias*" y "*reference group bias*". Estos sesgos ocurren cuando los encuestados tienen una tendencia a dar respuestas que son más socialmente aceptables ("*social desirability bias*"), o cuando difieren en el uso de las categorías de respuesta dada su cultura o grupo social ("*reference group bias*"). Por ejemplo, cuando a los estudiantes se les pregunta sobre su nivel de motivación escolar y se les pide responder en una escala ordinal, como las presentadas para el inventario de los cinco grandes o el cuestionario de perseverancia a largo plazo o "Grit", sus estándares individuales de lo que significa ser un estudiante motivado pueden afectar a sus respuestas. En particular, la idea de lo que significa ser un estudiante motivado está probablemente influida por factores culturales y el contexto de sus colegios. Es decir, aquello que hace que dos estudiantes con un mismo nivel objetivo de motivación, pero que viven en distintas regiones, que vienen de distintas culturas, que son de distinto género o que acuden a colegios distintos, respondan de manera distinta. Por ejemplo, hay estudios que han demostrado

como los estudiantes que viven en EE. UU. y que se identifican con un origen asiático tienden a responder a las respuestas de la escala de perseverancia a largo plazo o "Grit" de manera que resulta en valores menores que aquellos que no se identifican con el origen asiático, lo que contrasta con un mayor rendimiento académico medio de los estudiantes de origen asiático.

El método de anclaje por viñetas hipotéticas se ofrece como una posible solución a estos problemas de medida producidos por diferencias en los valores de referencia o "*reference group bias*". Este método se basa en pedir a los encuestados que evalúen no sólo un concepto determinado, por ejemplo, sus niveles de motivación, sino también los mismos niveles representados en una serie de viñetas hipotéticas. Estas viñetas hipotéticas representan historias cortas describiendo una persona hipotética con un cierto nivel objetivo en el concepto a medir. Como todos los encuestados responden a las mismas viñetas, asumiendo que todos las interpretan de la misma manera, las diferencias observadas en la evaluación de estas historias hipotéticas se pueden interpretar como diferencias en sus estándares individuales y puntos de referencia que usan en sus respuestas. Bajo el supuesto de que los encuestados usan estos mismos estándares individuales cuando evalúan las viñetas hipotéticas y cuando se evalúan a sí mismos, entonces se pueden usar las evaluaciones de estas viñetas para ajustar sus respuestas y corregir sesgos. Para más detalles sobre este método y sus aplicaciones en el estudio PISA en España véase Zamarro (2021).

### 2.2. Medidas directas de comportamiento

Dadas las limitaciones de las medidas basadas en encuestas, los investigadores han buscado medidas alternativas para medir directamente comportamientos que estén relacionados con distintas habilidades no cognitivas. En algunos casos los investigadores utilizan datos administrativos ya disponibles relacionados con el comportamiento de los estudiantes en el sistema educativo. Otra alternativa consiste en medidas de conducta a través de experimentos donde se pide a los encuestados que realicen una tarea directa en el laboratorio y se intenta cuantificar el rendimiento durante la tarea como medida de una cierta habilidad no cognitiva.

#### *2.2.1. Usando datos administrativos*

En algunos casos, educadores e investigadores, han usado datos administrativos para medir comportamientos relacionados con las habilidades no cognitivas de los

estudiantes. Distintos comportamientos académicos como las ausencias, las notas del curso, las incidencias de problemas de comportamiento, y expulsiones escolares están relacionados con la motivación y el esfuerzo académico de los estudiantes y pueden ayudar a construir sistemas de indicadores de alerta temprana para el posible abandono escolar. Estos comportamientos de manera habitual manifiestan bajos niveles de autocontrol, conciencia, o perseverancia a largo plazo o "Grit", que están asociados con una mayor probabilidad de abandono escolar. Para una mayor discusión del problema del abandono escolar, véase el capítulo 5b.

Un problema importante de las medidas de habilidades no cognitivas basadas en datos administrativos es que estos comportamientos están solamente determinados parcialmente por las habilidades no cognitivas. Existen otros factores que pueden estar asociados con estos comportamientos escolares, independientemente de los niveles de habilidades no cognitivas. Por ejemplo, aunque las notas de curso pueden reflejar bajos valores de autocontrol, conciencia, o perseverancia, también pueden reflejar otros factores familiares y de disponibilidad de recursos de manera independiente. Además, estos factores familiares y personales pueden a su vez estar relacionados con las habilidades no cognitivas tal y como se describe en el gráfico 1 arriba, lo que a su vez puede afectar la autoestima de los estudiantes y su rendimiento. Esto puede resultar en problemas importantes de endogeneidad si se está interesado en estudiar el efecto causal de las habilidades no cognitivas en otras variables. Para más discusión de este tipo de problemas empíricos, véase el capítulo 2 de este libro.

### *2.2.2. Medidas de conducta a través de experimentos con tareas directas*

Dadas las limitaciones de las medidas basadas en datos administrativos y encuestas, los investigadores han desarrollado medidas directas de habilidades no cognitivas a través de experimentos y tareas directas en laboratorios. Con estas medidas en laboratorio se intenta controlar lo más posible las condiciones en las que se recogen los datos para así limitar la influencia de otros factores externos en las medidas de habilidades no cognitivas. Estas medidas se basan en pedir a los encuestados que realicen una tarea determinada que se relacione con una habilidad no cognitiva particular. Los investigadores entonces cuantifican el rendimiento de los encuestados en la tarea.

Un ejemplo clásico en esta categoría es la famosa prueba de la golosina desarrollada por Mischel et al. (1972). Esta prueba se desarrolló para estudiar la ca-

pacidad de los niños de entre cuatro y seis años para postergar las recompensas. Los investigadores trabajaron y recogieron resultados de niños que estudiaban prescolar en una escuela que era parte de la Universidad de Stanford. En una serie de pruebas, los investigadores pedían a los niños que esperaran para comerse una golosina y alteraban las circunstancias de espera. Si los niños conseguían esperar entonces se les dan dos golosinas. Los investigadores midieron el número de minutos y segundos que los niños conseguían esperar como una medida de su habilidad para postergar recompensas, una habilidad relacionada con el autocontrol. Mischel y sus colegas concluyeron que los niños podían aumentar el tiempo de espera con estrategias de distracción. Los niños que conseguían esperar más se distraían mirando a otro lado, se cubrían los ojos, o hablaban consigo mismos, para evitar centrar la atención en la golosina. En el tercer apartado analizamos en mayor detalle la prueba de la golosina y sus implicaciones.

Este tipo de medidas a través de tareas directas tienen otras limitaciones importantes. Tienden a ser costosas, lo que limita la recolección de muestras grandes de datos, aunque esta es una limitación que está mejorando mucho con el uso de nuevas tecnologías. Segundo, no está claro cómo de representativos son los resultados que se obtienen en este tipo de pruebas realizadas en el contexto de laboratorio para un contexto más general del día a día (validez externa). Finalmente, es difícil realizar este tipo de pruebas múltiples veces para estudiar la evolución de las habilidades no cognitivas, dado que los participantes ganan familiaridad con la prueba la primera vez que la realizan lo que hace que los resultados estén sesgados al alza en pruebas siguientes.

Dados los desafíos para poder medir distintas habilidades no cognitivas de manera rigurosa, los investigadores han continuado estudiando posibles medidas alternativas. De manera más reciente se ha propuesto el uso de metadatos, eso es "datos sobre los datos", de exámenes y encuestas para obtener medidas de habilidades no cognitivas.

### 2.3. Usando metadatos para inferir medidas de habilidades no cognitivas

Tomar un examen o encuesta de manera exitosa requiere no sólo el conocimiento necesario de la materia sino también la autoeficacia, la conciencia y la autogestión necesaria para mantenerse centrado en la tarea. Por lo tanto, medidas basadas en los metadatos de exámenes y encuestas tales como el tiempo que se dedica a contestar las preguntas difíciles, si se responde rápidamente a las

preguntas sin dedicar un mínimo tiempo de respuesta, si se dejan preguntas en blanco, o si se contesta de manera aleatoria pueden ayudar a medir habilidades no cognitivas relacionadas con la autoeficacia, la conciencia, la autogestión, o la perseverancia.

Estas medidas basadas en metadatos se presentan como el equivalente a un experimento natural en economía, los metadatos no están diseñados para ser medidas directas de habilidades no cognitivas (tal y como un experimento natural no está diseñado como un experimento de laboratorio), pero la información que capturan de manera natural puede usarse para indicar distintos niveles de habilidades no cognitivas relevantes. El uso de metadatos de exámenes o encuestas se puede interpretar como un método intermedio entre el uso de datos administrativos y las medidas de conducta con experimentos en tareas directas. Aunque los metadatos que se suelen usar en este caso provienen de exámenes y encuestas que los estudiantes toman en el contexto del centro educativo, las condiciones en las que estas pruebas normalmente se desarrollan están más estandarizadas que los comportamientos que constituyen los datos administrativos en el curso de un día normal en el colegio o instituto.

Con respecto a los metadatos de exámenes, la investigación emergente ha mostrado que el tiempo que los estudiantes dedican a responder a las preguntas y en particular la tasa de respuesta rápida de preguntas ("*rapid guessing*"), sirve para aproximar los niveles de motivación académica del alumnado, lo que se relaciona con otras medidas de habilidades no cognitivas como la autoeficacia, autogestión, diligencia y la perseverancia. En este sentido, la investigación también ha mostrado cómo los estudiantes que no dedican suficiente tiempo a responder a las preguntas de un examen e intentan adivinar de manera rápida el 10% de las preguntas o más, presentan notas más bajas de curso, mayores ausencias escolares y problemas de conducta en el colegio.

Además del tiempo de respuesta, los investigadores han estudiado otros comportamientos cuando los estudiantes realizan un examen. Por ejemplo, Borghans y Schils (2012) proponen estudiar la tasa de decaimiento en el rendimiento a lo largo de pruebas estandarizadas como la prueba PISA desarrollada por la OCDE. Esto es, estudian como la probabilidad de responder de manera correcta a las preguntas de examen baja según el orden de las preguntas, del principio al final del examen. Usando datos de PISA, donde el orden de las preguntas varía en distintas versiones de la prueba y donde la versión de la prueba se asigna

a cada estudiante de manera aleatoria, lo que permite controlar por los niveles de dificultad de las preguntas, muestran como la tasa de decaimiento en la prueba no está muy relacionada con la puntuación final, lo que indica que el decaimiento no captura en gran parte las habilidades cognitivas de los estudiantes. Por otro lado, usando datos de exámenes con características similares a la prueba PISA, pero desarrollados en Inglaterra, Borghans y Schils (2012) muestran cómo los estudiantes de 16 años que presentaban una mayor caída del rendimiento a lo largo del examen presentaban niveles más bajos de conciencia, ambición y motivación escolar, así como una menor tasa de empleo y salarios en la edad adulta, 20 años más tarde, incluso cuando se tenía en cuenta la puntuación total en la prueba. Con estos resultados los autores concluyen que la tasa de decaimiento puede servir como una medida de habilidades no cognitivas relacionadas con la motivación intrínseca y la perseverancia.

### *2.3.1. El poder predictivo del esfuerzo de los estudiantes en encuestas*

Además de los exámenes, los estudiantes también completan de manera creciente encuestas en los centros educativos para recoger su opinión sobre distintos temas como su percepción sobre su centro y profesores, o para recoger datos sobre su situación económica y social. Estas encuestas, como era el caso de los exámenes estandarizados, también requieren que los estudiantes presten atención y se esfuercen en contestar las preguntas. Los investigadores también han propuesto medir el esfuerzo que los estudiantes ponen en estas encuestas en los colegios e institutos como una posible medida de habilidades no cognitivas relacionadas con la conciencia, la motivación y la diligencia. En particular, se propone estudiar dos tipos de comportamiento que demuestran poco esfuerzo por parte del estudiante: el porcentaje de preguntas de la encuesta que dejan en blanco y hasta qué punto los estudiantes responden de manera impredecible o aleatoria. Estos dos comportamientos se han encontrado relacionados con medidas de habilidades no cognitivas basadas en encuestas como la conciencia y la perseverancia a largo plazo o "Grit". Además, también se ha encontrado que la falta de esfuerzo en encuestas por parte de los estudiantes está correlacionada con los niveles de educación que se alcanzan, empleo y salarios más tarde en la vida, incluso cuando se tienen en cuenta los niveles de capacidad cognitiva y nivel socioeconómico de la familia.

Por ejemplo, Hitt et al. (2016) usan datos de seis estudios longitudinales representativos de la población adolescente de EE. UU. Todos los participantes en

estos estudios realizaron una prueba de habilidades cognitivas en el primer año de participación y encuestas recogiendo datos demográficos y su opinión sobre distintos temas. Los autores argumentan que, para los estudiantes, estas encuestas que se hacen en el colegio se perciben como deberes y, por tanto, se pueden usar para medir la motivación académica de los estudiantes a través del esfuerzo que ponen en estas encuestas. Los autores, entonces miden el porcentaje de preguntas de la encuesta que los estudiantes dejan en blanco o responden "no lo sé", en las encuestas de cada uno de los estudios. Usando un modelo de regresión lineal, los autores estudian la relación entre el esfuerzo en la encuesta en la adolescencia y el nivel de estudios alcanzado ya en la edad adulta (para más discusión sobre modelos de regresión lineal, véase el capítulo 2 de este libro). Los autores controlan en su modelo empírico por características familiares y personales que se relacionan con los costes y beneficios de adquirir un mayor nivel educativo. De manera importante, los autores también controlan por el nivel de habilidad cognitiva de los encuestados. Los autores encuentran que cuando se incluye tanto la tasa de no respuesta en la encuesta como el nivel de habilidades cognitivas, la tasa de no respuesta resulta ser un predictor estadísticamente significativo en cuatro de los seis estudios. Los efectos, cuando son significativos, varían entre 0,10 y 0,30 años menos de educación por cada incremento de una desviación estándar en la tasa de no respuesta. Este efecto compara favorablemente con el efecto de aumentar la habilidad cognitiva en una desviación estándar, lo que corresponde con un incremento de entre 0,10 y 1,44 años más de educación, dependiendo del estudio. Los autores también estudian la relación entre la tasa de no respuesta y los niveles de renta y empleo en la edad adulta y encuentran que, a pesar de observarse una relación inicial significativa, la relación se debe en gran parte a los mayores niveles de educación alcanzados.

