

Informe Técnico:

Análisis cuantitativo de Escuelas Resilientes de Educación Primaria en Canarias y Secundaria en Catalunya.

Autores:

Lucía Cobreros Vicente

Lucas Gortazar

Enero de 2026

Acerca de este informe

Este documento se enmarca en el proyecto **Escuelas Resilientes**, promovido por el Observatorio Social de la Fundación “la Caixa”, Save the Children España y EsadeEcPol. El objetivo del proyecto es doble: en primer lugar, se pretende identificar, estimando modelos de valor añadido, la contribución de los centros educativos que, escolarizando a un elevado porcentaje de alumnado vulnerable, logran que su alumnado progrese por encima de lo esperado; en segundo lugar, mediante un exhaustivo trabajo de campo, se trata de caracterizar y comprender qué hacen estos centros educativos diferente para conseguir estos resultados diferenciales.

Concretamente, este informe es la versión técnica y ampliada del análisis cuantitativo que acompaña al Policy Brief “*Escuelas Resilientes: Excelencia en centros educativos desafiados por la pobreza. Aprendizajes de Canarias y Catalunya*”. Su objetivo es profundizar en la identificación de la vulnerabilidad de los centros educativos, la selección y comparación de modelos de valor añadido y la caracterización de los centros resilientes utilizando datos para todo el alumnado de carácter administrativo y de cuestionarios de contexto. Se abordan dos casos de estudio —Educación Primaria en Canarias y Educación Secundaria en Catalunya— con el propósito de ofrecer un análisis en profundidad de cada contexto, sin buscar comparaciones directas entre ellos.

1. Revisión de la literatura

La literatura sobre eficacia escolar ha mostrado de manera consistente que las escuelas y el profesorado ejercen un efecto sustancial en los resultados académicos del alumnado, especialmente de aquellos con mayores desventajas socioeconómicas (Reynolds et al. 1996; Scheerens, 2004). Este planteamiento contrasta con visiones previas que concluían que la escuela ofrecía un valor limitado en términos de mejora educativa. Se entiende, de manera general, que una escuela es eficaz cuando su estudiantado obtiene logros superiores a los esperados, dadas sus condiciones de partida. Dentro de este campo, merece atención especial la identificación de lo que en este informe se denominan “escuelas resilientes”, es decir, aquellas que exhiben un progreso elevado atendiendo a un alumnado proveniente de entornos especialmente desfavorecidos. La relevancia de estas escuelas radica en que ilustran la posibilidad de compensar parcial o totalmente las desventajas asociadas al origen familiar mediante políticas y prácticas institucionales adecuadas.

¿Cómo se mide la eficacia escolar?

En la literatura se han desarrollado métodos estadísticos para aislar la contribución real del centro, teniendo en cuenta factores como el rendimiento previo o las características socioeconómicas del alumnado: los Modelos de Valor Añadido (VAM). Estos enfoques parten de la idea de que comparar “resultados brutos” (por ejemplo, la media de calificaciones o la tasa de graduación) puede llevar a juicios poco justos sobre el alumnado y el centro si no se tiene en cuenta que las escuelas atienden poblaciones con niveles de partida muy distintos (Thomas, 1996; Saunders, 1999). Para evitarlo, los VAM tienen en cuenta medidas de rendimiento previo del alumnado (u otras aproximaciones al logro inicial o previo al momento en el que se mide el resultado), de forma que la eficacia se pueda interpretar como el progreso *adicional* atribuible a la acción escolar, una vez descontado el punto de partida. En este sentido, el concepto de valor añadido refleja precisamente la ganancia educativa neta que aporta el centro.

Aunque existe consenso sobre la necesidad de tener en cuenta el rendimiento previo para estimar eficacia escolar, no es así cuando se trata de qué otros factores deben incluirse en estos modelos. La mayoría de las investigaciones coinciden en que el rendimiento previo, idealmente en la misma asignatura, es indispensable para medir la contribución escolar (Marks, 2017, 2021; Leckie & Prior, 2022; IFS, 2023). No obstante, algunos trabajos recomiendan añadir también características sociodemográficas y de composición del alumnado para afinar la estimación del verdadero efecto del centro, reduciendo así la posibilidad de favorecer o perjudicar a escuelas en contextos muy distintos (Timmermans et al., 2011; Muñoz-Chereau & Thomas, 2016; Leckie & Goldstein, 2019). Frente a ello, autores como Marks (2021) o Prior et al. (2021) advierten que incluir numerosas covariables puede complicar la interpretación de los resultados y, a la vez, conllevar problemas de calidad de datos o de especificación del modelo.

Esta controversia ha llevado a varios investigadores a comparar empíricamente los resultados obtenidos al añadir diferentes covariables. Estos trabajos ilustran la

variabilidad en la clasificación de las escuelas cuando se van incorporando distintos elementos, como las características sociodemográficas del alumnado (Timmermans & Thomas, 2015; Levy et al., 2022). Un caso notable es el de Inglaterra, donde el modelo de rendición de cuentas oficial, *Progress 8*, controla exclusivamente el rendimiento previo de los estudiantes. Si bien algunos estudios (Leckie & Goldstein, 2019) han planteado la preocupación de que este enfoque pueda penalizar a las escuelas con una alta proporción de alumnado en situaciones de vulnerabilidad, otros trabajos recientes (IFS, 2023) demuestran que el indicador es razonablemente robusto y que ajustes adicionales por características sociodemográficas del alumnado no mejoran sustancialmente la medición de la eficacia. Así, los autores sugieren que los modelos se elijan teniendo en cuenta un equilibrio entre la validez del indicador y la sencillez de su interpretación. Algo similar ocurre al considerar factores de escuela como el rendimiento previo promedio del centro, donde el potencial beneficio de controlar el efecto de la “composición escolar” se ve a menudo contrarrestado por la complejidad y la posible inestabilidad en las estimaciones (Marks, 2021).

En paralelo a este debate, ha cobrado relevancia la cuestión de hasta qué punto los modelos de valor añadido permiten una interpretación causal de la eficacia escolar. Angrist et al. (2021) muestran que, en contextos donde la admisión a los centros se resuelve mediante sorteos de plazas o umbrales de notas, es posible identificar de forma convincente efectos causales de la escuela sobre el rendimiento. Deming (2014) compara resultados de VAM con los obtenidos en situaciones de asignación aleatoria y concluye que, cuando se controla por rendimiento previo y algunas covariables básicas, las estimaciones de valor añadido son en general robustas y llevan a conclusiones similares que los métodos cuasiexperimentales. Sin embargo, Reardon & Raudenbush (2009) advierten que la validez causal de estos modelos depende de supuestos exigentes, entre ellos que todos los estudiantes puedan ser asignados a cualquier centro. Dada la falta de consenso, los resultados que se presentan deben entenderse como medidas de eficacia relativa y no como efectos causales estrictos; sin embargo, los VAM siguen constituyendo una herramienta sólida y ampliamente utilizada para caracterizar el desempeño de los centros.

La aplicación de los modelos de Valor Añadido

La aplicación de modelos de valor añadido (VAM) ha tenido un recorrido mucho más amplio en otros sistemas educativos que en España. En Australia, el portal My School publica anualmente indicadores de valor añadido a partir de las pruebas NAPLAN, con el objetivo de ofrecer información comparable sobre el progreso del alumnado y la eficacia de los centros, además de identificar escuelas que logran resultados sobresalientes en contextos difíciles (Australian Curriculum, Assessment and Reporting Authority, 2021). En Chile, diversas investigaciones han mostrado el potencial de estos indicadores para construir un sistema de rendición de cuentas más justo, capaz de reconocer a los centros que aportan mayor crecimiento educativo a sus estudiantes, aunque su implementación sigue siendo un desafío pendiente (Muñoz-Chereau, Anwandter & Thomas, 2020). En Inglaterra, los modelos de valor añadido se han convertido en el eje de la política de

evaluación externa: tras la experiencia con el contextual value added y el expected progress, desde 2016 se emplea la medida Progress 8 como indicador oficial de la contribución de cada escuela al progreso académico del alumnado (Department for Education, 2020; Leckie & Goldstein, 2017). En Estados Unidos, numerosos estados han incorporado medidas de valor añadido tanto para evaluar la eficacia docente como para clasificar escuelas en el marco de políticas de accountability de alta consecuencia, aunque la literatura crítica advierte de sus limitaciones técnicas y de sus posibles efectos no deseados sobre la enseñanza y el currículo (Amrein-Beardsley, 2014; Koretz, 2017).

La evidencia internacional muestra que estos modelos pueden ser útiles para complementar las evaluaciones externas y orientar la mejora educativa. En particular, estudios recientes han documentado que las estimaciones de valor añadido poseen cierta validez predictiva: por ejemplo, en Estados Unidos se ha encontrado que las puntuaciones de eficacia escolar ajustadas por valor añadido ayudan a anticipar la matriculación universitaria, la persistencia en la educación superior e incluso el desempeño en el primer año académico (Chetty, Friedman & Rockoff, 2014). De manera similar, en Australia se ha destacado el potencial de estos indicadores para reconocer a escuelas altamente eficaces incluso en contextos de vulnerabilidad, y usarlos como base para estudiar y difundir buenas prácticas (Lamb et al., 2015).

Qué dice la evidencia sobre las escuelas eficaces

La literatura clásica en eficacia escolar subraya el papel decisivo del liderazgo con altas expectativas, la cohesión docente, un proyecto educativo compartido y la atención al progreso individual como factores de éxito. Revisiones extensas (Murillo, 2022) destacan la importancia de un liderazgo escolar sólido y con altas expectativas, la cohesión y formación del profesorado, la existencia de un proyecto educativo bien definido y compartido por la comunidad, y la atención individualizada al progreso de cada alumno. Estos resultados son consistentes con los hallazgos de Lizasoain et al., (2014) para el País Vasco. En las llamadas escuelas resilientes, es decir, aquellas que, actuando en contextos de desventaja socioeconómica consiguen con un progreso elevado de su alumnado, resultan especialmente relevantes la cultura de altas expectativas (Bellei et al., 2004), el uso sistemático de datos para planificar la enseñanza, la implicación del claustro y el clima de respeto y apoyo mutuo (Pinskaya et al., 2019). Estos rasgos, combinados con una gestión institucional eficaz, pueden contrarrestar en parte la influencia negativa de un bajo nivel socioeconómico, lo que explica su frecuente mención como ejemplos valiosos para la formulación de políticas educativas.

En relación con quién se beneficia más de asistir a escuelas eficaces, la evidencia apunta a que los estudiantes con menor ventaja inicial suelen experimentar ganancias más pronunciadas a largo plazo. Jackson et al. (2023) muestran que los estudiantes más vulnerables son los que obtienen un mayor descenso en detenciones policiales, abandono escolar y mayores incrementos en graduación al asistir a centros eficaces. Por su parte, Mbekeani (2023) analiza institutos públicos de Massachusetts y concluye que, para los jóvenes de bajos ingresos, asistir a un centro de la parte alta de la distribución de eficacia conlleva mejoras significativas en el acceso a la educación superior y en los salarios a la

edad de 30 años. Estos resultados coinciden con otros que subrayan la relevancia de la elección de escuela en términos de movilidad social y de disminución de brechas de rendimiento (Britton et al., 2023).

Resultados académicos y socioemocionales

Aunque el foco de gran parte de la literatura y de las evaluaciones a gran escala permanece en los logros académicos, algunos estudios han resaltado la necesidad de abarcar dimensiones no cognitivas (Beuerman et al. 2023; Loviglio, 2023). En esta línea, trabajos como los de Kraft (2019) y Jackson et al. (2020) muestran que los docentes y las escuelas pueden influir también en variables socioemocionales —por ejemplo, la motivación, la mentalidad de crecimiento o la perseverancia— y que dichos efectos a menudo no están captados por los indicadores habituales de rendimiento en pruebas estandarizadas. Asimismo, Jackson et al. (2023) evidencian que cuando las escuelas mejoran tanto los resultados de las pruebas de diagnóstico como las competencias socioemocionales, se observan repercusiones particularmente positivas en la graduación, el acceso a la universidad y la reducción de comportamientos de riesgo, sobre todo entre el estudiantado menos aventajado. Esto sugiere que el valor añadido de las escuelas podría ser más amplio de lo que miden las evaluaciones académicas, si bien la mayoría de los sistemas de “rendición de cuentas” actuales siguen restringidos a la información sobre logros en exámenes. No obstante, para estudios centrados únicamente en el rendimiento académico, estas contribuciones invitan al menos a reconocer la posibilidad de que las “buenas” prácticas y la cultura escolar generen beneficios que van más allá de las pruebas.

2. Metodología: cómo medir la eficacia escolar y la complejidad en los centros educativos

2.1. Estimando la eficacia escolar: los modelos de valor añadido

Los Modelos de Valor Añadido (VAM) permiten estimar la contribución específica de los centros educativos al aprendizaje de su alumnado, una vez consideradas las diferencias iniciales y las características del contexto. De esta forma, se evita sobrevalorar (o infravalorar) el desempeño de centros con condiciones especialmente favorables (o adversas) y se genera una estimación del valor añadido más robusta. Para estimar el valor añadido, lo primero que hay que tener en cuenta es que los datos de escuelas tienen una estructura jerárquica: el alumnado está dentro de centros, y quienes comparten centro tienden a parecerse más entre sí que con quienes estudian en otros centros. Así, si se aplicara una regresión lineal simple —asumiendo independencia entre todas las observaciones— se correría el riesgo de atribuir al azar o a características individuales efectos que en realidad corresponden a la escuela.

Por este motivo, en este estudio se emplean modelos mixtos (también conocidos como de efectos aleatorios o multinivel), que incorporan la estructura jerárquica de los datos. En estos modelos, cada centro cuenta con un intercepto aleatorio que recoge la parte de variabilidad de los resultados individuales atribuible a su influencia particular. Este enfoque ofrece, además, dos ventajas clave además de la anterior: (i) permite aislar qué parte de la variación en los resultados se debe a diferencias entre centros y cuánta a diferencias entre estudiantes; y (ii) evita que los resultados de los centros más pequeños parezcan extremos solo porque tienen pocos alumnos. En la práctica, el modelo “suaviza” las estimaciones de esos casos y las acerca a la media general, de modo que no se sobreinterpreten diferencias que podrían deberse al azar.

En este informe, donde las variables de interés son los resultados del alumnado en las evaluaciones de diagnóstico —4º de ESO en Catalunya y 6º de Primaria en Canarias—, el objetivo es, mediante modelos mixtos, descomponer la variación en los resultados y estimar qué parte puede atribuirse a la acción de los centros educativos, una vez considerados:

- Los resultados de las evaluaciones previas (6º de Primaria en Catalunya; 3º de Primaria en Canarias).
- Las características socioeconómicas y demográficas del alumnado.
- Información socioeconómica y de rendimiento agregada a nivel de centro.

A continuación, se representa el modelo —en su versión más extendida— para, posteriormente, ir analizando diferentes versiones del mismo:

$$Y_{ijt} = \beta_0 + \alpha Y_{ij(T-t)} + \gamma' X_{i,j} + \delta' Z_j + u_{0j} + \epsilon_{ij}$$

Donde:

Y_{ijt} : resultado del estudiante i en el centro j en el momento T (por ejemplo, resultado en 4º ESO).

Y_{ijt-t} : resultado del estudiante i en el centro j en el momento $T - t$ (por ejemplo, resultado en 6º Primaria).

X_{ij} : vector de características del estudiante i (género, edad, origen migrante, nivel socioeconómico de los progenitores, etc.).

Z_j : vector de características del centro (titularidad del centro, nivel socioeconómico agregado, región en la que se encuentra, rendimiento agregado a nivel de centro, etc.).

u_{0j} es el intercepto aleatorio a nivel de centro.

e_{ij} es el término de error individual.

Descripción de los cinco modelos a estimar

En línea con el trabajo de Leckie y Prior (2022), se describen los cinco modelos más utilizados en la literatura de eficacia escolar, resumidos en la Tabla 1.

⇒ Modelo bruto o sin ajustes: Raw

El modelo Raw se concibe como la aproximación más simple posible: se toman los resultados académicos del alumnado al final de la etapa sin introducir ningún tipo de ajuste estadístico. Este método equivale, en la práctica, a comparar las medias de los centros en los exámenes finales, de modo que los colegios o institutos cuyos estudiantes obtienen las calificaciones más altas aparecen como “los mejores”. Sin embargo, la literatura sobre eficacia escolar (por ejemplo, Goldstein, 1997; Raudenbush & Willms, 1995) señala, desde hace décadas, que este enfoque no diferencia entre lo que “aporta” realmente el centro y las diferencias iniciales de nivel que traen los estudiantes al comenzar la etapa educativa. Así, el modelo Raw suele descartarse como un modelo de valor añadido, si bien se incluye en el análisis para resaltar el rol de las variables adicionales.

⇒ Modelo de Valor Añadido: VA

El modelo VA (Value-Added) constituye la primera forma de modelización del valor añadido. Su fundamento radica en introducir la nota previa del alumnado al inicio de la etapa como covariable principal, a fin de aislar el progreso logrado durante el periodo escolar analizado. En este caso, el coeficiente de la nota previa indica la relación general entre el nivel inicial y el rendimiento final, y se responsabiliza a los centros por la parte de desviación (positiva o negativa) que presentan sus estudiantes respecto a esta relación media. Este modelo, ampliamente difundido en la investigación y en algunos sistemas de evaluación, clasifica a los centros según si su valor añadido se sitúa significativamente por encima, por debajo o alrededor de la media.