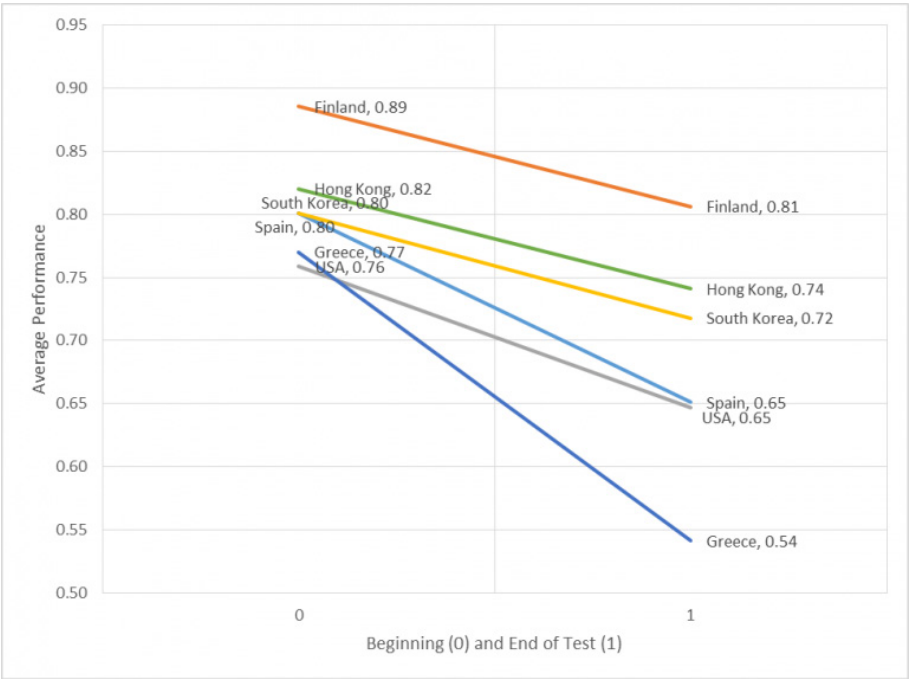
### *2.3.2. Diferencias internacionales, entre Comunidades Autónomas y por género*

La investigación en economía de la educación ha llegado a un entendimiento creciente de que los resultados de los estudiantes en pruebas estandarizadas, como es el caso de la prueba PISA, capturan tanto el conocimiento de los estudiantes en la materia, como el esfuerzo y motivación académica del alumnado. Este hecho se corrobora dados los resultados de trabajos como el de Gneezy et al. (2019) que muestran como los resultados en este tipo de pruebas se pueden alterar ofreciendo a los estudiantes incentivos monetarios. Como resultado, los investigadores han usado metadatos en este tipo de pruebas para estudiar has-

ta qué punto las diferencias entre países, regiones o grupos de estudiantes (por ejemplo, por género) se pueden explicar por distintos niveles de motivación y esfuerzo de los estudiantes (el lector interesado puede consultar también el capítulo 5c para más detalle sobre como diferencias en habilidades no cognitivas por género, impactan algunas dimensiones educativas).

En su análisis de la tasa de decaimiento del rendimiento usando datos de la prueba PISA, Broghans y Schils (2012) documentan diferencias importantes entre países en esta medida de esfuerzo y estiman como alrededor del 19% de la variación observada en la prueba PISA entre países podría explicarse por una distinta tasa de decaimiento del rendimiento a lo largo de la prueba. Zamarro et al. (2019) corroboran estos resultados usando tanto la tasa de decaimiento del rendimiento a lo largo de la prueba como otras medidas de esfuerzo basadas en metadatos de la encuesta que sigue a la prueba PISA, tales como la tasa de preguntas que los estudiantes dejan en blanco y la tasa de respuestas que se pueden considerar impredecibles o aleatorias. Zamarro et al. (2019) estiman que todas estas medidas de esfuerzo podrían explicar entre el 32% y el 38% de la variación

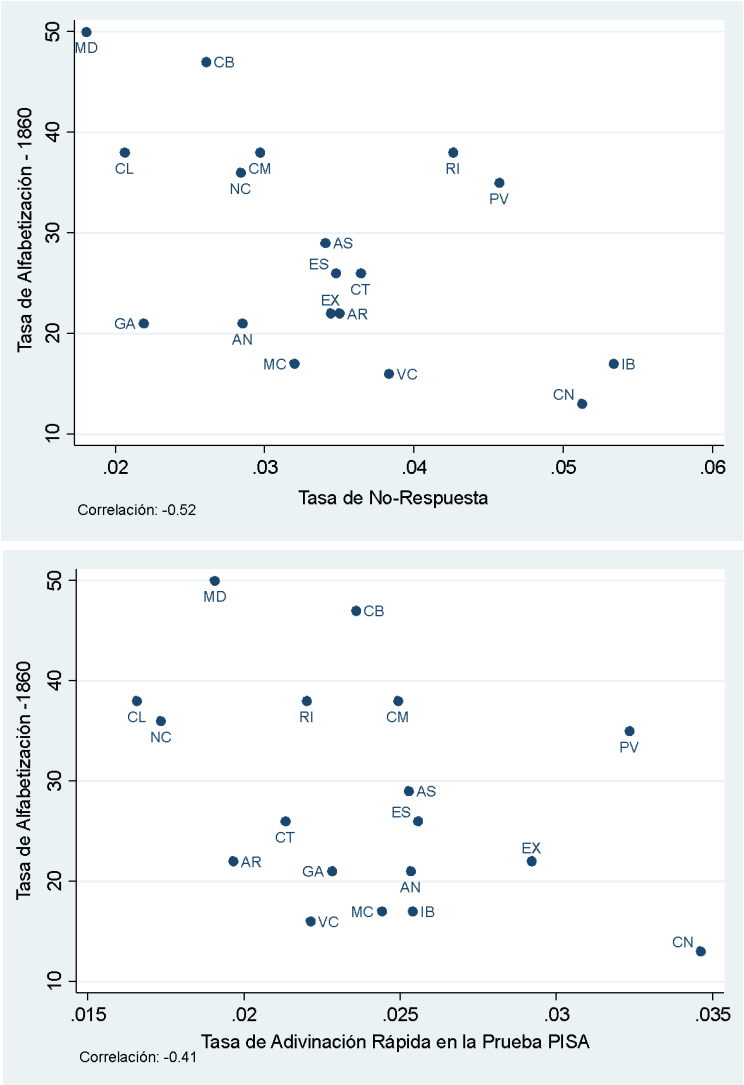
**Gráfico 2. Rendimiento al principio y al final del examen PISA, por país.**




Fuente: Zamarro, Hitt y Mendez (2019).

observada entre países en la puntuación de la prueba PISA. Tal y como revela el gráfico 2, ambos trabajos encuentran que, junto a Grecia, España es uno de los países que muestra una mayor caída del rendimiento a lo largo de la prueba, ya que la línea que muestra la caída del rendimiento desde el principio al final de la

**Gráfico 3. Tasa de no-respuesta en la encuesta PISA, tasa de respuesta rápida en la prueba y medidas históricas de alfabetización por regiones de España.**




Fuente: Zamarro (2021).

prueba tiene una mayor pendiente en estos países, lo que sugiere un déficit en las habilidades no cognitivas relacionadas con la diligencia y la persistencia.

Balart y Cabrales (2014) muestran diferencias importantes en la tasa de decaimiento en la prueba PISA también entre Comunidades Autónomas en España, pero concluyen que estas diferencias se pueden explicar en gran parte por las diferencias regionales en la composición de las características socioeconómicas de los estudiantes. Zamarro (2021) además señala cómo puede existir también un componente cultural que explique las diferencias entre Comunidades Autónomas dadas las correlaciones encontradas entre la tasa de no respuesta en el cuestionario PISA y la tasa de respuesta rápida de preguntas en la prueba, con las tasas históricas de alfabetización en España (véase gráfico 3).

La literatura también ha mostrado como, de manera consistente, las chicas presentan mayores niveles de esfuerzo en este tipo de pruebas que los chicos. Balart y Cabrales (2014) y Balart y Oosterveen (2019) muestran como ser chica y tener una mayor proporción de chicas en clase están asociados con un menor decaimiento en el rendimiento en la prueba PISA. De manera similar, Zamarro (2021) también muestra un mayor esfuerzo por parte de las chicas en España en medidas de motivación académica, la tasa de preguntas que los estudiantes dejan en blanco en la encuesta PISA y la tasa rápida de respuesta en la prueba. Este mayor esfuerzo por parte de las chicas supone que ellas compensan parte de la diferencia en rendimiento en la prueba con un mayor esfuerzo. Como resultado, una vez que se corrige por diferencias de género en esfuerzo usando métodos estadísticos, se encuentra que las brechas de género en el rendimiento en asignaturas como Matemáticas y Ciencias podrían ser mayores de lo que se creía inicialmente. Para más detalle sobre las diferencias de género en educación, se puede consultar el capítulo 5c.

## 3. ¡NO TE COMAS LA GOLOSINA!

En la década de los 60 del siglo pasado se realizó uno de los experimentos más conocidos y replicados de la historia de la psicología. Walter Mischel, psicólogo austríaco de la Universidad de Stanford, ideó una original prueba de postergación de recompensas dirigida a niños. La idea era sencilla. Un niño de 5-6 años que asiste a la guardería de la Universidad de Stanford es acompañado a una sala sin estímulos o distracciones y se le pide que se siente junto a una mesa en la que hay un plato con una golosina. Entonces se le informa de que tiene dos opciones.

Puede comerse la golosina en cualquier momento o bien puede esperar unos minutos concretos, que puede gestionar con un reloj analógico, y entonces tendrá dos golosinas.

Dado que para un niño de esa edad una golosina es un reforzador potente, un objeto de deseo, la prueba de la golosina es una lucha entre emoción y razón. Lo interesante de esta batalla es la disparidad de fuerzas, ya que a los 5-6 años las áreas del cerebro que gestionan las emociones, el sistema límbico y la amígdala, están maduras, mientras que la corteza prefrontal, la última capa evolutiva del cerebro humano, desde donde se gestionan las funciones ejecutivas y, por ende, la racionalidad, no madura hasta los 20-25 años. Así, pues, en esta batalla entre Goliat, el ahora, el carpe diem, la emoción, y un David que representa la racionalidad, la postergación, la inhibición del impulso de comerse ya la golosina, lo normal es que gane Goliat. Mischel no sólo tomó nota del tiempo que tardaba en comerse la golosina cada niño, sino también de las estrategias que utilizaban aquellos que tenían más éxito en la inhibición del deseo de comer la golosina cuanto antes.

La sorpresa llegó cuando, al realizar un seguimiento de los entonces niños que realizaron el test original, Mischel y coautores encontraron que, a igualdad de otras características del entorno familiar, aquellos niños que habían resistido mejor la tentación de comerse la golosina a los 5-6 años de edad tenían, en general, mejor rendimiento escolar, trabajos mejor remunerados, mayor nivel de satisfacción con su vida, menor índice de masa corporal y soportaban mejor el estrés que aquellos que habían sucumbido a la tentación antes. Este resultado, publicado en la última década del pasado siglo, generó un profundo debate sobre el papel que la educación del carácter, como lo denominan algunos en referencia a la parte moldeable de la personalidad, las habilidades de autorregulación o socioemocionales, como enfatizan otros, tiene en el bienestar a lo largo de la vida.

Es en ese contexto en el que surgen estudios, desde la economía, la psicología y la combinación de ambas disciplinas, que analizan el poder explicativo de estas dimensiones moldeables de la personalidad o habilidades en el desarrollo humano y en la determinación de la desigualdad en la sociedad. La aplicación de la modelización económica y de estrategias de identificación de efectos causales a medidas disponibles de estas habilidades ha destacado el notable papel que estas juegan en el rendimiento escolar, el abandono educativo temprano, la estabilidad personal y laboral y la salud. Asimismo, se ha demostrado que estas habilidades, independientes del cociente intelectual, se transmiten de padres a hijos

tanto por la vía genética como por el ejemplo. Las inversiones en programas de mejora temprana de las habilidades no cognitivas de colectivos socialmente desfavorecidos, por sus efectos en diversas dimensiones del desarrollo de los participantes y de sus hijos, son una de las líneas de actuación más prometedoras para la reducción de las desigualdades sociales. En el siguiente apartado revisamos la evidencia disponible.

En 2018, Watts, Duncan y Quan se plantearon replicar los resultados de los estudios originales de Mischel y sus colegas, pero usando una muestra más grande y representativa de los Estados Unidos. Watts et al. (2018) corroboraron correlaciones positivas entre el tiempo de espera de los niños en su muestra y el rendimiento académico. Sin embargo, la asociación resultó bastante menor que la que originalmente documentaron los trabajos de Mischel y coautores y desaparecía cuando se comparaban niños del mismo nivel socioeconómico y desarrollo cognitivo. Watts et al. (2018) tampoco encontraron una relación en su estudio entre el tiempo de espera de los niños a los 5 años y sus medidas socioemocionales de comportamiento a los 15 años. Estos nuevos resultados en una muestra más representativa demostraron la importancia de los factores socioeconómicos y culturales en los niveles de habilidades no cognitivas a pesar de que la prueba de la golosina estaba diseñada para limitar la influencia de este tipo de factores.

## 4. DETERMINANTES DE LAS HABILIDADES NO COGNITIVAS

### 4.1. Genética

La investigación realizada en las últimas décadas ha demostrado que los atributos, disposiciones, actitudes y recursos que dan contenido a las habilidades no cognitivas se van configurando desde la más tierna infancia, e incluso antes del nacimiento, como resultado de la interacción entre factores genéticos y ambientales. Buser et al. (2023), por ejemplo, demuestran que los genes juegan un papel relevante en la determinación de nuestra dotación de habilidades cognitivas y no cognitivas y, por tanto, explican parcialmente las desigualdades en logro educativo y laboral que, a su vez, determinan la desigualdad en la distribución de la renta, la riqueza y el bienestar en una sociedad.

El papel de los genes en la determinación de las habilidades no cognitivas está particularmente documentado en población clínica, esto es, de personas con una dificultad de aprendizaje, como el trastorno por déficit de atención con o

sin hiperactividad (TDA/H), y/o personas con necesidades educativas especiales como, por ejemplo, el autismo. Así, el factor de riesgo más importante para tener un hijo/a con autismo es el tener uno o más familiares con autismo. En este escenario, la probabilidad de tener un hijo con autismo se sitúa en torno al 20%, entre 10 y 20 veces más que en la población general. En el caso del TDAH, dicha probabilidad es del 80%.

Bouchard y Loehlin (2001) concluyen que, en la población general, los indicadores de habilidades no cognitivas obtenidos a partir de información sobre el comportamiento o la conducta de la generación *n* se transmiten a los miembros de la generación *n+1* en un porcentaje que oscila entre el 30% y el 60%. Cabe, pues, concluir que, si bien los genes tienen un papel relevante en la determinación de las habilidades no cognitivas, los factores ambientales juegan, para la mayoría de los grupos o colectivos sociales, sobre todo aquellos que se encuentran en riesgo de exclusión social, un papel aún más relevante, afortunadamente, en la determinación y evolución de dichas habilidades.

### 4.2. Factores ambientales y renta familiar

Entre los factores ambientales destacan, de acuerdo con la investigación, el estrés, el estilo de crianza de los padres o cuidadores habituales (del que también se habla en el capítulo 3b de este libro), el tamaño de la familia y un factor que podríamos definir, en un sentido general, como acompañamiento, que refleja el tiempo de calidad que, de forma individualizada, disfruta el menor con sus padres, cuidadores habituales o maestros.

La renta familiar se ha utilizado en economía como proxy de varios factores cuya presencia deprime el desarrollo de las habilidades no cognitivas: el estrés, la inestabilidad emocional de los padres, su menor probabilidad de acompañar al menor en su desarrollo y de pasar tiempo de calidad con él, o la exposición a entornos poco seguros o violentos. Fletcher y Wolfe (2016) encuentran, utilizando una muestra de escolares residentes en Estados Unidos con edades comprendidas entre los 5 y los 11 años, que la renta familiar tiene un destacado y creciente poder explicativo en el ajuste socioemocional de los niños. En concreto, estos autores encuentran que la probabilidad de que un estudiante muestre habilidades sociales adecuadas, como cooperar con compañeros, o evite problemas causados por su impulsividad, como la agresión a un compañero, está condicionada causalmente por la renta familiar.