⇒ Modelo de valor añadido contextualizado: CVA - A

El modelo CVA - A (Contextualised Value-Added) amplía el modelo VA al incluir, además del rendimiento previo, una serie de características socioeconómicas y demográficas de los estudiantes. Diversos trabajos empíricos (Ballou et al., 2004; Leckie & Goldstein, 2019) argumentan que variables como la situación familiar o el nivel cultural de origen condicionan el aprendizaje adicional que puede brindar un centro, de modo que la eficacia escolar no debería juzgarse sin considerar estos factores. En el presente estudio, se incluirán como variables de contexto el género y la edad (como aproximación de la

repetición), la condición de estudiante de origen migrante, las necesidades especiales de apoyo educativo y la vulnerabilidad económica. Así, se buscará contextualizar los resultados teniendo en cuenta factores que el centro no controla directamente, evitando penalizar (o favorecer) a centros con un alumnado especialmente desfavorecido en cuanto a la evaluación de su progreso educativo.

⇒ **Modelo de valor añadido contextualizado con efecto de pares: CVA - B**

Este modelo parte del CVA-A y añade un paso más al incorporar, además de las variables individuales, el promedio de la nota previa del centro, con el fin de capturar los conocidos “efectos de composición” o “efectos de pares” (Timmermans & Thomas, 2015). Algunos estudios señalan que la presencia mayoritaria de alumnado con niveles previos altos puede conllevar beneficios adicionales para todo el grupo (Thomas, 2001); no obstante, también se ha debatido la posibilidad de que esta incorporación sobreestime el efecto de seleccionar alumnos o docentes más eficientes en centros con estudiantes de mayor rendimiento (Castellano et al., 2014; Prior et al., 2021). Esto ocurriría porque los pares influyen a través de normas y expectativas académicas más altas, apoyo entre iguales y menor heterogeneidad que permite un ritmo de instrucción más exigente; y, al mismo tiempo, porque los centros con mayor rendimiento previo suelen atraer mejores recursos y docentes, familias más implicadas y climas escolares más favorables, de modo que el modelo puede atribuir a la composición efectos que en realidad provienen de esa selección no observada.

⇒ **Modelo de valor añadido contextualizado con información socioeconómica: CVA - X**

Por último, el modelo CVA-X incorpora factores poco modificables a corto plazo, como el tamaño, la tipología o la ubicación del centro, así como información socioeconómica agregada del alumnado (Keeves et al., 2005). Aunque estos elementos permiten un mayor grado de contextualización, también se ha advertido que pueden conducir a un “ajuste excesivo”, en la medida en que podrían ocultar la identificación de centros con bajo desempeño general (Timmermans et al., 2011).

Análisis de los cinco modelos de eficacia escolar en este informe

Como se explicaba en la sección anterior, en la literatura se han realizado algunos estudios que comparan cómo estas distintas especificaciones alteran la medición de la eficacia escolar y la clasificación de centros (Thomas & Mortimore, 1996; Marks, 2017, 2021; Muñoz-Chereau & Thomas, 2016; Leckie & Prior, 2022). El consenso es amplio respecto a la necesidad de incluir el rendimiento previo, pero no sobre la conveniencia de añadir factores socioeconómicos o de composición. En la práctica, se opta por estrategias intermedias, condicionadas por la disponibilidad de datos, los criterios de equidad y la necesidad de mantener la sencillez en la comunicación de resultados.

En este informe, para los dos casos de estudio que se presentan para Educación Primaria y Secundaria, se estiman los cinco modelos planteados y se comparan sus resultados poniendo el foco en la varianza del intercepto aleatorio (es decir, la varianza de los resultados explicada por los centros), la varianza del error (es decir, la varianza de los resultados explicada por diferencias en el alumnado dentro de cada centro) y el porcentaje de centros que tienen un valor añadido positivo y estadísticamente significativo. Además, se profundiza en cómo varían los centros seleccionados como eficaces en función del modelo; especialmente, en el rol de las variables de contexto.

Tabla 1. Resumen de los VAM considerados en este informe

Modelo	Variables incluidas	Permite...
Raw	Solo resultados finales	Comparar desempeño absoluto entre centros
VA	Resultados finales + nota previa individual	Atribuir a cada centro la parte de progreso neto respecto al nivel inicial del estudiante
CVA - A	VA + características individuales (género, edad, origen, NEAE, vulnerabilidad...)	Ajustar por contexto personal del alumnado y no penalizar a centros con perfiles más desfavorecidos
CVA - B	CVA-A + promedio de la nota previa del centro	Captar efectos de composición/pares en el rendimiento
CVA - X	CVA-B + factores estructurales del centro (ubicación, nivel socioeconómico agregado)	Incorporar variables poco modificables para contextualizar comparaciones entre centros

2.2. Identificación de los centros que atienden a alumnado vulnerable: el índice socioeconómico de los centros

Conocer la eficacia de los centros educativos que consiguen una mejora en los resultados académicos superior a la media en un contexto de complejidad requiere conocer con precisión el contexto socioeconómico del centro. Así, es esencial contar con un indicador robusto de la complejidad socioeconómica del alumnado. Dicho indicador posibilita, por un lado, evitar valoraciones injustas de centros con una alta concentración de vulnerabilidad y, por otro, destacar aquellos que, a pesar de los retos, logran impulsar de manera significativa el rendimiento de sus estudiantes.

La literatura coincide en subrayar una serie de factores socioeconómicos clave que influyen de forma decisiva en los resultados escolares: la educación de los progenitores, su ocupación, el origen del alumnado y el barrio, entre otros. Uno de los más estudiados es el nivel educativo de los progenitores, ampliamente señalado desde hace más de cincuenta años (Coleman et al., 1968) como uno de los predictores fundamentales del rendimiento académico del alumnado. Investigaciones posteriores (Black et al., 2005; Oreopoulos et al., 2006; Carneiro et al., 2013) han robustecido esta conclusión, mostrando cómo la transmisión intergeneracional de la educación —especialmente la

materna— tiende a perpetuar desigualdades en el acceso a títulos universitarios, a ingresos futuros y, en definitiva, a oportunidades de movilidad social ascendente. De esta forma, incluir información sobre la titulación de los progenitores permite detectar a aquellos estudiantes que parten de entornos con menor capital humano y social y que, por lo tanto, podrían requerir mayor apoyo.

El origen migrante, por su parte, también se ha relacionado consistentemente con el rendimiento escolar (Chiswick & DebBurman, 2004; Meunier 2011; Dustmann et al., 2012; Aslund et al., 2012). Van Ewijk y Sleegers (2017), en una revisión de estudios en quince países, muestran que el alumnado de primera o segunda generación procedente de contextos menos desarrollados enfrenta retos añadidos asociados no solo a factores económicos, sino también a la falta de dominio lingüístico, al choque cultural y a las percepciones o expectativas del profesorado. A nivel español, en Catalunya y Canarias, esta cuestión toma especial relevancia: en Catalunya, el 20% de la población entre 15 y 19 años en 2022 procedía de familias migrantes de primera generación y el 23% de los nacimientos en 2006 fueron de madres extranjeras (Idescat, 2023). En Canarias las cifras son similares: el porcentaje de niños menores de 15 años nacidos en el extranjero es del 9% en Canarias y 8,5% en Catalunya (INE, 2022). Estudios como los de Calero et al. (2010) y Cebolla et al. (2016) evidencian cómo el origen migrante afecta al rendimiento en distintas etapas educativas, y más recientemente, Cobreros y Gortazar (2023) confirman que el alumnado migrante, tanto de primera como de segunda generación, sigue obteniendo resultados inferiores en lengua y matemáticas, siendo estas diferencias muy relevantes en Catalunya pero prácticamente inexistentes en Canarias.

En cuanto a la situación socioeconómica del entorno o barrio de residencia de los estudiantes, Chetty et al. (2014) encontraron una fuerte correlación entre crecer en un contexto económicamente desfavorecido y menores posibilidades de movilidad social en la edad adulta. Estos barrios suelen ofrecer menos recursos y, al mismo tiempo, influyen en las redes sociales y en las aspiraciones de los jóvenes, perpetuando la desigualdad educativa. Por consiguiente, incorporar información relativa al nivel socioeconómico de la zona de residencia contribuye a afinar la medida de vulnerabilidad, dado que el contexto territorial puede amplificar o atenuar el capital familiar de origen.

También cabe destacar el papel que juega la ocupación de los progenitores. Más allá de su relación con el nivel educativo y los ingresos, la literatura (OECD, 2010) sugiere que la ocupación parental influye en las aspiraciones y expectativas educativas de las familias. Quienes se encuentran en situaciones laborales más precarias o desempeñan ocupaciones de baja cualificación tienden a presentar mayores dificultades para apoyar el desarrollo académico de sus hijos, incidiendo en su probabilidad de éxito escolar. Estas variables —nivel educativo de los progenitores, origen migrante, ocupación y

características del barrio— constituyen la base de numerosos índices de vulnerabilidad educativa en distintos países.¹

Para resumir y sintetizar toda esta información en un único indicador, resulta especialmente útil el Análisis de Componentes Principales (PCA, por sus siglas en inglés). El PCA es un método estadístico estándar en la construcción de índices que permite manejar conjuntos de datos complejos de forma eficiente, reduciendo la dimensionalidad sin perder información relevante. A través de combinaciones lineales de las variables originales —los llamados componentes principales— se consigue maximizar la variabilidad explicada. El primer componente, que captura la mayor parte de dicha variabilidad, suele tomarse como referencia para generar la medida que en este caso se interpretará como un índice socioeconómico o de complejidad de los centros.²

De este modo, para los dos casos de estudio planteados en este informe, se utilizará la metodología PCA con el fin de aproximar el nivel socioeconómico del alumnado que acude al centro. Las variables elegidas en cada caso serán el resultado del consenso en la literatura, la calidad de la información —evitando variables con muchos valores ausentes, especialmente si están correlacionados con los resultados—, la capacidad para explicar la mayor proporción de la varianza, la claridad interpretativa de las cargas asociadas al primer componente principal y su relación con los resultados académicos del alumnado.

2.3. Los retos metodológicos: valores ausentes y limitación de variables

Como todo estudio de estas características, no está exento de limitaciones, como el uso exclusivo del rendimiento académico, la presencia de valores ausentes o el periodo temporal. La primera se relaciona con el hecho de que, si bien se emplean resultados académicos procedentes de evaluaciones externas y estandarizadas —lo que facilita la comparación entre centros y cohortes—, sería deseable contar también con otras variables que abarquen dimensiones socioemocionales o de bienestar del alumnado (Beurman et al. 2023; Loviglio, 2023), pues el rendimiento escolar no agota la noción de calidad educativa. Además, la creciente tendencia al *teaching to the test* (o enseñar para la prueba) puede acabar sesgando tanto la utilidad formativa de estas pruebas como la verdadera magnitud del progreso logrado por el alumnado, al centrarse en exceso en los aspectos más evaluables y dejar de lado competencias menos formales.

¹ Francia, por ejemplo, utiliza el Índice de Posición Social para captar desigualdades basadas en la formación y la ocupación parental; Alemania cuenta con un Índice Social orientado a combatir la segregación escolar; e incluso a nivel nacional, Catalunya emplea el Index de Complexitat para asignar recursos adicionales a los centros con mayor concentración de alumnado vulnerable. Este consenso internacional avala la pertinencia de incluir dichas dimensiones al evaluar la complejidad de los centros y, en consecuencia, su eficacia relativa.

² Cabe destacar que, antes de aplicar el PCA, deben estandarizarse las variables para que cada una contribuya de manera equitativa y no se vea penalizada o sobrerrepresentada por diferencias en sus escalas o rangos.

En segundo lugar, para que los modelos de valor añadido funcionen con la máxima fiabilidad, es esencial contar con información de contexto lo más completa posible y con un mínimo de valores ausentes (Leckie & Prior, 2022). Lo mismo ocurre en la construcción del índice de complejidad, el cual depende, en gran medida, de la calidad de los datos. Cuando los valores no reportados se concentran en colectivos concretos o están asociados a peores resultados académicos, se pueden introducir sesgos importantes. Adicionalmente, una vez estimados los modelos, caracterizar a los centros que obtienen un valor añadido positivo —frente al resto— requiere de información suficiente. En este informe, cada uno de los casos de estudio cuenta con sus ventajas e inconvenientes: mientras que en Catalunya se dispone de variables de registro administrativo, pero escasas para identificar qué caracteriza a los centros con valor añadido positivo, en Canarias el abanico de información es mayor, pero proviene de encuestas con altos porcentajes de valores ausentes no aleatorios, que correlacionan con características del alumnado y el centro.

En tercer lugar, para que la medición del efecto escuela sea lo más robusta posible, es deseable contar con los resultados en las evaluaciones al inicio y final de etapa en un mismo centro. Como se profundizará a continuación, en el caso de Canarias, la única información censal disponible de las evaluaciones de diagnóstico a nivel de sección censal proviene de 3º y 6º de Primaria, lo que podría llevar a infraestimar la contribución real de algunos centros, al haber progreso que se ha conseguido entre 1º y 3º de Primaria. En Catalunya, en cambio, se dispone de datos de 6º de Primaria y 4º de ESO, pero no de los cursos iniciales de la Secundaria, lo que impide una valoración completa del progreso a lo largo de todo el ciclo. Estas restricciones temporales y de cobertura en la evaluación subrayan la necesidad de interpretar los resultados con cautela y de impulsar mejoras continuas en la recolección de datos educativos.

El último reto metodológico considerado se relaciona con la estabilidad temporal del valor añadido. La literatura muestra que los efectos de centro calculados a partir de una sola cohorte pueden variar cuando se consideran distintas generaciones o varios puntos en el tiempo (Leckie & Goldstein, 2019; Emslander et al., 2022). Esto no invalida los resultados, pero sí subraya la conveniencia de plantear, de cara al futuro, una estimación periódica y longitudinal de la resiliencia escolar. De este modo, las administraciones educativas podrían identificar con mayor solidez qué centros mantienen de manera consistente un efecto positivo sobre el aprendizaje y cuáles requieren apoyos adicionales, facilitando así un seguimiento más preciso de la eficacia escolar a lo largo del tiempo.

3. Escuelas Resilientes en Educación Primaria: un caso de estudio para las Islas Canarias

3.1. Datos, variables y complejidad de las escuelas en Canarias

Fuentes de datos y variables

El análisis del caso de estudio para Canarias se ha basado fundamentalmente en información de estudiantes y centros proporcionada por la Agencia Canaria de Calidad Universitaria y Evaluación Educativa (ACCUEE) y bases de datos abiertas del Instituto Nacional de Estadística (INE). La principal fuente de datos ha sido la ACCUEE, que, bajo estrictos acuerdos de confidencialidad, compartió con el equipo de investigación los microdatos a nivel de estudiante de las evaluaciones de diagnóstico que se llevan a cabo anualmente en Canarias en Primaria y Secundaria, así como los cuestionarios de contexto a estudiantes, familias, equipos directivos y tutores que las acompañan.

Aunque la ACCUEE dispone de pruebas de diagnóstico desde 2015/2016 hasta 2022/2023, no todos los años son comparables: durante la pandemia (2019/2020 y 2020/2021) no se aplicaron, y en muchos de los cursos las pruebas son muestrales. Así, solo en 2015/2016 (3º de EP) y 2018/2019 (6º de EP) las pruebas censales permiten seguimiento en Educación Primaria.³

En este sentido, un aspecto importante a considerar es que el indicador de valor añadido (VAM) que se utiliza a lo largo del informe parte de la calificación en 3º de EP en lugar de 1º de EP. Esto implica que el punto de partida en 3º de EP puede ya estar incorporando parte de la mejora que el centro haya propiciado entre 1º y 3º, lo que llevaría a infravalorar el efecto real del centro de 3º a 6º para aquellos centros especialmente efectivos en los primeros cursos. En otras palabras, si un centro ejerce un impacto positivo considerable entre 1º y 3º, esa ganancia quedará “oculta” cuando usemos 3º de EP como la línea de base, reduciendo la magnitud estimada del VA de 3º a 6º.

Aun cuando no se espera que un centro pase de tener VA positivo a negativo en tan corto plazo—o a la inversa—, esta limitación de los datos nos recuerda que el VA estimado puede no captar el efecto completo que un centro ejerce a lo largo de toda la etapa de Primaria. En un escenario ideal, se contaría con datos de rendimiento desde 1º de EP para capturar de forma más fiel la contribución acumulada de cada centro. Por tanto, a la hora de interpretar nuestros resultados, debe estar presente que el efecto real de los centros, en especial de aquellos con mejoras tempranas, podría ser mayor de lo que muestran los modelos basados únicamente en la calificación de 3º de EP.

A continuación, se describen las variables utilizadas a lo largo del informe.

³ Debido a una gran pérdida de observaciones entre 6º de EP y 4º de ESO y a que el trabajo de campo se ha enfocado en centros educativos en Primaria, no se ha contemplado por el momento llevar a cabo estimaciones de eficacia escolar para Secundaria en las Islas Canarias.