En este sentido, la evidencia revisada en Diamond (2016) confirma que factores tales como crecer en un hogar en riesgo de exclusión social, ser miembro de una minoría socialmente discriminada o la exposición a la violencia o a la falta de seguridad, tienen un efecto muy nocivo sobre el desarrollo de las funciones ejecutivas del cerebro. En concreto, los niños que viven en hogares en riesgo de exclusión social muestran un funcionamiento ejecutivo promedio a la edad de 4 años, esto es, al comienzo de su vida escolar, un 30% inferior al de los niños no expuestos a la pobreza. Este diferencial adverso se traduce en una menor probabilidad de mantener la atención en clase, de transitar de forma flexible entre actividades, o de relacionarse de forma socialmente válida con los compañeros.

La evidencia contenida en Farah et al. (2006) sugiere que esas diferencias basadas en la renta familiar se incrementan con la edad. En concreto, los adolescentes de entre 10 y 13 años que viven en hogares en riesgo de exclusión social muestran, en promedio, una memoria de trabajo un 50% inferior a la de los estudiantes que viven en hogares de clase media. El diferencial desfavorable a esas edades alcanza el 60% en el caso de la inhibición conductual, el componente de la función ejecutiva más estrechamente relacionado con el concepto económico de habilidades no cognitivas.

Diamond (2016) señala al estrés que caracteriza a los hogares en riesgo de exclusión social como el factor determinante en la relación causal entre la renta familiar y el desarrollo de las habilidades no cognitivas. Abundante evidencia en el campo de la psicología y la medicina destaca el efecto pernicioso de la exposición prolongada al estrés en el desarrollo de la corteza prefrontal y, por ende, del funcionamiento ejecutivo. Asimismo, resulta evidente que los hogares menos favorecidos en términos de renta son hogares donde hay mayor estrés, donde lo urgente se impone permanentemente a lo importante, proporcionando un contexto menos favorable a la estimulación de la reflexión en los menores, auténtico determinante de un funcionamiento ejecutivo adecuado (Zelazo, 2003).

### 4.3. Estilos de crianza

En relación al estilo de crianza, Diamond (2016) señala que los hijos cuyos padres son muy protectores o controladores, y aquellos que basan la educación de sus hijos en el castigo, muestran sistemáticamente niveles inferiores de funcionamiento ejecutivo que aquellos con padres que favorecen el desarrollo de la

autonomía de sus hijos. En general, el nivel de autorregulación de los niños es más elevado en aquellos hogares en los que se utiliza un lenguaje más enriquecido, en los que los padres plantean preguntas abiertas que fomentan la reflexión de sus hijos y moldean en voz alta sus estrategias de resolución de problemas. Estos factores también guardan una estrecha relación con la renta familiar como indicador sintético de contexto.

Otro factor relevante para el desarrollo de las habilidades no cognitivas de un niño es quién es su cuidador habitual. Este factor está estrechamente relacionado con el anterior, el estilo de crianza de los niños. Ao et al. (2021) encuentran que aquellos menores residentes en China que son cuidados habitualmente por sus abuelos muestran niveles sistemáticamente superiores de *locus of control* externo que aquellos que son cuidados por sus padres. Esto es, los menores que son cuidados habitualmente por sus abuelos son, a igualdad de otras características como la renta familiar, más proclives a pensar que el éxito o la mejora se deben a factores ajenos a ellos mismos como, por ejemplo, la suerte, el destino o las condiciones familiares iniciales. Este resultado se explicaría por el diferente estilo de crianza de padres y abuelos. Estos últimos, en concreto, muestran un estilo más tendente al disfrute de los nietos, al ocio compartido, con menos límites, disciplina, implicación en las tareas académicas y, en general, una menor estimulación de la autonomía de los nietos.

Este último resultado sugiere que ser cuidado habitualmente por los abuelos puede conducir, a igualdad de otras características como la renta familiar, a una menor mentalidad de crecimiento en los niños. Este concepto, originalmente desarrollado por la psicóloga Carol Dweck, se ha mostrado como una de las dimensiones no cognitivas con mayores implicaciones en el rendimiento escolar y en la autoestima de los menores (Boneva y Ertac, 2016).

En concreto, podemos decir que un niño tiene mentalidad de crecimiento cuando es consciente de que es el esfuerzo continuado, la perseverancia, la que permite mejorar los resultados que obtenemos. La mentalidad de crecimiento permite a los niños comprender que los errores son parte intrínseca del proceso de aprendizaje y, por tanto, son valiosos. Ante todo, una dosis adecuada de mentalidad de crecimiento preserva su autoconcepto al hacerles conscientes de que los resultados mejoran con práctica y esfuerzo, y no se explican por dotaciones estáticas de factores supuestamente inalterables como la inteligencia o las aptitudes físicas para el deporte.

La antítesis de la mentalidad de crecimiento es una conceptualización de la inteligencia y otros factores personales como estáticos, insensibles al esfuerzo y a la práctica. Así, los niños con mentalidad fija suelen tener miedo a equivocarse, no quieren ponerse a prueba en nuevos contextos por miedo a no hacerlo igual de bien y, en general, suelen conjugar con frecuencia el verbo ser en frases del estilo "yo no soy bueno en esto" poco después de haber comenzado con *esto*.

El resultado obtenido en Ao et al. (2021) sugiere que países como España, en los que la conciliación de la vida profesional y familiar depende de los apoyos informales proporcionados en el entorno familiar y, muy en concreto, por los abuelos, alcanzan dotaciones de habilidades no cognitivas subóptimas en los jóvenes como resultado de esa forma de organización social.

### 4.4. Tamaño de la familia

El número de miembros de la unidad familiar es otro de los factores señalados en la literatura como determinante de la dotación de habilidades no cognitivas de una persona. Juhn y otros (2015) demuestran, utilizando una muestra representativa de la población de Estados Unidos, que el efecto sustitución entre número de hijos (cantidad) e inversión por hijo (calidad) también afecta a las habilidades no cognitivas. En concreto, estos autores encuentran que un incremento en el tamaño de la familia por el nacimiento de un nuevo miembro conduce a una menor inversión parental por niño, reduciendo tanto las habilidades cognitivas como las no cognitivas, aproximadas por las dificultades conductuales, de los menores.

El concepto de inversión parental no hace referencia únicamente a la inversión monetaria en centros educativos, academias, etc., sino también a la inversión de tiempo de calidad con cada uno de los menores. Este último aspecto nos conduce directamente a otro de los factores determinantes destacados en la literatura, que podemos definir como acompañamiento. El acompañamiento hace referencia al tiempo que un adulto (padre, abuelo o maestro) dedica de forma individual a un menor. La importancia de sentirte parte del grupo, sea éste la familia o el aula, ha sido destacada en numerosas publicaciones como un determinante de la motivación y el esfuerzo que un menor pone en lo que hace. Asimismo, disponer de un tiempo individualizado con los adultos de referencia de cada contexto es básico para sentir que uno pertenece a dicho contexto.

### 4.5. Entorno escolar

La relevancia del acompañamiento en la determinación de las habilidades no cognitivas se pone de manifiesto, por ejemplo, en las investigaciones que señalan que los grupos escolares más reducidos tienen mejores resultados en habilidades no cognitivas (Dee y West, 2008, véase el capítulo 3a del tamaño de la clase en este mismo libro), o en aquellas que muestran que el cuidado de los abuelos ofrece mejores resultados en habilidades no cognitivas que las alternativas formales como las guarderías por el trato individualizado que obtienen los nietos (Morando y Platt, 2022). La evidencia sobre el tamaño de la unidad familiar también apunta en este sentido.

Los maestros también juegan un papel muy destacado en la determinación de las habilidades no cognitivas de sus alumnos. En este sentido apuntan los resultados obtenidos en Cheng y Zamarro (2018) y Jackson (2018). Este último estudio demuestra que los profesores tienen efecto no sólo en el rendimiento escolar de sus alumnos, sino también en un amplio abanico de conductas como absentismo, expulsiones, repetición, etc. En concreto, Jackson (2018) encuentra que teniendo en cuenta el efecto de los profesores tanto en el rendimiento como en la conducta de sus alumnos prácticamente se dobla el impacto que los profesores tienen en el desarrollo global a largo plazo de sus estudiantes. Por su parte, Cheng y Zamarro (2018) desarrollan medidas de habilidades no cognitivas de los profesores que proporcionan información no recogida en indicadores habitualmente considerados, al tiempo que demuestran el impacto que tienen en el rendimiento de sus alumnos. Para más información sobre el papel del profesorado en el desarrollo de las habilidades cognitivas y no cognitivas de sus estudiantes (véase el capítulo 3d).

### 4.6. Cultura

La cultura es, sin duda, otro factor determinante de la dotación de habilidades no cognitivas de las cohortes jóvenes de un país. Fernández y Fogli (2009) definen la cultura como el conjunto de creencias y preferencias que condiciona las decisiones individuales, que varían de forma sistemática entre grupos definidos por su ubicación geográfica y/o posición social, y que se transmite de padres a hijos. La prolífica literatura sobre el impacto de la cultura así definida en la sociedad ha mostrado su relevancia en decisiones de fertilidad, oferta de trabajo, movilidad geográfica, formación de nuevos hogares, regulación del mercado de

trabajo, etc. En este sentido, Méndez (2015) demuestra que una fracción destacada, entre el 10% y el 30%, de las diferencias entre países desarrollados en rendimiento de sus escolares de 15 años de edad en la prueba estandarizada PISA se debe al efecto de la cultura en la determinación de las habilidades no cognitivas dominantes en cada país. En concreto, Méndez (2015) analiza en qué medida el rendimiento en PISA de estudiantes de segunda generación de inmigración con diferentes orígenes pero que residen en un mismo país se ve influido por las cualidades que se consideraba importante fomentar en un niño en el país de origen de su padre cuando éste era joven. Así, aquellos jóvenes cuyos padres vivían en países que consideraban esencial fomentar el esfuerzo, la perseverancia, la autonomía y la postergación de recompensas obtienen, a igualdad de otras características personales, del centro escolar, institucionales y de ciclo, mejores resultados en Matemáticas, Ciencias y Lengua. Por su parte, Méndez y Zamarro (2018) encuentran que este factor cultural también tiene un destacado efecto en el nivel educativo finalmente alcanzado.

Cabe señalar que la evidencia disponible también demuestra que los factores ambientales tienen impacto en el desarrollo de las habilidades no cognitivas del menor desde su concepción. En este sentido, Chang y otros (2020) encuentran que la frecuencia e intensidad de los episodios de lluvias torrenciales acaecidos durante el embarazo tienen impacto en el desarrollo cognitivo y no cognitivo de los menores de la India. En concreto, los episodios que tienen lugar en el primer trimestre de embarazo tienen un mayor impacto negativo en el desarrollo del niño.

## 5. INTERVENCIONES EFICACES EN LA MEJORA DE LAS HABILIDADES NO COGNITIVAS

Son numerosas las intervenciones que se han diseñado con el objetivo de mejorar las habilidades no cognitivas de grupos concretos de población, normalmente definidos por la presencia de un déficit en dichas habilidades, o de la población general.

Heckman y Kautz (2013) realizan una exhaustiva revisión de la literatura sobre evaluación de intervenciones diseñadas para mejorar el bienestar de determinados grupos sociales, normalmente colectivos desfavorecidos. Estos autores revisan más de 20 programas cuyas evaluaciones cumplen con los requisitos de identificar el efecto de interés de forma metodológicamente adecuada, testar la

presencia de dicho efecto varios años después de la intervención e identificar el impacto del mismo en medidas de habilidades no cognitivas y en diferentes dimensiones del desarrollo como, por ejemplo: cociente intelectual, rendimiento escolar, nivel educativo alcanzado, salud, criminalidad, abuso de sustancias estupefacientes, situación laboral, ingresos laborales, etc.. Todos los programas analizados tienen una duración mínima de un año, siendo dos años la duración más frecuente de los programas dirigidos a hogares con niños menores de 4 años.

El análisis realizado en Heckman y Kautz (2013) constata que las intervenciones más eficaces, esto es, aquellas que mejoran el bienestar de la población objetivo en mayor medida y cuyos efectos aparecen más persistentes en el tiempo, son aquellas que van dirigidas a hogares desfavorecidos en los que hay menores en edad preescolar. En concreto, las intervenciones más coste-eficaces son aquellas dirigidas a hogares en los que hay niños menores de 5 años. Estos menores lograron, gracias a los programas evaluados, mejorar tanto su rendimiento escolar como el nivel educativo finalmente alcanzado en la edad adulta, su salud, sus ingresos laborales y reducir la probabilidad de verse inmersos en problemas con la justicia respecto de lo que habría pasado de no haberse realizado el programa en sus hogares. Cabe, pues, afirmar que un programa eficaz lo será tanto más cuanto más jóvenes sean los menores del hogar al que va dirigido.

Heckman y Kautz (2013) encuentran que los únicos programas eficaces en la mejora sostenida en el tiempo del cociente intelectual son los que van dirigidos a hogares con niños menores de 3 años. Este resultado sugiere que el periodo crítico para la mejora de esta medida de inteligencia se extiende desde la concepción hasta, aproximadamente, los 3 años. A partir de esa edad, los factores determinantes del desarrollo cognitivo están, en buena medida, predeterminados. Por el contrario, las habilidades no cognitivas de una persona se pueden mejorar con intervenciones bien definidas tanto en su más tierna infancia, interviniendo con sus padres o cuidadores habituales, como a lo largo de la escolarización obligatoria, la adolescencia y la vida adulta de esa persona. Eso sí, la eficacia de esta intervención será tanto mayor cuanto menor sea la edad de esa persona.

Cunha et al. (2010) explican este último resultado al encontrar que las habilidades, cognitivas o no, muestran una marcada complementariedad dinámica. En otras palabras, los niveles de una habilidad en una edad determinan la productividad de inversiones futuras en habilidades. Estos autores encuentran que, a diferencia de las habilidades no cognitivas, las habilidades cognitivas sólo

son moldeables durante los primeros años de la vida de una persona. Aunque la complementariedad dinámica explica la mayor productividad de las inversiones en habilidades no cognitivas a edades más tempranas, inversiones en estas habilidades en edades posteriores todavía muestran una destacada rentabilidad. Finalmente, estos autores encuentran que las habilidades no cognitivas contribuyen a la mejora de las habilidades cognitivas, no siendo tan relevante la causalidad en sentido opuesto.

Estos resultados sugieren que las intervenciones más eficaces para la mejora de la calidad de vida de menores en hogares socialmente desfavorecidos serán aquellas que logren mejorar las habilidades no cognitivas de estos menores de forma sostenida en el tiempo. Los resultados analizados en Heckman y Kautz (2013) confirman esta intuición.