Información a nivel de estudiante

⇒ **Información de carácter administrativo: fecha de nacimiento, género, país de nacimiento y edad del alumnado.**

En 6º de EP se cuenta con un total de 18.986 estudiantes, conservando prácticamente la totalidad de las observaciones en todas las variables de carácter administrativo excepto el país de nacimiento del estudiante, con un 5,2% de valores ausentes.

⇒ **Resultados de aprendizaje censales del alumnado en las evaluaciones diagnósticas en 3º y 6º de Primaria.**

Los resultados de 3º de EP actuarán como punto de partida y variable explicativa en los modelos, mientras que los resultados de 6º de EP serán la variable dependiente. Por lo tanto, aunque en 6º de Primaria se disponga de información de competencia matemática, comunicación lingüística en lengua castellana (lengua, a partir de ahora), ciencias e inglés, solamente se incluirán **matemáticas y lengua**, al estar disponibles en ambos cursos.

En las evaluaciones de diagnóstico de Canarias de 6º de EP, se calculan los resultados de aprendizaje en las evaluaciones mediante el modelo de respuesta al ítem (IRT, por sus siglas en inglés), una metodología estadística que permite estimar el nivel de competencia de los estudiantes a partir de sus respuestas a un conjunto de preguntas. Es una metodología habitual en educación y es la utilizada también en pruebas internacionales como PISA, TIMSS o PIRLS. En lugar de asignar una puntuación basada únicamente en el número de aciertos, Canarias lleva a cabo un modelo IRT de dos parámetros que tienen en cuenta la dificultad de cada ítem y su capacidad de discriminar las respuestas de alumnos de distintos niveles. Esto permite estimar la competencia de cada estudiante de una manera muy precisa.

Los resultados se presentan en una escala con media 500 para el primer año en que se realizan las pruebas y desviación estándar 100, lo que facilita la comparación de los niveles de desempeño entre estudiantes. Sin embargo, en aras de facilitar la interpretación de los resultados y permitir una comparación más clara y tradicional de los efectos estimados en el modelo, se ha optado por estandarizar los puntajes resultantes de la metodología IRT, de manera que tengan una media de 0 y una desviación estándar de 1. Esto permite expresar los efectos en términos de desviaciones estándar, lo cual suele ser más intuitivo para realizar comparaciones. Además, este procedimiento no altera las relaciones estadísticas entre las variables ni introduce sesgos en la estimación.

Respecto a la población, en 3º de EP, se dispone de información de 18.825 estudiantes, ascendiendo el número a 18.986 en 6º de EP. Sin embargo, 2.621 estudiantes realizaron las pruebas en 3º de EP, pero no en 6º de EP (“salientes”) por distintas razones: o bien son alumnos que han abandonado el sistema educativo canario o bien han repetido curso. A su vez, 2.635 estudiantes realizaron las pruebas en 6º pero no en 3º de EP (“entrantes”, potencialmente por repetición o llegada al sistema educativo canario). Por lo tanto, la población de 6º de EP consta de 18.500 estudiantes y 504 centros, mientras que la población objeto de estudio en los modelos de regresión constará de **15.813 estudiantes**

que llevan a cabo las pruebas de competencia matemática tanto en 3º de EP como en 6º (15.836 en el caso de Lengua) en 503 centros.

⇒ **Cuestionarios de contexto realizados al alumnado y a las familias.**

Las evaluaciones de diagnóstico en el caso de las Islas Canarias van acompañadas de cuestionarios de contexto que se administran a los estudiantes y sus familias con el fin de recabar información sobre su entorno, sus características personales, familiares y escolares. El 89% de estudiantes de 6º de EP matriculados en el curso 2018/2019 contestaron a los cuestionarios donde se recoge información variada sobre la experiencia y el contexto educativo, con cerca de 90 variables: esto incluye aspectos de su trayectoria académica (repetición, absentismo), sus hábitos de estudio (frecuencia y dedicación a los deberes), su percepción de la enseñanza en clase (metodologías utilizadas, uso de recursos tecnológicos), las relaciones interpersonales y el clima escolar (ambiente de clase, participación, motivación, autopercepción en diferentes asignaturas), así como variables de autoconcepto y disfrute.

Por su parte, los cuestionarios administrados a las familias del alumnado de 6º de EP en ese mismo año (2018/2019) permiten describir el contexto sociofamiliar del alumnado y su relación con la experiencia educativa mediante alrededor de 70 variables. Las tasas de respuesta son sustancialmente menores con respecto al alumnado, con solo el 57,5% de familias que responden, siendo el porcentaje de 51,4% cuando se trata de la variable de nivel de ingresos. Las variables de estos cuestionarios incluyen información diversa sobre la composición y características del hogar, el nivel de estudios y la situación laboral de los progenitores, así como sus fuentes de ingresos y nivel de ingresos. Además, incluyen datos sobre la implicación familiar en la educación y la satisfacción con diferentes aspectos del centro educativo.

Información a nivel de centro educativo

⇒ **Cuestionarios de contexto dirigidos al equipo directivo.**

Además de los estudiantes y las familias, los equipos directivos también responden a un cuestionario asociado a las evaluaciones de diagnóstico con alrededor de 100 preguntas en el que se recoge información sobre la trayectoria y características del equipo directivo, la organización y recursos del centro, así como su percepción de la gestión, el clima escolar y los problemas más comunes. Asimismo, incluyen ítems relacionados con la satisfacción de la dirección respecto a la marcha del centro, la relación con el profesorado y el alumnado, y las medidas adoptadas para impulsar mejoras educativas. De los 504 centros educativos disponibles, 467 responden a estos cuestionarios.

⇒ **Cuestionarios de contexto dirigidos a los tutores de 6º de Primaria.**

El último de los cuestionarios de contexto se dirige a los tutores de 6º de EP, donde las y los tutores de 471 de los 504 centros —más de uno por centro— contestan a cerca de

130 preguntas.⁴ Estas variables hacen referencia a la labor y percepciones del profesorado tutor: se indaga en la actitud y participación del alumnado, las metodologías de enseñanza y evaluación que emplean, así como su percepción sobre recursos disponibles, problemas de infraestructura o materiales, y la forma de abordar incidencias disciplinarias. También se recogen datos sobre la relación con las familias, el equipo directivo y entre compañeros. Finalmente, se pregunta por la satisfacción general del tutor respecto al rendimiento del alumnado y el clima de trabajo en el centro.

Nivel socioeconómico de los centros educativos

Persistencia de valores ausentes en la información socioeconómica

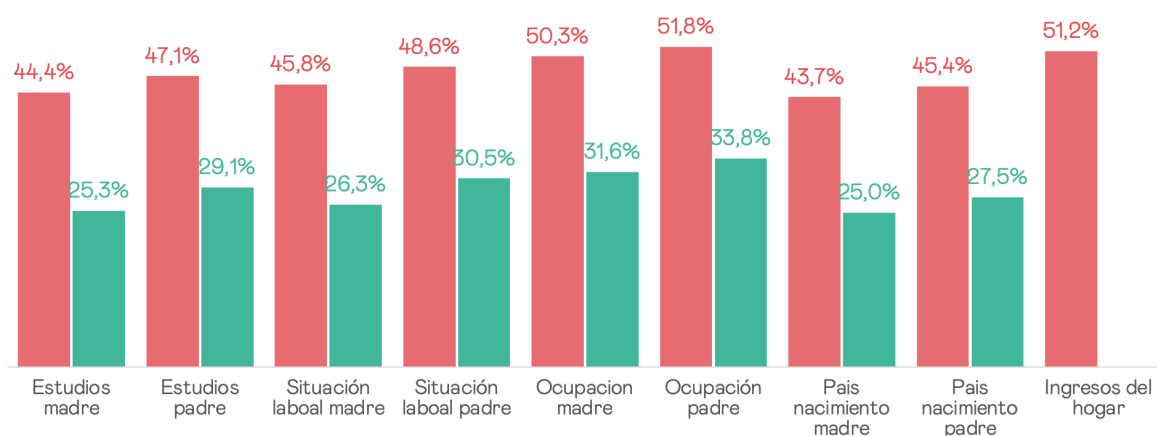
En aras de seleccionar aquellos centros que son eficaces en situaciones de complejidad, resulta necesario calcular el nivel socioeconómico de los centros educativos. Para ello, se crea un indicador socioeconómico de las escuelas utilizando información reportada por las familias en los cuestionarios de contexto. En este proceso, surge una dificultad: las variables relacionadas con el nivel socioeconómico de los hogares muestran un elevado porcentaje de valores ausentes, lo que puede dificultar la detección de la complejidad. Para tratar de reducir esta problemática, se imputan los valores ausentes utilizando la información de los cuestionarios de contexto para la misma cohorte tres años antes (en 2015/2016), con la excepción de la variable de ingresos, por su carácter cambiante y su escasa disponibilidad en la ola previa. En el Gráfico 1 se observa que este proceso consigue disminuir sustancialmente el número de valores ausentes para todas las variables consideradas.