Otra característica común a los programas más eficaces es que son programas que implican a los padres, bien de forma obligatoria o bien porque logran motivarlos para participar. Este resultado es lógico si se tiene en cuenta que los programas más eficaces trabajan con familias con menores no escolarizados y, por tanto, la formación va dirigida a los padres que se convierten en un molde más adecuado para que, por imitación, sus hijos asimilen esos nuevos hábitos o habilidades no cognitivas. En concreto, algunos de los programas más eficaces, como los programas de preescolar Perry o Head Start, incluyen sesiones realizadas en el hogar objetivo de la intervención, para moldear directamente a los padres. La menor eficacia estimada para los programas dirigidos a adolescentes coincide con el hecho de que ninguno de estos programas implica a los padres de los adolescentes, ni realiza sesiones en el hogar.

Uno de los programas más evaluados y eficaces, el Programa preescolar Perry, iba dirigido a niños de 3 o 4 años, de familias de ingresos bajos, de raza negra y con un cociente intelectual inferior a 85 a la edad de 3 años. El programa ofrecía una intervención de dos horas y media de lunes a viernes en un centro para los niños y visitas al hogar para moldear a los padres y reforzar los vínculos con sus hijos. El Programa Perry tuvo una duración de dos años. Transcurrido este tiempo, tanto las familias aleatoriamente seleccionadas para el programa como las excluidas eran escolarizadas en el mismo centro.

Los resultados de la evaluación aleatorizada muestran que el Programa Perry mejoró el rendimiento escolar y redujo la probabilidad de que estos menores tuvieran problemas con la justicia años después, con efectos positivos, aunque

menos robustos, sobre el nivel educativo finalmente alcanzado y los ingresos laborales. La clave de esta eficacia se encuentra en la mejora de las habilidades no cognitivas de estos menores, ya que el programa sólo tuvo efectos transitorios en el cociente intelectual.

Los menores que participaron en el Programa Perry realizaban diariamente rutinas de pensamiento del tipo "planifica-ejecuta-revisa" en las que se les enseñaba a elaborar un plan, ejecutarlo y reflexionar sobre las desviaciones del plan original y su ejecución, así como sobre los resultados obtenidos. La reflexión la realizaban con maestros del centro y con sus compañeros de clase. Asimismo, se les acostumbró a trabajar con sus compañeros para encontrar soluciones a los problemas que iban encontrando en la consecución de sus objetivos.

Finalmente, otro aspecto a destacar de las intervenciones o programas que estimulan eficazmente el desarrollo de las habilidades no cognitivas es que, en algunos casos, como el del *Programa Perry*, los beneficios del programa se transmiten a la siguiente generación. Así, Heckman y Karapakula (2019) encuentran que los hijos de los participantes en el programa Perry se vieron notablemente beneficiados de la participación de sus padres en dicho programa. En concreto, estos niños tuvieron menos problemas por su conducta en el colegio, niveles más elevados de educación y empleo y menores tasas de participación en actividades delictivas que los hijos de los no participantes. Similares resultados se obtienen en García, Heckman y Ronda (2021) para los hijos de participantes afroamericanos residentes en hogares desfavorecidos. Asimismo, Heckman y Karapakula (2019) encuentran que los hermanos de los participantes en el programa también se vieron beneficiados de forma notable.

El Programa Perry bien podría definirse como un programa de estimulación del funcionamiento ejecutivo de los menores. En concreto, el objetivo del Programa Perry podría definirse como que los menores adquiriesen el hábito de pensar antes de hacer, esto es, estimular la inhibición. Para ello, probablemente con apoyos visuales, se les recordaba la necesidad de generar un plan, ejecutarlo y, finalmente, realizar una reflexión sobre las posibles desviaciones entre lo planificado y lo obtenido.

Asimismo, y tal y como sugiere Vygotsky en sus aportaciones seminales, la reflexión se realizaba entre compañeros, lo que contribuye a un aprendizaje más significativo y a la mejora del funcionamiento ejecutivo. Así, el Programa Perry generó el hábito de pensar antes de hacer, estimulando el funcionamiento ejecu-

tivo o habilidades no cognitivas de los alumnos participantes, algo que se infiere de los resultados de la evaluación del programa que indican que este mejoró la conducta de los menores.

Utilizando criterios similares a los de Heckman y Kautz (2013) sobre selección de las evaluaciones de intervenciones a analizar, Diamond y Ling (2016) ofrecen una revisión exhaustiva de la eficacia de programas diseñados para la mejora del funcionamiento ejecutivo.

Diamond y Ling (2016) revisan más de 80 intervenciones que van desde programas de ordenador para la estimulación de determinadas dimensiones del funcionamiento ejecutivo, como la memoria operativa, a intervenciones que constituyen auténticos currículos diseñados para la mejora del funcionamiento ejecutivo de la población escolar general, similares a los evaluados en Heckman y Kautz (2013), siendo algunos programas, como *Tools of the Mind*, incluidos en ambas revisiones.

En general, las intervenciones revisadas en Diamond y Ling (2016) tienen un tamaño muestral inferior a las analizadas en Heckman y Kautz (2013), al tiempo que realizan un menor seguimiento temporal del efecto de la intervención No obstante, los resultados son cualitativamente coherentes con los obtenidos en Heckman y Kautz (2013) y, ante todo, complementarios, al informar sobre detalles del diseño e implementación del programa, así como su desigual eficacia en alumnos de distinto perfil de neurodesarrollo, aspecto que pasa habitualmente desapercibido en las evaluaciones realizadas desde la economía.

Los resultados obtenidos en Diamond y Ling (2016) confirman que la estimulación de las funciones ejecutivas requiere de un tiempo de asimilación del nuevo hábito, esto es, de una exposición coherente a una secuencia estímulo-respuesta en un contexto determinado. En ese sentido, cabe destacar que los programas revisados en Heckman y Kautz (2013) tenían una duración mínima de un año, que se extendía hasta los dos años en el caso de los programas más eficaces.

Asimismo, Diamond y Ling (2016) concluyen que las intervenciones que se centran en una dimensión concreta del funcionamiento ejecutivo como, por ejemplo, la inhibición o la memoria operativa, no suelen transferir sus beneficios, si los hubiera, a otras dimensiones del funcionamiento ejecutivo, por lo que obtienen beneficios menores de los que podrían obtener con un enfoque más global. Este resultado sugiere que contemplar todas las dimensiones del

funcionamiento ejecutivo en el diseño del programa puede ser particularmente relevante para su eficacia.

El concepto de funciones ejecutivas tiene una ventaja destacada sobre el de habilidades no cognitivas, y es que ya está integrado en el mundo educativo. Así, mientras que tanto maestros (Infantil y Primaria) como profesores (Secundaria y posteriores etapas educativas) suelen ser ajenos a un concepto construido por exclusión del cociente intelectual como es el caso del término habilidades no cognitivas, casi todos ellos, sobre todo los maestros, están familiarizados con las funciones ejecutivas del cerebro y su relevancia en el desarrollo cognitivo, emocional y social de niños y adolescentes.

Las funciones ejecutivas han sido destacadas como determinantes para el aprendizaje y la adquisición de habilidades socioemocionales por parte del alumnado con dificultades de aprendizaje y/o necesidades educativas especiales. Así, alumnos con valoraciones de dislexia, discalculia, déficit de atención con o sin hiperactividad, autismo, altas capacidades, etc., requieren habitualmente de profesionales con un conocimiento preciso de cómo estimular el desarrollo de las funciones ejecutivas para facilitar su aprendizaje y su inclusión social.

En las últimas décadas se han constatado los beneficios de la estimulación de las funciones ejecutivas del cerebro para el conjunto de la población, no sólo para personas con una valoración clínica, sobre todo cuando dicha estimulación se produce en edades tempranas. Esta cuestión es particularmente relevante habida cuenta de que las actuales cohortes de niños de 5 años están rezagadas en funcionamiento ejecutivo en comparación con niños de 5 años de hace unas generaciones (Smirnova, 1998; Smirnova y Gudavera, 2004). Pese a la relevancia del funcionamiento ejecutivo para el desarrollo equilibrado (cognitivo, emocional, social y físico) de los menores, estudios recientes indican que, según los propios maestros de Educación Infantil, más de la mitad de sus alumnos no poseen destrezas de funcionamiento ejecutivo efectivas, al tiempo que señalan a la función ejecutiva deficiente como el principal desafío para el proceso de enseñanza-aprendizaje (Rimm-Kaufman et al., 2000).

A pesar de la evidencia científica a favor de la relevancia de la capacidad de autorregulación y de la necesidad de su estimulación, sobre todo desde edades tempranas, el modelo de escuela vigente en España, y en la práctica totalidad de países desarrollados, sigue centrado en los aspectos cognitivos, en la acumulación progresiva y secuencial de conocimientos reglados. Esta forma de concebir la educación,

más acertada para un país de incipiente industrialización que para uno que aspira a un cambio de modelo productivo centrado en la innovación y la creatividad, deja en manos del buen hacer, el criterio, el voluntarismo y la personalidad particular de cada maestro, y de cada estudiante, aspectos cruciales como la motivación de los estudiantes, el desarrollo de su autoconfianza, la estimulación de su control inhibitorio, de otras dimensiones de las habilidades no cognitivas como, por ejemplo, la mentalidad de crecimiento, y, en general, de las funciones ejecutivas.

La evidencia aquí revisada sugiere que podrían derivarse notables beneficios en términos de reducción de la pobreza y la desigualdad en España de intervenciones dirigidas a hogares desfavorecidos, o menos favorecidos, con niños en edades muy tempranas e incluso antes de su nacimiento. En concreto, dado que la escolarización en España es prácticamente universal a partir de los tres años de edad, nuestra sugerencia tiene dos componentes: a) desarrollar y evaluar programas dirigidos a futuros padres y a padres y cuidadores habituales (los propios padres, abuelos, guarderías, etc.) de niños menores de 3 años de edad; y b) desarrollar guías de buenas prácticas que ilustren cómo alcanzar los objetivos socioemocionales de desarrollo que las Comunidades Autónomas suelen establecer para niños 3 a 6 años de edad.

La eficacia de estas intervenciones será optimizada si estas se definen sobre la base de las buenas prácticas para la estimulación de las habilidades no cognitivas y las funciones ejecutivas, tiene una duración de al menos un año e implica de forma activa a padres y maestros, creando una cultura o proyecto de centro educativo y hogar, generalizable a otros contextos.

Al formar a los maestros de un centro educativo en las buenas prácticas conducentes a la estimulación de las habilidades no cognitivas y las funciones ejecutivas, estaremos contribuyendo no solo al desarrollo cognitivo, emocional y social de todos los estudiantes, sino también a crear colegios en los que la inclusión de los estudiantes que provienen de contextos desfavorecidos, tienen dificultades de aprendizaje, necesidades educativas especiales o altas capacidades es más probable. Este es precisamente el objetivo del proyecto *Educar para Ser* desarrollado, entre otros, por los autores del presente capítulo, presente en más de 40 colegios de la Comunidad de Madrid y actualmente en fase de evaluación experimental en China[1]. Cabe, asimismo, destacar el programa Pentabilities, de

[1] Más información sobre el proyecto *Educar para Ser* en www.educarparaser.org

la Fundación Bofill, que utiliza herramientas digitales para estimular el desarrollo de las habilidades socioemocionales de los estudiantes[2].

## 6. CONCLUSIONES

Las habilidades no cognitivas son un concepto amplio que incluye los rasgos de personalidad, atributos, disposiciones, habilidades sociales, actitudes, capacidades y recursos personales, independientes de la capacidad intelectual.

Dado su origen interdisciplinar, estas habilidades han recibido múltiples nombres en las literaturas económica, psicológica, pedagógica o filosófica incluyendo rasgos de personalidad, carácter o temperamento, valores, funciones ejecutivas, habilidades socioemocionales, y habilidades no cognitivas, entre otros.

La medición de las habilidades no cognitivas es una tarea compleja. Las medidas basadas en encuestas sufren de sesgos importantes por lo que los investigadores han buscado medidas alternativas basadas en datos administrativos y metadatos de exámenes y encuestas que los estudiantes toman en el colegio.

Aunque existe un componente hereditario importante de las habilidades no cognitivas estas habilidades se desarrollan a lo largo de la vida con periodos críticos durante la infancia, y su desarrollo se ve influido por factores culturales y el entorno familiar y educativo.

La estimulación de las habilidades no cognitivas desde edades tempranas tiene notables efectos positivos en el rendimiento y la convivencia escolar, el abandono educativo temprano, la delincuencia, el estado de salud y, en concreto, el abuso de sustancias estupefacientes, la categoría ocupacional, los ingresos salariales y, en definitiva, en el bienestar a lo largo de la vida.

Asimismo, los beneficios de las intervenciones eficaces en la estimulación de las habilidades no cognitivas suelen ser mayores para colectivos socialmente más vulnerables y se transmiten de padres a hijos.

---

[2] Más información sobre el proyecto *Pentabilities* en https://fundaciobofill.cat/en/pentabilities-en

## EJERCICIOS

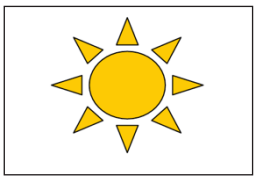
Ejercicio 1. *Imagina que estás interesado en estudiar la relación entre la motivación y el rendimiento escolar. Tienes datos de una encuesta de estudiantes al principio del curso académico y mides su rendimiento al final. En un análisis de regresión entre estas medidas encuentras una correlación negativa. Parece que una mayor motivación está asociada con un menor rendimiento. Define los conceptos de "social desirability bias" y "reference group bias". Explica como estos podrían explicar la relación negativa que se encuentra entre motivación y rendimiento educativo.*

Ejercicio 2. *En tus propias palabras, explica los posibles mecanismos por los que, en tu opinión, las intervenciones más costo-eficaces son aquellas dirigidas a los niños más jóvenes.*

Ejercicio 3. El poder de la reflexión

*La tarea de la noche y el día se diseñó para analizar las habilidades de autorregulación de niños de 4 o 5 años. Es un sencillo juego en el que se les pide que digan "día" si ven la imagen de la luna y "noche" si ven la imagen del sol. En una primera ronda del juego se estimó que el 56% de los niños acertaban en su respuesta. A continuación, se pidió a una muestra comparable a la anterior en características socioeconómicas y escolares, que jugaran al juego, pero dejando que sonara durante tres segundos una canción infantil antes de responder y después de ver la imagen. El porcentaje de acierto en este caso se situó en un 89%.*

*Día*



*Noche*

*¿Qué te sugieren estos resultados? ¿Cómo se te ocurre que se puede fomentar el hábito de la reflexión en un aula de Secundaria? ¿Preguntar a mano alzada fomenta la reflexión?*

Ejercicio 4. El poder de un M&M

*Un estudio midió el cociente intelectual (CI) de un amplio grupo de alumnos de Primaria y los clasificó en tres grupos: alto, medio y bajo. Transcurridos unos meses, los investigadores volvieron a medir el CI de estos alumnos con diferentes preguntas. A la mitad de los alumnos de cada grupo, aleatoriamente elegidos, se les ofreció un caramelo M&M por acierto.*

*Los CI no variaron entre el primer y el segundo test, independientemente de que se les ofreciera el M&M o no, entre los alumnos de CI alto y medio. No obstante, los alumnos de CI bajo que tuvieron el incentivo de los caramelos M&M incrementaron en un 45% su CI promedio, manteniéndose éste inalterado entre los alumnos de CI bajo que no tuvieron dicho incentivo.*

*¿Cómo explicas estos resultados? ¿Crees que comer M&M te puede hacer más inteligente? ¿Qué crees que miden las medidas tradicionales de inteligencia?*

Ejercicio 5. Eligiendo extraescolares

*Unos padres están decidiendo, con su hijo de 8 años, las actividades extraescolares de este curso. Los padres, que han leído sobre las habilidades no cognitivas, quieren elegir una actividad que contribuya a un desarrollo global (cognitivo, emocional y social) de su hijo. Consideran estas opciones: fútbol sala, Tae-Kwon-Do, cantar en un coro, tocar un instrumento musical y el método Kumon de cálculo. ¿Cuál/es de estas actividades les recomendarías? ¿Por qué?*

*Tras consultar con su hijo, este deja claro que sólo contempla fútbol sala y Tae-Kwon-Do. ¿Cuál recomendarías? ¿Por qué?*

## BIBLIOGRAFÍA


Arco-Tirado, J.L., Fernández-Martín, F.D. y Hoyle R.H. (2018). Development and Validation of a Spanish Version of the Grit-S Scale. *Front Psychol, 9*, 96. doi: 10.3389/fpsyg.2018.00096. PMID: 29467705.

Balart, P. y Cabrales, A. (2014). *La maratón de PISA: la perseverancia como factor del éxito en una prueba de competencias.* Monografías Sobre Educación. Fundación Ramón Areces y Fundación Europea Sociedad y Educación.

Balart, P. y Oosterveen, M. (2019). *Females show more sustained performance during test-taking than males*. Nat Commun 10, 3798. https://doi.org/10.1038/s41467-019-11691-y.

Borghans, L. y Schils, T. (2012). *The leaning tower of Pisa: Decomposing achievement test scores into cognitive and noncognitive components*. Retrieved from http://www.sole-jole.org/13260.pdf.

Bouchard, T. J., Jr. y Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics, 31*(3), 243-273. https://doi.org/10.1023/A:1012294324713.

Duckworth, A. L., Peterson, C., Matthews, M. D. y Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087-1101. https://doi.org/10.1037/0022-3514.92.6.1087.

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology, 44*(3), 329-344. https://doi.org/10.1037/h0057198.

Gneezy, U., List, J. A., Livingston, J.A., Qin, X., Sadoff, S. y Xu, Y. (2019). Measuring Success in Education: The Role of Effort on the Test Itself. *American Economic Review: Insights, 1*(3): 291-308. DOI: 10.1257/aeri.20180633.

Hitt, C., Trivitt, J. y Cheng, A. (2016). When you say nothing at all: The predictive power of student effort on surveys. *Economics of Education Review, 52*, 105-119. https://doi.org/10.1016/j.econedurev.2016.02.001.

John, O. P., Naumann, L. P. y Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement and Conceptual Issues, en O. P. John, R. W. Robins y L. A. Pervin (eds.), *Handbook of personality: Theory and research*. New York, NY: Guilford Press, 114-158.

John, O. P., Donahue, E. M. y Kentle, R. L. (1991). *The Big Five Inventory-Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

Mischel, W., Ebbesen, E.B. y Zeiss, A.R. (1972). Cognitive and attentional mechanisms in delay of gratification. *J Pers Soc Psychol, 21*(2), 204-18. https://doi.org/10.1037/h0032198.

Watts, T.W., Duncan, G.J. y Quan, H. (2018). Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Delay of Gratification and Later Outcomes. *Psychol Sci., 29*(7), 1159-1177. doi: 10.1177/0956797618761661.

Zamarro, G. (2021). *Motivación académica, habilidades no cognitivas y brecha de género en matemáticas y ciencias. El caso de España.* Monografías sobre Educación. Fundación Ramón Areces y Fundación Europea Sociedad y Educación.

Zamarro, G., Hitt, C. y Mendez, I. (2019). When students don't care: Reexamining international differences in achievement and student effort. *Journal of Human Capital, 13*(4), 519-552. https://doi.org/10.1086/705799.

# Anexo
# ACTUALIZACIÓN DE LOS RESULTADOS DE PISA 2022

Ismael Sanz
Departamento de Economía Aplicada I e Historia e Instituciones Económicas
Universidad Rey Juan Carlos
Department of Social Policy
London School of Economics
ORCID: https://orcid.org/0000-0003-1286-4124

Después de la finalización de los capítulos del libro, la OCDE ha publicado los resultados de PISA 2022. El 5 de diciembre de 2023 se conocía el Informe PISA centrado en esta ocasión en Matemáticas. A continuación, se presenta un anexo al libro con los principales resultados de esta evaluación internacional de la OCDE.

## Introducción. Principales resultados

La prueba de PISA 2022 de Matemáticas, Lectura y Ciencias a estudiantes de 15 años ha contado con la participación de los 37 países de la OCDE, y otros 44 países y economías que sin ser de este organismo internacional también realizan los test. La tabla 1 de este Anexo muestra la evolución de los resultados de España en comparación a la OCDE desde 2009 a 2022. Comenzamos por Matemáticas, el área principal en PISA 2022, como también lo fue en 2012, lo que significa que el alumnado realizó más preguntas de Matemáticas que de Lectura y Ciencias, por lo que los resultados son más precisos. En concreto, el 94% de los estudiantes realizó la prueba de 60 minutos de Matemáticas, y dedicaron otros 60 minutos a una de estas tres áreas: Lectura, Ciencias o pensamiento creativo. El otro 6% de los estudiantes realizó las pruebas de Lectura y Ciencias. Es la primera vez que PISA evalúa pensamiento creativo y sus resultados serán conocidos más tarde que los de Matemáticas, Lectura y Ciencias, a lo largo de 2024.

La puntuación de los estudiantes españoles en PISA 2022 ha sido de 473 puntos, registrando el peor resultado desde 2009, e incluso desde el primer

PISA centrado en Matemáticas en 2003. El alumnado de nuestro país ha bajado su rendimiento desde los 484 puntos de 2012 a los 473 de 2022. La disminución en los resultados de la OCDE ha sido incluso más intensa, desde los 494 de 2012 puntos a los 472 de 2022, de modo que los estudiantes españoles se encuentran ahora en el promedio de los países desarrollados. Por construcción, el resultado de los estudiantes de 15 años en PISA 2003, la primera edición que se centró en Matemáticas fue de 500 puntos y desviación típica de 100. Es decir, que tanto el alumnado español como el de la OCDE tiene ahora en 2022 menos competencias en Matemáticas que los estudiantes que tenían esa misma edad en 2003. La OCDE (2023) establece que 20 puntos de PISA 2022 equivalen a lo que un estudiante aprende durante un curso escolar. En ediciones anteriores la OCDE había fijado en una puntuación más elevada de esta equivalencia a un curso escolar. La regla general que emplean Hanushek y Woessmann (2020) "encontrada a partir de comparaciones de aprendizaje en pruebas diseñadas para evaluar el rendimiento a lo largo del tiempo" es que los estudiantes, en promedio, aprenden alrededor de un tercio de una desviación estándar por año escolar. En consecuencia, se puede concluir que los estudiantes de 15 años de la OCDE y de España en 2022 tienen casi un curso escolar menos de competencias en Matemáticas que el alumnado de esa edad en 2003, puesto que los 28 puntos de diferencia entre 2003 y 2022 se aproximan a los 33 puntos.

En dos décadas de historia de PISA, la puntuación promedio de la OCDE nunca había cambiado más de cuatro puntos en Matemáticas entre evaluaciones consecutivas (OCDE, 2023). La caída de 17 puntos en Matemáticas de la OCDE entre 2018 y 2022 señala que el cierre de los centros educativos durante la pandemia del Covid-19 ha sido un shock general que ha afectado a muchos países al mismo tiempo. La disminución de los resultados del alumnado de 17 puntos, de 489 a 472 puntos, es similar al 14% de la desviación estándar de pérdida de aprendizaje por el Covid estimada en el metaanálisis de Betthäuser et al. (2023) o el 18% que estiman Sanz y Tena (2023). Pero el análisis de tendencias de los resultados de PISA revela un declive de décadas que comenzó mucho antes de la pandemia. La pérdida de competencias en Matemáticas de los estudiantes se había producido ya antes de PISA 2018, cuando la puntuación de la OCDE era de 489 puntos, 11 menos que los 500 de PISA 2003, indicando que el sistema educativo a nivel internacional tiene problemas de largo plazo (OCDE, 2023).

**Tabla 1. Resultados de España y OCDE en PISA 2009, 2012, 2015, 2018 y 2022.**

| | | 2009 | 2012 | 2015 | 2018 | 2022 |
|---|---|---|---|---|---|---|
| Matemáticas | España | 483 | 484 | 486 | 481 | 473 |
| | OCDE | 495 | 494 | 490 | 489 | 472 |
| Ciencias | España | 488 | 496 | 493 | 483 | 485 |
| | OCDE | 501 | 501 | 493 | 489 | 485 |
| Lectura | España | 481 | 488 | 496 | 477* | 474 |
| | OCDE | 493 | 496 | 493 | 487 | 476 |

* Los resultados de Lectura de España en PISA 2018 se tienen que tomar con cautela porque la OCDE concluyó que "no se puede garantizar completamente la comparabilidad de los datos de PISA 2018 para España con los de evaluaciones anteriores de PISA" (OCDE, 2020, https://www.oecd.org/pisa/PISA2018-AnnexA9-Spain.pdf).
Fuente: PISA 2022. Informe español. INEE, Ministerio de Educación, FP y Deporte. INEE.

El análisis de tendencias de los resultados de PISA en Lectura y Ciencias antes de 2018 también revela que el rendimiento comenzó a disminuir mucho antes de la pandemia. En estas materias, el rendimiento alcanzó su punto máximo en 2012 y 2009, respectivamente, antes de disminuir. De hecho, la evolución de los resultados en Lectura entre 2018 y 2022 ha sido muy similar a la de Matemáticas, con España, OCDE registrando caídas muy pronunciadas que las sitúan en una puntuación de 474 y 476 puntos respectivamente. La diferencia entre los resultados del alumnado español y de la OCDE no es estadísticamente significativa. Los estudiantes de 15 años de nuestro país y los del conjunto de países desarrollados tienen menos competencias lectoras que los estudiantes que realizaron PISA en su primera edición en 2000, centrada en Lectura. En Ciencias, España ha registrado 485 puntos en PISA 2022, un resultado muy similar a los 483 de PISA 2018. Nuestro país ha logrado detener el declive en las puntuaciones en Ciencias que se había venido produciendo desde los 496 puntos alcanzados en 2012. Además, España se encuentra ahora también en Ciencias en el promedio de la OCDE que ha caído en 4 puntos entre PISA 2018 y PISA 2022 desde los 489 a los 485 puntos.

**Tabla 2. Resultado en Matemáticas por países y para las Comunidades Autónomas españolas en PISA 2022.**

| | Media | 95% Intervalo de confianza | Ranking OCDE | |
|---|---|---|---|---|
| | | | Límite superior | Límite inferior |
| Japón | 536 | 530 - 541 | 1 | 2 |
| Corea | 527 | 520 - 535 | 1 | 2 |
| Estonia | 510 | 506 - 514 | 3 | 4 |
| Suiza | 508 | 504 - 512 | 3 | 5 |
| Castilla y León | 499 | 492 - 507 | | |
| Canadá* | 497 | 494 - 500 | 5 | 13 |
| Asturias | 495 | 486 - 504 | | |
| Cantabria | 495 | 486 - 504 | | |
| Comunidad de Madrid | 494 | 487 - 501 | | |
| Países Bajos* | 493 | 485 - 500 | 4 | 20 |
| La Rioja | 493 | 485 - 501 | | |
| Navarra | 492 | 484 - 501 | | |
| Irlanda* | 492 | 488 - 496 | 5 | 18 |
| Bélgica | 489 | 485 - 494 | 5 | 20 |
| Dinamarca* | 489 | 485 - 493 | 5 | 19 |
| Reino Unido* | 489 | 485 - 493 | 5 | 20 |
| Polonia | 489 | 485 - 493 | 5 | 20 |
| Austria | 487 | 483 - 492 | 5 | 20 |
| Australia* | 487 | 484 - 491 | 6 | 20 |
| República Checa | 487 | 483 - 491 | 5 | 20 |
| Aragón | 487 | 478 - 496 | | |
| Galicia | 486 | 479 - 494 | | |
| Eslovenia | 485 | 482 - 487 | 6 | 21 |
| Finlandia | 484 | 480 - 488 | 6 | 24 |
| Letonia | 483 | 479 - 487 | 6 | 25 |
| País Vasco | 482 | 474 - 490 | | |
| Suecia | 482 | 478 - 486 | 6 | 27 |
| Nueva Zelanda | 479 | 475 - 483 | 7 | 28 |
| Lituania | 475 | 472 - 479 | 16 | 29 |

| | | | | |
|---|---|---|---|---|
| Alemania | 475 | 469 - 481 | 8 | 30 |
| Francia | 474 | 469 - 479 | 15 | 29 |
| España | 473 | 470 - 476 | 18 | 29 |
| Hungría | 473 | 468 - 478 | 16 | 30 |
| Comunidad Valenciana | 473 | 465 - 480 | | |
| Portugal | 472 | 467 - 477 | 17 | 30 |
| Italia | 471 | 465 - 477 | 16 | 31 |
| Islas Baleares | 471 | 463 - 478 | | |
| Cataluña | 469 | 458 - 481 | | |
| Extremadura | 469 | 459 - 479 | | |
| Noruega | 468 | 464 - 472 | 19 | 31 |
| Estados Unidos* | 465 | 457 - 473 | 18 | 32 |
| República Eslovaca | 464 | 458 - 470 | 20 | 32 |
| Castilla-La Mancha | 464 | 457 - 470 | | |
| Murcia | 463 | 455 - 472 | | |
| Islandia | 459 | 456 - 462 | 26 | 32 |
| Israel | 458 | 451 - 464 | 23 | 32 |
| Andalucía | 457 | 448 - 467 | | |
| Turquía | 453 | 450 - 456 | 28 | 32 |
| Islas Canarias | 447 | 438 - 456 | | |
| Grecia | 430 | 426 - 435 | 33 | 33 |
| Chile | 412 | 408 - 416 | 34 | 34 |
| Melilla | 404 | 392 - 416 | | |
| México | 395 | 391 - 399 | 35 | 37 |
| Ceuta | 395 | 382 - 407 | | |
| Costa Rica | 385 | 381 - 388 | 35 | 37 |
| Colombia | 383 | 377 - 389 | 35 | 37 |

* Se requiere precaución al interpretar las estimaciones porque uno o más estándares de muestreo de PISA no se cumplieron en estos países.
Fuente: OCDE, PISA 2022 Database, tabla I.B1.2.1 y tabla I.B2.1.

La tabla 2 muestra que Japón y Corea lideran la clasificación de la OCDE en Matemáticas, la materia central de PISA 2022. Además, de estos dos países, otros 17 más tuvieron un desempeño significativamente por encima del promedio de la OCDE (472 puntos), desde Estonia (510 puntos) hasta Nueva Zelanda (479 puntos). La brecha de rendimiento entre los países de mayor y menor

rendimiento es de 153 puntos en Matemáticas entre los países de la OCDE en PISA 2022.