Aun así, continúa existiendo una elevada concentración de los valores ausentes de estas variables en determinados centros que, además, no es aleatoria. En el Gráfico 2 se representa la correlación intraclase (ICC – Intraclass Correlation), que indica que entre el 14% y el 16,6% de la variación total de no reportar información se debe a diferencias entre centros. También influye la titularidad: la probabilidad de que las familias no reporten información socioeconómica es mayor en centros concertados y privados que en los públicos. Asimismo, los estudiantes sin información reportada presentan, en promedio, un rendimiento inferior. De hecho, al estimar un modelo con efectos fijos de centro, se observa que incluso dentro de un mismo centro estos estudiantes obtienen peores resultados —por ejemplo, en lengua 0,26 SD menos en promedio—.

⁴ Cabe destacar que no se dispone del fichero de registro para el cuestionario de tutores del curso 2018/2019, sino solo del cuestionario de 2022/2023. Por lo tanto, solo se utilizan aquellas variables en las que se puede determinar con certeza la definición de la variable y el significado de los valores.

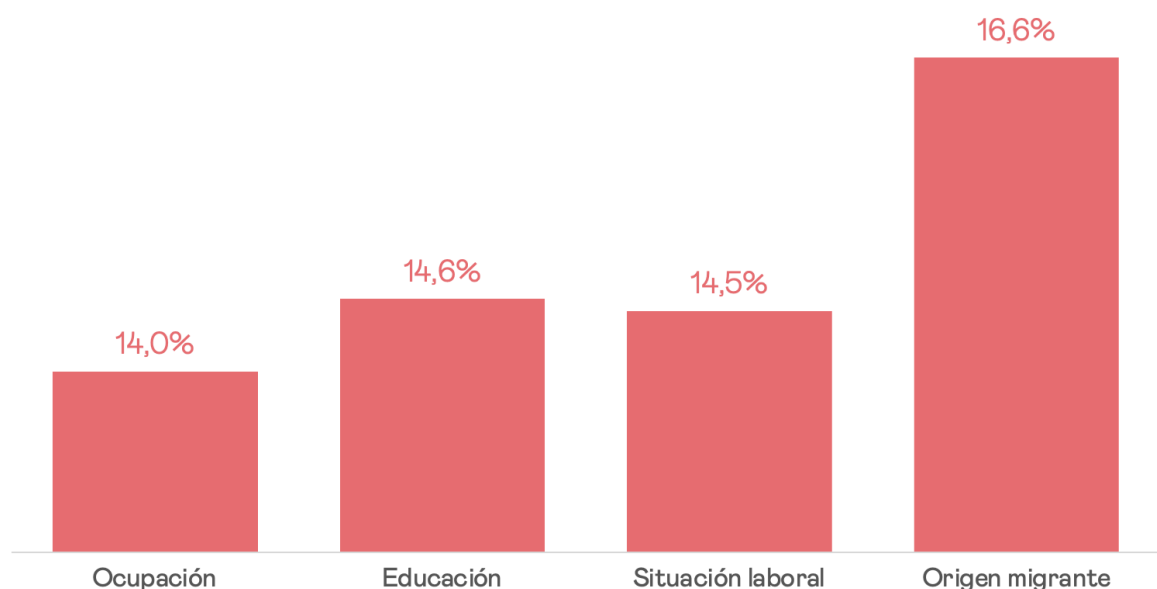
Gráfico 1. Porcentaje de valores ausentes de variables socioeconómicas antes y después del proceso de imputación

Original e imputada



Fuente: elaboración propia con datos de la ACCUEE | EsadeEcPol

Gráfico 2. Correlación intraclase de los valores ausentes



Fuente: elaboración propia con datos de la ACCUEE | EsadeEcPol

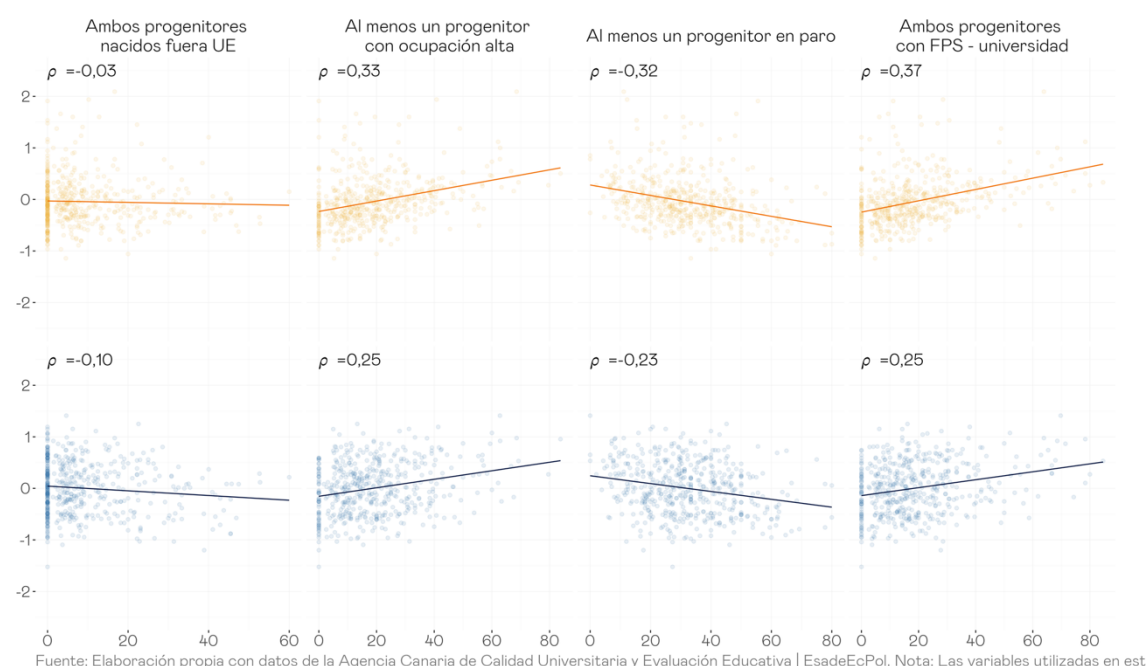
Cálculo del índice socioeconómico de los centros

La elección de las variables que se incluirán en el índice socioeconómico del centro se fundamenta en su relación con los resultados en competencia matemática y lingüística, la persistencia de valores ausentes y su capacidad para explicar la mayor proporción de la **varianza**. Como se mencionaba con anterioridad, dado que no se dispone de información de carácter administrativo del nivel socioeconómico del alumnado y sus progenitores, se utilizan las variables de educación, ocupación, situación laboral y origen de los progenitores recogidas en los cuestionarios de contexto asociados a las evaluaciones en 2018/2019 (6º de EP), con información imputada de 2015/2016 (3º de EP). De este modo, una limitación evidente más allá de los valores ausentes es que las

variables incluidas reportan información de la complejidad del alumnado que acude a 6º de EP, no de todo el alumnado matriculado en el centro.

Tanto la información de situación laboral y el nivel de ocupación de los progenitores, como su educación y origen, muestran una relación estadísticamente significativa con los resultados de las evaluaciones a nivel de centro. Concretamente, si se utilizan modelos lineales para estimar la relación entre los resultados y el porcentaje de estudiantes con un (dos) progenitor(es) en paro, un (dos) progenitor(es) con ocupación alta, un (dos) progenitor(es) con formación universitaria, y un (dos) progenitor(es) nacido fuera de la UE, se obtiene que, por pares, las variables que explican la mayor proporción de la varianza de resultados entre centros y tienen un coeficiente más elevado —y significativo— son el porcentaje de estudiantes con un progenitor en paro, el porcentaje de estudiantes con un progenitor con ocupación alta, con dos progenitores con formación universitaria y con dos progenitores nacidos fuera de la UE. Esta última variable es la que, como se observa en el Gráfico 3, guarda una menor relación con los resultados, siendo excluida cuando se utiliza el método LASSO⁵ para identificar las variables, así como todas las variables de origen migrante de los progenitores.⁶

Gráfico 3. Correlación entre los resultados en **matemáticas y lengua** en 6º de EP y las variables clave para la clasificación del nivel socioeconómico de los centros