Si nos centramos en la UE, Estonia se ha confirmado en PISA como el país líder y el único comunitario que aguanta el pulso a los países asiáticos. Ha obtenido 510 puntos en PISA Matemáticas por los 536 de Japón y los 527 de Corea. Países Bajos (493) e Irlanda (492) ocupan el segundo y tercer puesto de la UE en Matemáticas, con el promedio comunitario situado en 474 puntos. En Lectura, Irlanda (516) ha quedado la primera de toda la UE, e incluso de toda la OCDE seguida muy de cerca por Japón (516) y Corea (515), mientras que Estonia es el segundo Estado miembro de la UE con 511 puntos. La media comunitaria en Lectura está muy alejada de Irlanda y Estonia, con 475 puntos. En Ciencias, Estonia es la primera de la UE con 526 puntos, seguida por Finlandia (511) e Irlanda (504), muy por encima las tres de los 484 puntos de media comunitaria. Se puede concluir que Estonia e Irlanda dominan la clasificación de la UE en Matemáticas, Lectura y Ciencias, aunque los resultados de Irlanda tienen asociado un asterisco en el Informe PISA de 2022, indicando que hay que tener precaución en la comparación con otros países. Estonia e Irlanda tienen una ventaja de 30 puntos con respecto a la media de la UE si tenemos en cuenta las tres competencias de Matemáticas, Lectura y Ciencias que es casi el equivalente a lo que se aprende en un curso escolar. Es decir, que los estudiantes del 9º curso de Estonia e Irlanda (nuestro 3º de la ESO) tienen las mismas competencias que los estudiantes de 10º curso (4º de la ESO) españoles, de la OCDE o de la UE. Estonia, e Irlanda, son la "nueva Finlandia", el país europeo que en los 2000 se convirtió en la referencia del viejo continente por sus buenos resultados educativos en calidad y equidad.

Estonia e Irlanda comparten varias estrategias clave que han contribuido a su éxito educativo (OCDE, 2020). Algunas de las políticas que España podría implementar y que parecen estar dando resultado en Estonia e Irlanda serían medidas específicas que subrayan un enfoque de retroalimentación a los equipos directivos y profesores sobre su desempeño, una mayor autonomía pedagógica de los centros educativos y el diseño de una carrera de la profesión docente. Ambos países han implementado estrategias clave en áreas como:

- Autonomía escolar y gobernanza descentralizada: Estonia ha proporcionado a los centros educativos considerable autonomía para diseñar currículos y métodos de enseñanza adaptados a necesidades locales, como se evidencia en la reforma curricular de 2014.

- Desarrollo profesional de docentes: tanto Estonia como Irlanda han puesto un énfasis significativo en el desarrollo profesional de sus docentes y líderes escolares. Estonia ha implementado un sistema de carrera basado en competencias desde 2013, mientras que Irlanda ha enfocado esfuerzos en el desarrollo profesional continuo, especialmente en centros educativos desfavorecidos.
- Uso de tecnología y fomento de habilidades digitales: Estonia destaca por la integración de tecnología en su educación, con iniciativas como la "E-Schoolbag" y el programa "Espejo Digital".
- Adaptación a cambios demográficos y sociales: Irlanda ha respondido a cambios demográficos y sociales con programas como "First 5" y ha incrementado la inversión en educación temprana, ofreciendo preescolar gratuito y expandiendo el aprendizaje temprano.
- Evaluación y retroalimentación constructiva: Irlanda ha mejorado la calidad educativa a través de políticas de evaluación continua y retroalimentación a los estudiantes, docentes, y comunidad educativa en general.
- Consolidación de centros educativos: Estonia ha consolidado su red escolar, especialmente en áreas rurales, fusionando escuelas con pocos estudiantes y creando centros de secundaria en cada capital de condado bajo programas específicos.

También Polonia y Reino Unido han tenido unos buenos resultados en relación con la UE en Matemáticas, y sobre todo en Lectura y Ciencias. El buen rendimiento de Polonia comenzó en el año 2000 con registros muy similares a España. En 2022, se sitúa en las primeras posiciones en las tres competencias con 489 puntos en Matemáticas y Lectura y 499 en Ciencias, aunque el cierre de los centros educativos ha tenido un impacto muy negativo en el aprendizaje de sus estudiantes. Reino Unido, por su parte, ha alcanzado 489, 494 y 500 puntos respectivamente, confirmando el buen rendimiento educativo que este país ya logró en la prueba internacional de Lectura de 4º de Primaria de PIRLS 2021, cuyos resultados se conocieron en mayo de 2023. Polonia y Reino Unido son países de gran tamaño, a diferencia de Estonia e Irlanda. La gestión educativa puede ser más fácil en países de pequeña magnitud como lo demuestra el hecho de que los países líderes en la UE (Estonia, 1,3 millones, e Irlanda, 5,1 millones) o en la OCDE (Singapur, 5,6 millones) son países con poca población y de pequeña extensión.

Dos países que habían recibido mucha atención entre la comunidad educativa española, Finlandia y Portugal, han tenido caídas muy importantes entre PISA 2018 y PISA 2022 que se unen a las que ya habían registrado en la edición anterior. Finlandia ha caído más de 33 puntos entre PISA 2012 y PISA 2022 en las tres competencias, una disminución que además ha sido más intensa en el decil de menor rendimiento y en el cuartil de menor nivel socioeconómico. En consecuencia, las diferencias entre los estudiantes de mejor y peor rendimiento y entre el alumnado de entornos favorecidos y vulnerables ha aumentado de forma significativa.

Por Comunidades Autónomas en España, podemos distinguir tres grupos. Las Comunidades que tienen significativamente mejores resultados que la media de España (473) en Matemáticas, que son: Castilla León (499), Asturias (495), Cantabria (495), Comunidad de Madrid (494), La Rioja (493), Navarra (492), Aragón (487), Galicia (486) y País Vasco (482). Un segundo grupo ha alcanzado una puntuación que no es estadísticamente diferente del conjunto del país: Comunidad Valenciana (473), Islas Baleares (471), Cataluña (469) y Extremadura (469). Y finalmente, se sitúan las Comunidades y ciudades autónomas con un resultado significativamente inferior a nuestro país: Castilla La Mancha (464), Murcia (463), Andalucía (457), Canarias (447), Melilla (404) y Ceuta (395). Estos tres grupos responden a una división geográfica, con el centro y noroeste (además de Aragón y Navarra) mostrando los mejores resultados, seguido del noreste (junto con Extremadura) en la media, mientras que el sur registra las puntuaciones más bajas. La clasificación que muestra PISA 2022 es similar a las de otras ediciones de esta misma prueba (exceptuando la de 2018 que tuvo problemas en su aplicación en España), u otras evaluaciones internacionales como TIMSS y PIRLS. El ranking del rendimiento educativo por Comunidades también es similar a la Evaluación General de Diagnostico de Primaria (2009) y Secundaria (2010) realizadas por el Ministerio de Educación, FP y Deportes. La única excepción podría ser el caso de Castilla La Mancha, que en las pruebas de Evaluación General de Diagnóstico del Ministerio se encontraría más bien en el segundo grupo, en la media de España.

La situación de las Comunidades Autónomas españolas cambia cuando se tiene en cuenta su índice del status socioeconómico y cultural (ISEC). El ISEC es una puntuación compuesta que combina en una sola la información tres factores: el nivel educativo más alto de los padres; la ocupación más elevada; y posesiones del hogar como una habitación propia; software educativo o aplicaciones; coches;

libros o dispositivos digitales con pantallas. Los estudiantes proporcionan esta información en el cuestionario de contexto, que responden después de completar las pruebas cognitivas. De este modo, el ISEC permite distinguir entre estudiantes desfavorecidos (el 25% del alumnado con el ISEC más bajo en su país) y estudiantes favorecidos (el 25% con el ISEC más elevado en su país). Este índice se construyó por primera vez en 2000, con media 0 para el promedio de la OCDE en ese año y desviación típica 1 (Ministerio de Educación, FP y Deportes, 2023). La tabla 3 muestra el ISEC para las comunidades y ciudades autónomas junto con el valor de España y la media de la OCDE. España tiene un ISEC de –0,03 indicando que el nivel socioeconómico y cultural en nuestro país es similar al de la OCDE (0,00), aunque algo más bajo. Melilla (–0,79), Ceuta (–0,47) y Murcia (–0,24) tienen los índices más bajos, con una situación socioeconómica peor que la de OCDE y España. En el lado contrario se encuentran la Comunidad de Madrid (0,25), País Vasco (0,18), Asturias (0,09) o Castilla León (0,08) con un ISEC superior a España y al conjunto de países desarrollados.

La última columna de la tabla 3, permite "descontar" la influencia del ISEC en la puntuación de Matemáticas de las Comunidades y ciudades Autónomas. Es decir, se muestran los resultados que habría registrado cada una de las regiones si las condiciones sociales, económicas y culturales de los estudiantes fueran las mismas. Así, por ejemplo, se comprueba que Melilla, Ceuta y Murcia tienen, sólo en parte, resultados inferiores en Matemáticas porque disponen de los ISEC más bajos. Una vez tenido en cuenta que los padres de los estudiantes de estas ciudades autónomas y la Región tienen, en promedio, un nivel educativo más bajo, ocupaciones para las que se requiere de menor cualificación y menos recursos, Melilla mejora 23 puntos, pasando de 404 a 427, Ceuta 12 (llegando a 407) y Murcia 9 (situándose en 472). Madrid (–7 puntos) y País Vasco (–5 puntos) son las Comunidades que más reducen su puntuación al descontar el efecto del ISEC, sin ser una diferencia estadísticamente significativa (Ministerio de Educación, FP y Deportes, 2023). En general, la agrupación en tres bloques de Comunidades, las que tienen mejores resultados que España, las que se encuentran en la media, y las que están por debajo, no cambia mucho. País Vasco (477) dejaría el primer grupo para pasar al segundo con un resultado muy similar a la media de España (475) y Murcia pasaría del último grupo al segundo con una puntuación parecida al conjunto del país (472 frente a 475). El resto de Comunidades seguiría en los mismos grupos descritos con los resultados brutos de Matemáticas.

**Tabla 3. Resultados brutos en Matemáticas, índice socioeconómico y cultural y puntuación teniendo en cuenta el ISEC por Comunidades Autónomas.**

| | Resultados brutos | ISEC | Resultados descontando el ISEC |
|---|---|---|---|
| Castilla y León | 499 | 0,08 | 499 |
| Asturias | 495 | 0,09 | 493 |
| Cantabria | 495 | 0,03 | 495 |
| Comunidad de Madrid | 494 | 0,25 | 487 |
| La Rioja | 493 | –0,06 | 495 |
| Navarra | 492 | 0,02 | 493 |
| Aragón | 487 | 0,00 | 489 |
| Galicia | 486 | 0,07 | 485 |
| País Vasco | 482 | 0,18 | 477 |
| España | 473 | –0,03 | 475 |
| Comunidad Valenciana | 473 | –0,10 | 476 |
| OCDE | 472 | 0,00 | 475 |
| Islas Baleares | 471 | –0,05 | 473 |
| Cataluña | 469 | –0,02 | 471 |
| Extremadura | 469 | –0,14 | 474 |
| Castilla La Mancha | 464 | –0,16 | 469 |
| Murcia | 463 | –0,24 | 472 |
| Andalucía | 457 | –0,18 | 463 |
| Islas Canarias | 447 | –0,21 | 455 |
| Melilla | 404 | –0,79 | 427 |
| Ceuta | 395 | –0,47 | 407 |

Fuente: Ministerio de Educación, FP y Deportes (2023): PISA 2022. Figura 3.5.b. y figura 3.10.b. https://www.educacionyfp.gob.es/inee/evaluaciones-internacionales/pisa/pisa-2022.html

## Distribución de resultados

PISA establece seis niveles de competencia en Matemáticas, Lectura y Ciencias. El nivel 2 es el considerado como el necesario para que los estudiantes de 15 años puedan continuar sus estudios de secundaria superior (Bachillerato o FP Media) o estudios superiores (Universidad o FP Superior) y afrontar con garantías su inserción laboral y hasta social en el futuro. Pues bien, durante el período de 2012 a 2022, la proporción de estudiantes españoles de 15 años que obtuvie-

ron puntaciones por debajo del nivel 2, aumentó en tres puntos porcentuales en Matemáticas; en seis puntos porcentuales en Lectura; y en cinco puntos porcentuales en Ciencias. A pesar de este aumento, España continúa disponiendo de una proporción de estudiantes rezagados inferior a la OCDE. Por ejemplo, en Matemáticas el 27% del alumnado español, frente al 31% de la OCDE, no alcanza la competencia de nivel 2 en Matemáticas. El umbral 2 es en el que los estudiantes pueden "interpretar y reconocer, cómo se puede representar matemáticamente una situación simple como comparar la distancia total entre dos rutas alternativas, o convertir precios a una moneda diferente" (OCDE, 2023). Es decir, que un promedio del 69% de los estudiantes de la OCDE tiene al menos una competencia básica en Matemáticas. Esto significa que están comenzando a demostrar la capacidad e iniciativa para usar Matemáticas en situaciones reales simples. En Lectura nuestro país tiene un 24% de estudiantes en España que no alcanzaron el nivel 2, por el 26% de la OCDE. El nivel 2 es en el que el alumnado puede, como mínimo, "identificar la idea principal en un texto de longitud moderada, encontrar información basada en criterios explícitos, aunque a veces complejos, y pueden reflexionar sobre el propósito y la forma de los textos cuando se les dirige explícitamente a hacerlo". En Ciencias, España tiene un 21% de estudiantes que no alcanza el nivel 2, por el 24% en la OCDE, que es el umbral en el que el alumnado pueden, como mínimo, "reconocer la explicación correcta para fenómenos científicos familiares y pueden usar ese conocimiento para identificar, en casos simples, si una conclusión es válida basada en los datos proporcionados".

En el lado contrario, el alumnado excelente es el que se encuentra en los niveles 5 y 6. Nuestro país también tiene un porcentaje de estudiantes excelentes en Matemáticas (6%) inferior a la OCDE (9%), un déficit sistemático que se confirma en PISA 2022. En este nivel excelente, los estudiantes pueden "modelizar situaciones complejas matemáticamente y pueden seleccionar, comparar y evaluar estrategias de resolución de problemas apropiadas para abordarlas" (OCDE, 2023). En Lectura, España tiene un 5% del alumnado excelente, en comparación con el promedio del 7% de la OCDE. Este alumnado puede comprender textos extensos, tratar con conceptos que son abstractos o contraintuitivos, y establecer distinciones entre hecho y opinión, basados en pistas implícitas relacionadas con el contenido o la fuente de la información. Por último, un 5% de los estudiantes españoles se sitúa en el nivel excelente en Ciencias, también por debajo del 7% de la OCDE. Este alumnado puede

"aplicar de manera creativa y autónoma su conocimiento de y sobre la ciencia a una amplia variedad de situaciones, incluyendo aquellas no familiares".

En definitiva, PISA 2022 confirma una característica muy consistente del sistema educativo español: tenemos menos alumnos rezagados que en la OCDE y también una proporción inferior de estudiantes excelentes. Es decir, que nuestro país tiene más porcentaje de los alumnos en la parte media de la distribución. La falta de alumnos excelentes en España en comparación a la OCDE (y la UE) podría suponer un déficit a medio y largo plazo que podría lastrar las posibilidades de desarrollo de I+D, y la necesaria apuesta por la especialización en sectores de alto valor añadido, intensivos en capital humano y tecnológico. Alemania y Reino Unido han logrado tener también una proporción de estudiantes rezagados más baja que en el promedio de la OCDE, pero disponen de más alumnado excelente que España y que el conjunto de países desarrollados. Italia, por su parte, tiene un patrón similar al de nuestro país con una proporción algo más baja de estudiantes rezagados que la OCDE, y menos alumnos excelentes. Estados Unidos tiene una distribución de sus alumnos buena en Lectura y Ciencias, con más estudiantes excelentes y menos rezagados que el promedio de la OCDE, y una preocupante situación en Matemáticas, como ya le ocurrió en PISA 2018.

**Tabla 4. Distribución de los estudiantes españoles en PISA por niveles de rendimiento.**

| España | | 2009 | 2012 | 2015 | 2018 | 2022 |
|---|---|---|---|---|---|---|
| Lectura | Nivel bajo | 20 | 18 | 16 | 23 | 24 |
| | Nivel alto | 3 | 6 | 6 | 5 | 5 |
| Matemáticas | Nivel bajo | 24 | 24 | 22 | 25 | 27 |
| | Nivel alto | 8 | 8 | 7 | 7 | 6 |
| Ciencias | Nivel bajo | 18 | 16 | 18 | 21 | 21 |
| | Nivel alto | 4 | 5 | 5 | 4 | 5 |
| **OCDE** | | **2009** | **2012** | **2015** | **2018** | **2022** |
| Lectura | Nivel bajo | 20 | 19 | 20 | 22 | 26 |
| | Nivel alto | 7 | 8 | 8 | 8 | 7 |
| Matemáticas | Nivel bajo | 24 | 24 | 25 | 24 | 31 |
| | Nivel alto | 12 | 12 | 10 | 11 | 9 |
| Ciencias | Nivel bajo | 19 | 19 | 22 | 22 | 24 |
| | Nivel alto | 8 | 8 | 7 | 7 | 7 |

| Francia | | 2009 | 2012 | 2015 | 2018 | 2022 |
|---|---|---|---|---|---|---|
| Lectura | Nivel bajo | 20 | 19 | 21 | 21 | 27 |
| | Nivel alto | 10 | 13 | 13 | 9 | 7 |
| Matemáticas | Nivel bajo | 22 | 22 | 23 | 21 | 29 |
| | Nivel alto | 14 | 13 | 11 | 11 | 7 |
| Ciencias | Nivel bajo | 19 | 19 | 22 | 21 | 24 |
| | Nivel alto | 8 | 8 | 8 | 7 | 8 |
| **Alemania** | | **2009** | **2012** | **2015** | **2018** | **2022** |
| Lectura | Nivel bajo | 19 | 15 | 16 | 21 | 25 |
| | Nivel alto | 8 | 9 | 12 | 11 | 8 |
| Matemáticas | Nivel bajo | 19 | 18 | 17 | 21 | 30 |
| | Nivel alto | 18 | 18 | 13 | 13 | 9 |
| Ciencias | Nivel bajo | 15 | 12 | 17 | 20 | 23 |
| | Nivel alto | 13 | 12 | 11 | 10 | 10 |
| **Italia** | | **2009** | **2012** | **2015** | **2018** | **2022** |
| Lectura | Nivel bajo | 21 | 20 | 21 | 23 | 21 |
| | Nivel alto | 6 | 7 | 6 | 5 | 5 |
| Matemáticas | Nivel bajo | 25 | 25 | 23 | 24 | 30 |
| | Nivel alto | 9 | 10 | 11 | 10 | 7 |
| Ciencias | Nivel bajo | 21 | 19 | 23 | 26 | 24 |
| | Nivel alto | 6 | 6 | 4 | 3 | 4 |
| **Reino Unido** | | **2009** | **2012** | **2015** | **2018** | **2022** |
| Lectura | Nivel bajo | 18 | 17 | 18 | 17 | 20 |
| | Nivel alto | 8 | 9 | 9 | 12 | 10 |
| Matemáticas | Nivel bajo | 20 | 22 | 22 | 19 | 24 |
| | Nivel alto | 10 | 12 | 11 | 13 | 11 |
| Ciencias | Nivel bajo | 15 | 15 | 17 | 17 | 20 |
| | Nivel alto | 11 | 11 | 11 | 10 | 10 |
| **Estados Unidos** | | **2009** | **2012** | **2015** | **2018** | **2022** |
| Lectura | Nivel bajo | 18 | 17 | 19 | 19 | 20 |
| | Nivel alto | 10 | 8 | 10 | 13 | 14 |
| Matemáticas | Nivel bajo | 23 | 26 | 29 | 27 | 34 |
| | Nivel alto | 10 | 9 | 6 | 8 | 7 |
| Ciencias | Nivel bajo | 18 | 18 | 20 | 19 | 22 |
| | Nivel alto | 9 | 8 | 9 | 9 | 11 |

Fuente: OCDE 2023. PISA 2022 Results (Volume I). Gráfico I.3.1; gráfico I.3.4 y gráfico I.3.5.

## Resultados y nivel socioeconómico de los estudiantes

Como ya se ha comentado, el índice PISA de estatus económico, social y cultural se computa de tal forma que todos los estudiantes de todos los países pueden ser situados en la misma escala socioeconómica. Por tanto, es posible utilizar este índice para comparar el rendimiento de estudiantes en relación con su nivel socioeconómico independientemente del país en el que se encuentre. En España, el 36% de los estudiantes se encuentra en el quintil superior internacional de la escala socioeconómica. Las Tablas 5 y 6 del Anexo, muestran que la puntuación de este 36% de alumnado español de mayor nivel socioeconómico en comparación a todos países participantes en PISA en Matemáticas fue de 511 puntos, por los 517 de la OCDE. En Alemania (534 puntos) o Reino Unido (529) el alumnado con este mismo nivel socioeconómico tiene puntuaciones significativamente más altas. Por el contrario, el 6,3% de estudiantes españoles que se encuentra en el quintil de menor nivel socioeconómico de entre todos los países participantes en PISA tiene 412 puntos, significativamente por encima de los 396 del alumnado del nivel socioeconómico similar en la OCDE. Y no es sólo el quintil más bajo, el 10,7% de estudiantes españoles que se encuentra en el siguiente quintil internacional, el segundo de menor nivel socioeconómico

**Tabla 5. Proporción de estudiantes por quintiles internacionales de nivel socioeconómico.**

| | Porcentaje de estudiantes, por quintiles internacionales de ISEC | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Quintil inferior | | Segundo quintil | | Quintil medio | | Cuarto quintil | | Quintil superior | |
| | % | E.E. | % | E.E. | % | E.E. | % | E.E. | % | E.E. |
| España | 6,3 | (0,4) | 10,7 | (0,3) | 20,2 | (0,5) | 26,5 | (0,5) | 36,2 | (0,8) |
| **OCDE** | 4,9 | (0,1) | 11,8 | (0,1) | 20,4 | (0,1) | 25,7 | (0,1) | 37,1 | (0,2) |
| Francia | 3,2 | (0,3) | 11,7 | (0,4) | 22,5 | (0,8) | 28,5 | (0,6) | 34,2 | (1,0) |
| Alemania | 6,0 | (0,4) | 15,2 | (0,6) | 21,7 | (0,5) | 26,2 | (0,7) | 30,9 | (1,1) |
| Italia | 4,2 | (0,3) | 13,6 | (0,5) | 24,9 | (0,6) | 27,8 | (0,6) | 29,5 | (1,0) |
| Reino Unido | 1,9 | (0,3) | 9,7 | (0,4) | 21,7 | (0,8) | 26,7 | (0,7) | 40,0 | (1,0) |
| Estados Unidos | 3,9 | (0,5) | 11,5 | (0,8) | 20,7 | (1,0) | 24,8 | (0,6) | 39,0 | (1,9) |

Los quintiles están definidos a nivel internacional, para incluir al 20% de los participantes de PISA en cada quintil; dentro de cada muestra nacional, la proporción puede, por lo tanto, diferir del 20%.
Fuente: OCDE, Base de Datos PISA 2022, tablas I.B1.4.6 e I.B1.4.8. https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-i_53f23881-en

tiene 440 puntos en Matemáticas significativamente por encima de los 427 de la OCDE. En definitiva, los estudiantes españoles de mayor nivel socioeconómico tienen peores resultados que el alumnado de la OCDE en sus mismas circunstancias. Y, por el contrario, los estudiantes de nuestro país de menor nivel socioeconómico tienen mejores resultados que sus compañeros de la OCDE con el mismo contexto. Reino Unido destaca por su buen desempeño en equidad. El quintil de menor nivel socioeconómico, y eso que es sólo una proporción del 1,9% del que se podría esperar un muy mal desempeño, tiene 422 puntos en Matemáticas en PISA 2022, un resultado significativamente por encima de la media de la OCDE.

**Tabla 6. Rendimiento en Matemáticas por quintiles internacionales de nivel socioeconómico.**

| | | | Rendimiento en Matemáticas, por quintiles internacionales de estatus socioeconómico | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Todos los estudiantes | | Quintil inferior | | Segundo quintil | | Quintil medio | | Cuarto quintil | | Quintil superior | |
| | Punt. | E.E. | Punt. | E.E. | Punt. | E.E. | Punt. | E.E. | Punt. | E.E. | Punt. | E.E. |
| España | 473 | (1,5) | 412 | (3,9) | 440 | (2,3) | 448 | (1,8) | 473 | (1,8) | 511 | (1,8) |
| **OCDE** | 472 | (0,4) | 396 | (1,5) | 427 | (0,7) | 449 | (0,6) | 474 | (0,5) | 517 | (0,5) |
| Francia | 474 | (2,5) | 395 | (6,7) | 421 | (3,7) | 441 | (2,7) | 474 | (2,7) | 526 | (2,9) |
| Alemania | 475 | (3,1) | 406 | (5,9) | 435 | (3,9) | 458 | (3,4) | 482 | (3,5) | 534 | (3,4) |
| Italia | 471 | (3,1) | 400 | (6,3) | 432 | (3,4) | 454 | (3,4) | 476 | (3,1) | 512 | (4,9) |
| Reino Unido | 489 | (2,2) | 422 | (11,9) | 452 | (4,5) | 469 | (3,5) | 484 | (3,0) | 529 | (3,7) |
| Estados Unidos | 465 | (2,5) | 413 | (8,9) | 423 | (6,3) | 430 | (3,5) | 456 | (4,5) | 507 | (4,2) |

Los quintiles están definidos a nivel internacional, para incluir al 20% de los participantes de PISA en cada quintil; dentro de cada muestra nacional, la proporción puede, por lo tanto, diferir del 20%.
Fuente: OCDE, Base de Datos PISA 2022, tablas I.B1.4.6 e I.B1.4.8. https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-i_53f23881-en

La tabla 7 del Anexo aporta más indicadores de equidad educativa. La columna 3 muestra la fortaleza de la relación entre el índice de estatus económico, social y cultural y el rendimiento educativo (en este caso, en Matemáticas). Cuanto mayor sea el porcentaje de la varianza en el rendimiento en Matemáticas que es explicado por ese nivel socioeconómico, menor será la equidad. Lo ideal sería que no existiera ninguna relación entre el índice de estatus económico, social y cultural y el rendimiento educativo, y que fueran dos aspectos independientes.

A este respecto, Francia y Alemania serían los países con menor equidad, mientras que Reino Unido, Italia y España tienen mayor igualdad de oportunidades, que es la misma conclusión que se alcanzaba con los resultados de PISA 2018. La siguiente columna muestra la diferencia en los resultados en Matemáticas entre el cuartil de mayor nivel socioeconómico y el cuartil inferior. Los estudiantes socioeconómicamente favorecidos obtuvieron en promedio 93 puntos más en Matemáticas que los estudiantes desfavorecidos en los países de la OCDE. Francia y Alemania, son de nuevo los países con más diferencias y, por tanto, con peores resultados en relación con la igualdad de oportunidades. El capítulo 5.b incide en las diferencias señaladas en la columna 4 y muestra indicadores de equidad por Comunidades Autónomas españolas. Reino Unido, España e Italia de nuevo destacan por disponer de una mayor proporción de estudiantes acadé-

**Tabla 7. Indicadores de disparidades socioeconómicas y rendimiento académico.**

| | Puntuación media en Lectura en PISA 2022 | Índice de Cobertura: cobertura de la población de 15 años | Fortaleza: Porcentaje de varianza en el rendimiento en Matemáticas explicado por ESCS ($R^2$) | Diferencia entre estudiantes favorecidos y desfavorecidos en Matemáticas | Porcentaje de estudiantes desfavorecidos que son académicamente resilientes |
|---|---|---|---|---|---|
| | Media | (2) | (3) % | (4) Diferencias resultado | (5) % |
| Francia | 474 | 0,93 | 21,5 | 113 | 7,4 |
| Alemania | 475 | 0,92 | 18,7 | 111 | 9,5 |
| Italia | 471 | 0,87 | 13,5 | 85 | 11,3 |
| España | 473 | 0,90 | 14,2 | 86 | 11,7 |
| Reino Unido | 489 | 0,97 | 11,0 | 86 | 15,2 |
| Estados Unidos | 465 | 0,86 | 14,9 | 102 | 10,6 |
| OCDE | 472 | – | 15,5 | 93 | 10,2 |

Nota: ESCS se refiere al índice PISA de estatus económico, social y cultural. Un estudiante socioeconómicamente favorecido (desfavorecido) es un estudiante en el cuarto superior (inferior) de ESCS en su propio país/economía. Los estudiantes académicamente resilientes son estudiantes desfavorecidos que obtuvieron puntuaciones en el cuartil superior de rendimiento en Lectura entre los estudiantes de su propio país.
Fuente: PISA 2022 Database, tablas I.B1.4.1, I.B1.4.3 y I.B1.5.19. https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-i_53f23881-en

micamente resilientes: alumnos desfavorecidos que obtienen puntuaciones en el cuartil superior de rendimiento en Matemáticas entre los estudiantes de su propio país. En suma, los indicadores de equidad en educación de Francia y Alemania son peores que los del promedio de la OCDE, mientras que Reino Unido, Italia y España ofrecen mayor igualdad de oportunidades a sus estudiantes. Estados Unidos, finalmente, se encuentra en la media de la OCDE en cuánto a equidad en educación. La desigualdad en Alemania podría tener que ver con el hecho de la existencia de itinerarios muy pronto en la educación secundaria (a los 10 años se hace la primera elección educativa, *Table II.4. Snapshot of selecting and grouping students*, OCDE, 2023), de modo que una parte importante de los estudiantes rezagados o de entornos desfavorecidos no continúan con los estudios que conducen a la universidad. Por lo demás, Reino Unido, contradiciendo la imagen de una educación elitista, ha consolidado en PISA 2022 los buenos resultados tanto en calidad como en equidad de ediciones anteriores. La OCDE resalta en el Volume I de PISA que el sistema educativo de Reino Unido es, junto con los de Canadá, Dinamarca, Finlandia, Irlanda, Japón, Corea y Letonia, uno de los más equitativos entre los países desarrollados.