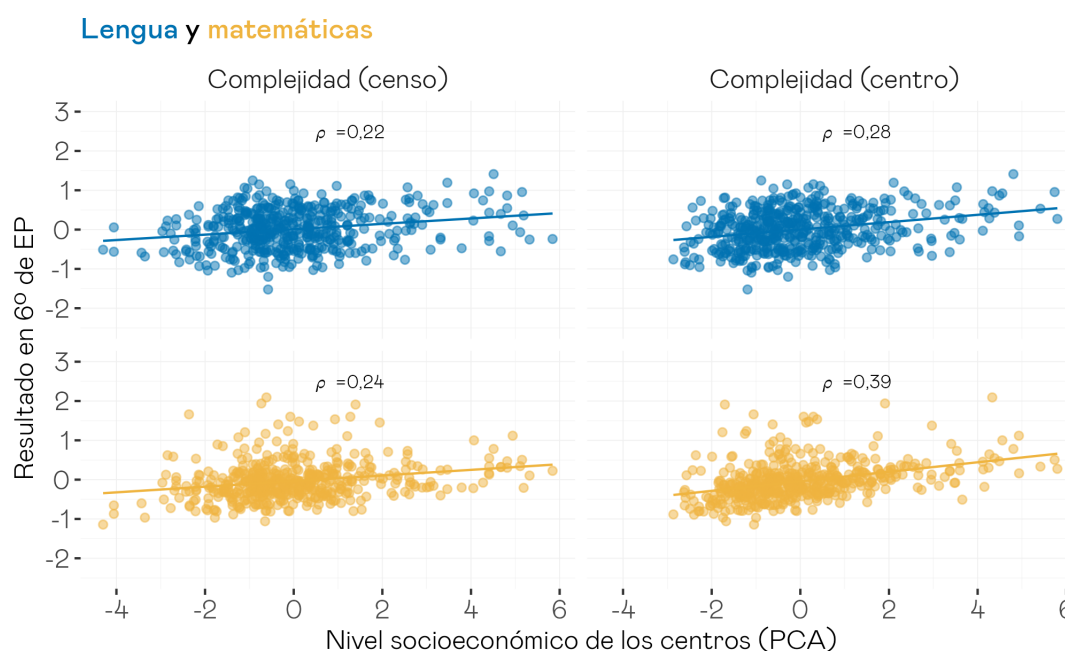
⁵ El *Least Absolute Shrinkage and Selection Operator* es un método de regresión penalizada que, al añadir un término de penalización sobre los coeficientes, fuerza a que algunos se reduzcan hasta cero y selecciona así las variables más relevantes. Esto resulta especialmente valioso cuando se cuenta con numerosas variables y se busca un modelo parsimonioso y fácil de interpretar, como sucede al estudiar el efecto de factores socioeconómicos en los resultados de evaluaciones. Sin embargo, dado que la penalización introduce un sesgo en los coeficientes, se aplica una etapa posterior de estimación en la que se reestima el modelo sin penalización (conocida como post-selection), utilizando únicamente las variables seleccionadas, para obtener estimaciones más precisas y confiables.

⁶ Los resultados de este informe son robustos a la inclusión o exclusión de esta variable.

Por último, para tratar de mitigar el efecto de los valores ausentes en las variables de interés, se incluye información a nivel de sección censal del INE,⁷ concretamente de la sección censal en la que se encuentra el centro, como la tasa de paro, la renta neta media del hogar y el porcentaje de hogares con educación superior. Pese a que todas ellas guardan una relación estadísticamente positiva, como se observa en el Gráfico 4, cuando se construye el índice utilizando información censal, la correlación con los resultados en las evaluaciones es menor. Adicionalmente, cuando se combina información a nivel de sección censal y centro, las variables de sección no aportan información adicional a lo que ya aporta la variación entre centros e introduce una capa adicional de complejidad.

Por lo tanto, el índice socioeconómico elegido, incluye información del porcentaje de progenitores en paro, con ocupación alta y educación terciaria, así como el porcentaje de progenitores nacidos fuera de la UE. Adicionalmente, se comprueba la robustez de los resultados a la exclusión de la variable de origen.

Gráfico 4. Correlación entre el índice de nivel socioeconómico de los centros y los resultados en 6º de EP en **lengua** y **matemáticas**



Fuente: Elaboración propia con datos de la ACCUEE | EsadeEcPol.

3.2. ¿Cómo se estima la eficacia escolar en Primaria?

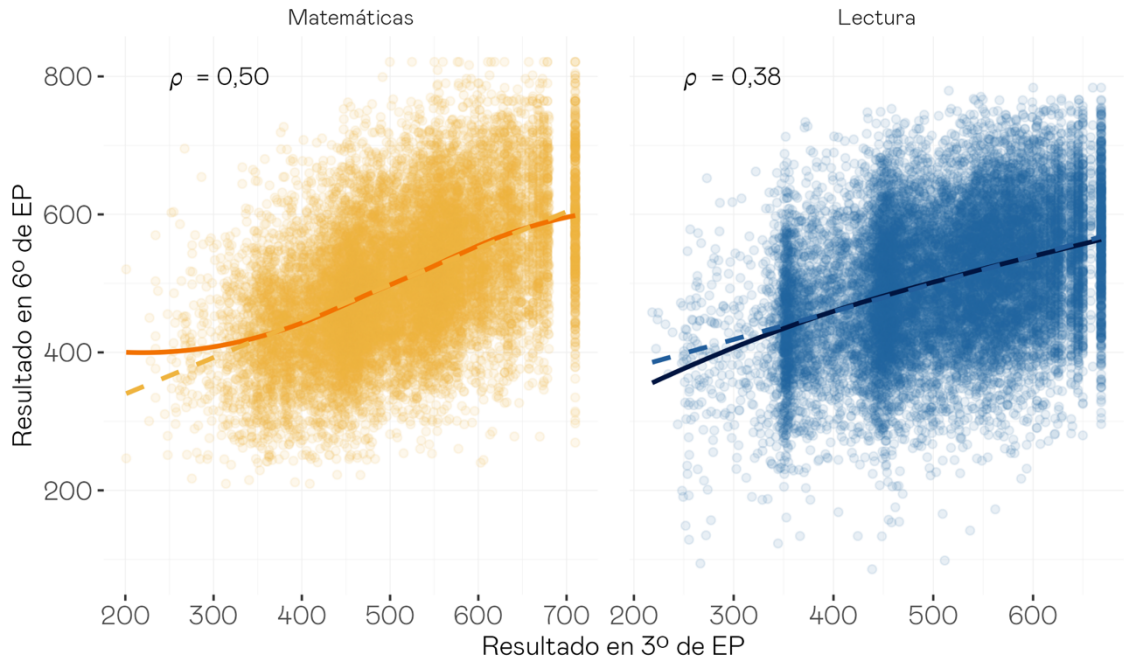
Comparación entre modelos de Valor Añadido en Canarias

En este apartado, se lleva a cabo una comparación entre modelos de eficacia escolar, siguiendo el análisis realizado por Leckie y Prior (2022).

⁷ Censo de Población y Viviendas (2021) y Atlas de Distribución de Renta de los Hogares (2021).

En primer lugar, el Gráfico 5 presenta la relación entre los resultados de aprendizaje de los estudiantes entre los cursos de 3º y 6º de EP. Como puede apreciarse, la correlación es alta, siendo más acusada para matemáticas (0,5) que para lengua (0,36), algo habitual en la literatura. Esta correlación alta ya apunta a la necesidad de considerar el uso de un modelo de valor añadido; es decir, a controlar por el aprendizaje previo en 3º.

Gráfico 5. Correlación entre los resultados de las competencias en 3º y 6º de EP en **matemáticas** y **lengua**



Fuente: Elaboración propia con datos de la ACCUEE | EsadeEcPol.

La Tabla 3 presenta los principales parámetros de los modelos de VA estimados mediante efectos mixtos con intercepto aleatorio: el coeficiente del rendimiento previo, la varianza del intercepto aleatorio (Var-u0, variación entre centros), la varianza de los residuos (Var-e), la correlación intraclase (ICC) y el porcentaje de centros con valor añadido positivo. Además, la tabla resume las variables incluidas en cada modelo.