PISA 2022 proporciona también información sobre el índice de aislamiento, que mide el grado en que ciertos tipos de estudiantes (por ejemplo, estudiantes desfavorecidos) están aislados de todos los demás tipos de estudiantes o de un grupo específico de estudiantes (por ejemplo, estudiantes favorecidos), basado en los centros educativos a los que asisten. Es decir, que el índice aumenta con la concentración de los estudiantes de un grupo en un número limitado de escuelas, con un rango de variación de 0 a 1, donde 0 corresponde a una situación sin ningún tipo de segregación y 1 a un aislamiento o segregación total. Este índice tiene como objetivo medir las oportunidades de interacción social entre diferentes grupos de estudiantes dentro de un centro educativo. Los compañeros de clase y del centro educativo pueden tener una fuerte influencia entre sí, para bien y para mal (véase el capítulo 3c de este libro y OCDE, 2023). Así, España tiene un índice de aislamiento de los estudiantes desfavorecidos (definidos como aquellos en el peor cuartil, de la distribución nacional del índice socioeconómico) de 0,14, significativamente por debajo del promedio de la OCDE de 0,18, lo que constituye una señal más de que el sistema educativo español tiene un buen comportamiento en cuanto a equidad. El índice de aislamiento de los estudiantes socioeconómicamente favorecidos (definidos como aquellos en el

mejor cuartil de la distribución nacional del índice socioeconómico) de España es de 0,18, muy similar al de la OCDE de 0,19.

Otros aspectos relacionados con igualdad de oportunidades son la diferencia de resultados entre chicas y chicos, con las primeras teniendo una ventaja de 24 puntos en Lectura en la OCDE (25 en España), mientras que los segundos tienen 9 puntos más en Matemáticas (10 en España). En Ciencias, la diferencia de rendimiento entre chicos y chicas no es significativa en la OCDE (en España sí lo es, pero de sólo 5 puntos). El alumnado nativo registró en PISA 2022, 29 puntos más que los estudiantes inmigrantes en Matemáticas en los países de la OCDE, pero solo cinco puntos más una vez que se tuvo en cuenta el nivel socioeconómico y el idioma hablado en casa.

**Conclusión**

Las evaluaciones de dos organismos internacionales diferentes (IEA y OCDE), que analizan niveles educativos distintos (Primaria y Secundaria), con metodologías heterogéneas (PISA evalúa un concepto más amplio como el de las competencias mientras que TIMSS y PIRLS se centran más en el currículo y los conocimientos) llegan a la misma conclusión de que Matemáticas es la asignatura pendiente en España. La mejora de las competencias Matemáticas de los alumnos españoles es un objetivo alcanzable.

La competencia en Matemáticas es un determinante importante de los salarios tanto en la OCDE como en España. Hanushek et al. (2015) muestran, que las personas con más competencias Matemáticas tienen mayores salarios. Como se puede ver en PIAAC (una evaluación como PISA también de la OCDE, pero a la población entre 15 y 65 años), las personas con una desviación estándar más de dominio de las Matemáticas tienen un 17,8% más de salario (22,8% más en el caso de España). En esta misma línea, Hernández y Serrano (2013) encuentran, utilizando también los datos de PIAAC para España, que los conocimientos del individuo tienen un efecto positivo sobre el salario. Las estimaciones que obtienen estos autores indican que cada punto adicional en PIAAC en comprensión lectora aumenta el salario un 0,37%, un efecto muy similar al de un punto adicional en competencia matemática (0,38%). Al considerar simultáneamente ambos tipos de conocimiento, los dos siguen teniendo efectos positivos y significativos sobre el salario, aunque los de las competencias Matemáticas son de mayor magnitud. Por cada punto PIAAC en Matemáticas habría un 0,27%

más de salario, mientras que cada punto de comprensión lectora supondría un 0,13% adicional. Al incluir los años de estudio del individuo, el coeficiente de la competencia lectora deja de ser significativo (y de hecho, es cero), mientras que el de Matemáticas continúa siendo significativo. Es decir, que el dominio de las Matemáticas tiene un impacto distintivo en los salarios con respecto a la formación general.

PISA 2022 muestra que Matemáticas genera ansiedad a los estudiantes, y lo hace con mayor intensidad a los alumnos españoles que a los de la OCDE, y muy en particular a las alumnas (un aspecto que se trata el capítulo 5c). Como señala el Informe del Ministerio de Educación, FP y Deportes (2023) "España se encuentra en el grupo de países donde el alumnado muestra más ansiedad ante las Matemáticas, muy por encima del promedio de los países de la OCDE y del Total UE". La OCDE combina las respuestas de los estudiantes a varias preguntas para crear el índice PISA de ansiedad matemática. Las preguntas son si los estudiantes están de acuerdo o en desacuerdo con afirmaciones sobre si "a menudo me preocupa tener dificultades en las clases de Matemáticas"; "me estreso mucho cuando tengo que hacer deberes de Matemáticas"; "me siento incapaz de resolver un problema de Matemáticas" o "me da miedo suspender Matemáticas". España tiene un índice de 0,37, más del doble que la media de la OCDE (0,17) y la UE (0,17), indicando que las Matemáticas generan más ansiedad en nuestro país.

Puede haber, por tanto, una especie de profecía que se cumple a sí misma. Dado que muchos alumnos españoles parten de la base de que Matemáticas es difícil y es sólo apta para los más inteligentes, a algunos de ellos se les termina dando mal. Hay mucho margen de mejora presentando las Matemáticas como una competencia que no es inaccesible, que es muy importante en la vida diaria y que cada vez lo va a ser más, en un mundo dominado por el big data y la robotización. De hecho, TIMSS muestra que, tanto en Matemáticas como en Ciencias, la variable que mayor impacto tiene sobre la resiliencia es la confianza que el niño o la niña tenga en el aprendizaje del área. En el caso de España el impacto es superior al del total UE por lo que cuanta mayor confianza se genere en los niños y niñas mejores resultados se pueden llegar a conseguir. La mentalidad del crecimiento que se analiza en el capítulo 5d y que hace hincapié en que la inteligencia no es fija, sino que puede crecer a base de esfuerzo, de aprender de los errores y de perseverancia puede ayudar a que una parte del alumnado no vea las Matemáticas como algo inaccesible.

En 2008, la IEA también realizó una evaluación de conocimientos y pedagogía en Matemáticas a los futuros maestros de primaria (estudiantes del último curso de la carrera de educación primaria). Entre las conclusiones para España se encontraba que el 26,8% de los entonces estudiantes de Magisterio había evitado la asignatura de Matemáticas en 1º de Bachillerato. Y un 40,3% adicional no había impartido esta materia en 2º de Bachillerato. Este organismo internacional también encontró que sólo el 6,6% de los créditos de la carrera de Magisterio (ahora Grado en Educación Primaria) se dedicaba a Matemáticas. Es decir, que otro de los factores que puede explicar el déficit de las competencias en Matemáticas de nuestros alumnos podría encontrarse en la escasa relevancia que tiene esta asignatura en la formación inicial de los maestros en Primaria. Tanto en el estudio que realizó la IEA en 2008 (TEDS-M, INEE 2012) como ahora en TIMSS 2019 se muestra que los docentes españoles tienen más formación en el área de los procedimientos pedagógicos para las Matemáticas que en los contenidos matemáticos en sí mismos. Equilibrar ambos aspectos, muy importantes y que se refuerzan entre sí, puede ser también otro de los aspectos que podría mejorar la preparación de los docentes en Matemáticas en Primaria.

**Bibliografía**

Betthäuser, B.A., Bach-Mortensen, A.M. y Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the Covid-19 pandemic. *Nature Human Behaviour, 7*, 375-385. https://doi.org/10.1038/s41562-022-01506-4

Hanushek, E. A., Schwerdt, G., Wiederhold, S. y Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review. 73*(C), 103-130. DOI: 10.1016/j.euroecorev.2014.10.006

Hanushek, E. y Woessmann, L. (2020). The economic impacts of learning losses. *OECD Education Working Papers*, *225*, OECD Publishing, Paris. https://doi.org/10.1787/21908d74-en

Instituto Nacional de Evaluación Educativa (INEE) (2012). *TEDS-M Informa Español. Estudio internacional sobre la formación inicial en Matemáticas de los maestros*. Ministerio de Educación y FP, Madrid. https://www.educacionyfp.gob.es/dctm/inee/internacional/teds-mlinea.pdf?documentId=0901e72b8143866e

Ministerio de Educación, FP y Deportes (2010). *Evaluación General de Diagnóstico 2009. Educación Primaria. Cuarto curso. Informe de resultados*. Ministerio de Educación, FP y Deportes, Madrid.

Ministerio de Educación, FP y Deportes (2011). *Evaluación General de Diagnóstico 2009. Educación Secundaria Obligatoria. Segundo curso. Informe de resultados.* Ministerio de Educación, FP y Deportes, Madrid.

Ministerio de Educación, FP y Deportes (2023). *PISA 2022 Programa para la Evaluación Internacional de los Estudiantes Informe español*. Ministerio de Educación, FP y Deportes, Madrid. https://www.educacionyfp.gob.es/dam/jcr:91f26ac3-0a3b-4efa-b2ce-a5d791229f4d/pisa-2022-informe-completo-digital-low.pdf

OECD (2020). Education Policy Outlook in Estonia. *OECD Education Policy Perspectives, 13*, OECD Publishing, Paris. https://doi.org/10.1787/9d472195-en

OECD (2020). Education Policy Outlook in Ireland. *OECD Education Policy Perspectives, 18*, OECD Publishing, Paris. https://doi.org/10.1787/978e377b-en

OCDE (2023), *PISA 2022 Results (Volume I): The State of Learning and Equity in Education, PISA*. OECD Publishing, Paris. https://doi.org/10.1787/53f23881-en

Sanz, I. y Tena, J.D. (2023). The Impact of the Covid-19 Pandemic on Education Learning, en Sainz, J. y Sanz, I. (eds.) *Addressing Inequities in Modern Educational Assessment*. Springer, Cham. https://doi.org/10.1007/978-3-031-45802-6_2

# Epílogo
# LA ECONOMÍA DE LA EDUCACIÓN A TRAVÉS DEL LENTE DE ESPAÑA: LECCIONES Y POSIBILIDADES

Matthew A. Kraft
Brown University

Este manual ofrece a los lectores un resumen actualizado de la Economía de la Educación con aplicación a España. La colección de capítulos contiene conocimientos profundos e innovadores sobre la situación del sistema educativo español. El libro brinda un contexto importante sobre el desarrollo de la educación en España a lo largo del tiempo y en comparación a otros países.

Sin embargo, los capítulos del manual abordan mucho más que la historia de la educación en España. Los autores proporcionan una introducción accesible a teorías y principios económicos, combinándolos con métodos econométricos innovadores aplicados a la Economía de la Educación. Los autores demuestran la eficacia de estas herramientas para analizar políticas educativas e identificar reformas prometedoras que amplíen el acceso y mejoren la calidad de la educación. Investigadores de todas partes del mundo pueden aprender mucho de estos autores, quienes arrojan luz sobre los puntos ciegos que todos tenemos dado nuestra diferente formación y experiencias individuales.

Como los mejores textos académicos, el manual está basado en un resumen de hechos empíricos y análisis rigurosos usando métodos descriptivos y causales. Leyendo el libro, ganamos una familiaridad profunda sobre los factores que influyen en los resultados del sistema educativo español y las posibilidades que existen para mejorarlo.

El libro destaca logros claves del sistema educativo español y grandes dificultades que requerirán del esfuerzo colectivo por parte del Gobierno, los profesores y las familias para superarlos. España ha tenido mucho éxito en los últimos años reduciendo el abandono escolar y aumentando las tasas de graduación. Esta inversión en el capital humano de la población tendrá beneficios individuales para millones de estudiantes, así como para la Economía y la sociedad en su conjunto. Más allá de eso, sabemos que las inversiones que aumentan el nivel

educativo de una nación son inversiones en la democracia y en la participación ciudadana.

España también está desarrollando un sistema de escolarización infantil amplío y gratuito. Esta es unas de las políticas más eficaces que existen en todos los países para disminuir la desigualdad de los rendimientos educativos. Una amplía literatura demuestra que los centros educativos infantiles que emplean profesores bien pagados y capacitados que enseñan currículos de alta calidad para incrementar las competencias cognitivas y socioemocionales, tienen efectos positivos muy grandes en los resultados de los alumnos a lo largo de sus vidas.

El compromiso de la educación bilingüe y trilingüe en España también tiene enorme importancia. Vivimos en un mundo cada vez más globalizado que requiere conocimientos de idiomas extranjeros. Esfuerzos para ofrecer más oportunidades para inmersión lingüística de los alumnos de Educación Infantil y Primaria son invaluables. La educación bilingüe en edades tempranas proporciona el desarrollo y destreza cognitiva a los alumnos para aprender idiomas extranjeros en la etapa en la que aprenden más rápidamente.

Este manual también destaca desafíos educativos antiguos y profundos. En lo más alto de la agenda está el desafío de aumentar la efectividad y eficiencia de las escuelas, al tiempo que se reducen las grandes desigualdades educativas existentes. Las pruebas internacionales demuestran que hay un gran potencial para mejorar el rendimiento académico de alumnos españoles en comparación a otros países. Un camino claro para lograrlo es dedicarse a mejorar la calidad de la enseñanza y el entorno del aprendizaje. Este objetivo requerirá un compromiso de la formación del profesional y un cambio de la cultura en las escuelas para reconocer diferencias de efectividad entre profesores. También se requiere una reconsideración de las políticas de asignación de profesores a centros para disminuir los altos y preocupantes niveles de rotación.

En España, así como en cada parte del mundo, hay que priorizar métodos pedagógicos que capaciten al alumnado para ser pensadores autónomos y críticos. Es esencial modernizar el estilo de didáctica para ser más activo y menos pasivo para preparar a los alumnos para un mundo en constante cambio y con avances tecnológicos acelerados. La educación también tiene un papel clave en el desempeño de las competencias socioemocionales y el bienestar mental del alumnado. Desarrollando entornos de aprendizaje donde los alumnos se sienten que pertenecen y son parte de la comunidad escolar es crítico para la formación integral del alumno.

Para aprovechar al máximo el potencial del sistema educativo español, España tendrá que abordar las disparidades que se hacen evidentes cuando se analiza el rendimiento escolar en función del estatus socioeconómico, país de origen y género de los estudiantes. Se necesitan cambios en la financiación de la educación para disminuir las brechas de recursos entre regiones y centros, y poder así crear un futuro más justo y equitativo para todos. Además de una distribución más equitativa de gastos por alumno, se requiere también una creencia compartida de que todos estamos mejor cuando no se deja a nadie atrás.

Las herramientas y estructuras de la Economía de la Educación repasadas en este manual ofrecen una guía de cómo se puede entender y analizar los múltiples factores interconectados que forman el sistema educativo español. Sin embargo, las posibilidades para utilizarlas serán limitadas sin un compromiso para vincular datos administrativos y abrir acceso a los investigadores y administradores. También requiere mayor atención el análisis coste-eficacia de cada euro gastado en la financiación de todos los niveles del sistema educativo: Infantil, Primaria, Secundaria y Terciaria. Podemos utilizar este libro como guía para entender el pasado y el presente del sistema educativo, y también para imaginar un futuro mejor para la próxima generación de alumnos españoles.