Tabla 3. Resumen de los estadísticos de los principales modelos de valor añadido

	Matemáticas (6º de EP)					Lengua (6º de EP)				
	Raw	VA	CVA	CVA-B	CVA-X	Raw	VA	CVA	CVA-B	CVA-X
Resultados 3º de EP		X					X			
Género, año de nacimiento, país de nacimiento		X	X				X	X		
Resultados centro 3º de EP (efecto de pares)		X	X	X			X	X	X	

Info socioeconómica de centro (educación y ocupación)	X	X	X	X	X	X	X	X	X
Coef. 3º de EP	0,524	0,512	0,516	0,516	0,386	0,357	0,358	0,358	
R ² Aj.	0	0,254	0,260	0,261	0,268	0	0,144	0,163	0,167
Var(u0)	0,194	0,153	0,151	0,149	0,142	0,220	0,205	0,202	0,200
Var(e)	0,803	0,597	0,591	0,591	0,591	0,743	0,616	0,600	0,600
ICC	0,195	0,204	0,204	0,201	0,194	0,228	0,250	0,252	0,250
% centros VA>0	33,99	40,76	38,97	37,77	36,78	43,54	45,73	44,73	44,93
N centros	503	503	503	503	503	503	503	503	503
N estudiantes	15813	15813	14971	14971	14971	15836	15836	14989	14989

En primer lugar, se observa que los resultados previos en las evaluaciones explican la mayor parte de la variabilidad en los resultados de 6º de EP, como muestra el R² ajustado en ambas competencias, especialmente en competencia matemática —en línea con la evidencia existente (Leckie & Prior, 2022)—. El rendimiento previo explica un 25% de la varianza en matemáticas y un 14% en lengua. Cabe mencionar que, además, estos resultados son robustos a la inclusión de no linealidades en el rendimiento previo. Al añadir variables de contexto del estudiante, aunque su relación es estadísticamente significativa, la varianza de los resultados explicada por el modelo solo aumenta ligeramente;⁸ de hecho, la diferencia es sustancialmente menor que la existente en la literatura para otros contextos (Leckie & Prior, 2022). Sin embargo, los datos son consistentes con los datos conocidos para el contexto de Canarias para la Educación Secundaria Obligatoria, donde se observa una menor variación de los resultados explicada por el origen migrante, el género y el nivel socioeconómico que en otras comunidades autónomas (Cobrerros & Gortazar, 2023) y que, acorde con la literatura, tendría lógica pensar que fuera todavía menor en Primaria. En el modelo CVA - B, en el que se trata de capturar el efecto de los pares mediante la inclusión de la media de los resultados previos a nivel de centro, se observa que la varianza se mantiene constante, en línea con los resultados obtenidos por Thomas y Mortimore (1996); de hecho, si bien esta variable es significativa cuando se incluye en los modelos que ignoran el contexto de los estudiantes, pasa a ser no estadísticamente significativa cuando se incluyen variables de contexto del estudiante. Lo mismo ocurre en el modelo CVA - X: si bien las variables del contexto socioeconómico promediadas a nivel del centro son relevantes para explicar

⁸ Hay que tener en cuenta que, al añadir variables de contexto, disminuye la muestra en casi 1.000 estudiantes, lo que también puede tener un efecto en el R². Sin embargo, si estimamos el modelo VA con una muestra comparable, el R² disminuye igualmente en la misma medida, por lo que no se trata de un efecto derivado de la variación muestral.

la variación en los resultados cuando se incluyen en solitario, cuando se incluyen en una regresión que ya controla por el nivel socioeconómico de los estudiantes, no son estadísticamente significativas y, en consonancia, tampoco tienen un impacto en el R^2 ajustado, en línea con los resultados obtenidos por Timmermans et al. (2011).

En segundo lugar, la correlación intraclase (ICC) indica que prácticamente un quinto de la variación en las notas de 6º de EP se debe a diferencias de resultados entre centros, y no dentro de un mismo centro. La ICC indica qué proporción de la variabilidad total en las notas de 6º depende de las diferencias de resultados entre centros (Var-u0) y no de las diferencias entre estudiantes dentro de cada centro (Var-e). Cuando no se incluye ningún control, la ICC en matemáticas es del 19,5% y en lengua del 22,8%, indicando que dichos porcentajes de la variación en los resultados se deben a diferencias entre centros. Al introducir los resultados de 3º de EP, la ICC aumenta ligeramente: esto se debe a que la varianza de resultados dentro de los centros desciende y lo hace más que la varianza entre centros. Así, el rendimiento previo parece estar capturando una parte sustancial de la heterogeneidad entre estudiantes de un mismo centro, pero no las diferencias estructurales entre centros. Dicho de otro modo, el rendimiento previo explica mejor por qué los estudiantes del mismo centro difieren entre sí, pero no por qué difieren entre centros. Adicionalmente, introducir variables adicionales de contexto no varía prácticamente ni la varianza entre centros ni la varianza dentro de los centros. Por otro lado, el hecho de que la varianza entre centros sea mayor en lengua que en matemáticas mientras que la varianza dentro de los centros es similar, parece indicar que las condiciones estructurales de los centros tienen mayor impacto en lengua.

Por último, como resulta lógico, al controlar las diferencias de partida del alumnado, cambia el porcentaje de centros con VA positivo y estadísticamente significativo. En matemáticas, cuando no se añade ningún control, solo el 33,99% de los centros tiene un efecto positivo y significativo; sin embargo, en el modelo VA (controlando por resultados previos) sube al 40,76%. Para lengua, el salto es menor, (2 p.p).

Como análisis de robustez, aunque por simplicidad no se añade en la tabla, se ha estimado el modelo CVA incluyendo variables socioeconómicas de los estudiantes a pesar de que presentan un porcentaje de valores ausentes no aleatorios cercano al 25%. Dado que el nivel de educación y ocupación de los progenitores presenta una correlación estadísticamente significativa con los resultados, al omitirlas también podría incurrirse en un sesgo por omisión de variables relevantes, de forma que se estima el modelo en aras de comprobar que los estadísticos principales no varían sustancialmente. Adicionalmente, puesto que los valores ausentes están correlacionados con los resultados, se añaden variables dicotómicas para controlar su inclusión. Se observa que, aunque los coeficientes son estadísticamente significativos, el coeficiente asociado al rendimiento previo se mantiene, el R^2 ajustado también se mantiene estable,⁹ así como la correlación

⁹ Dado que el número de observaciones se reduce a 11.840 sus magnitudes no son estrictamente comparables; sin embargo, incluso cuando se comparan regresiones con muestras comparables, el R^2 no aumenta al incluir información de educación y ocupación de los progenitores.

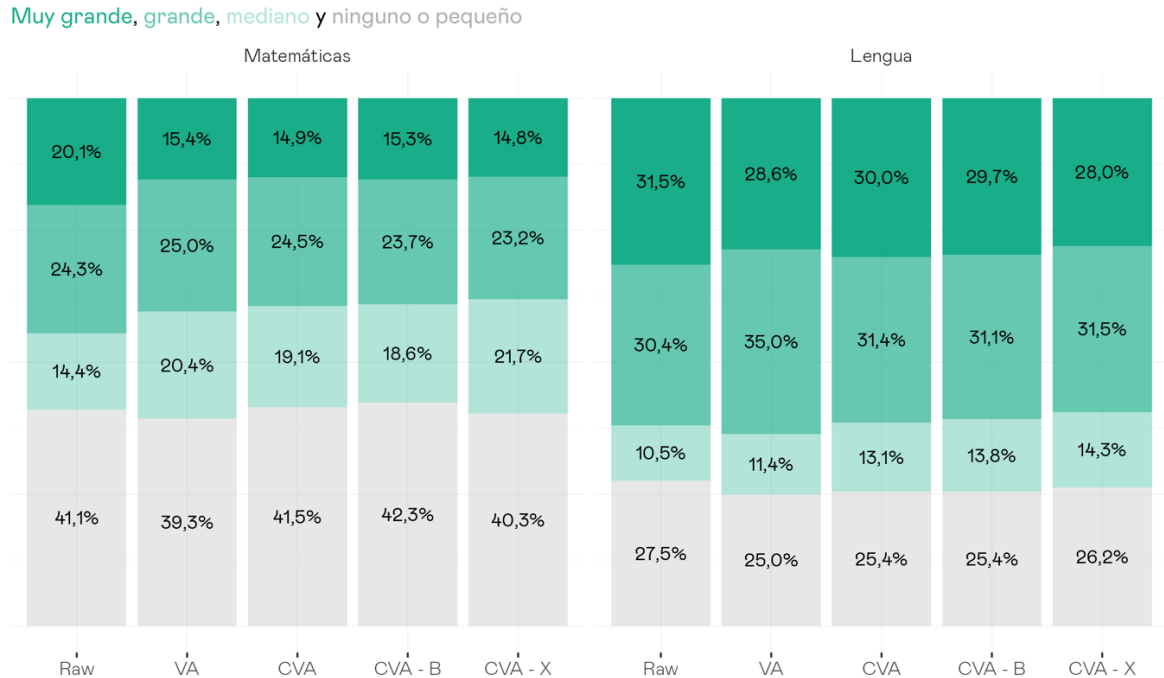
intraclase. Respecto al porcentaje de centros con valor añadido positivo, como cabe esperar, cae ligeramente (alrededor de 2 p.p).

Tamaño del efecto de los centros

Para comprender mejor el impacto que la elección que el modelo VAM tiene en la magnitud de los efectos individuales de las escuelas, se clasifica las escuelas con un valor añadido mayor o igual a cero en cuatro categorías según el valor absoluto de sus efectos estimados para las dos competencias. Siguiendo las definiciones utilizadas por la Education Endowment Foundation¹⁰ para evaluar intervenciones educativas (utilizadas por Leckie y Prior, 2022), los cuatro grupos se clasifican en función de la magnitud del VA expresado en unidades de desviación estándar del rendimiento académico actual de los estudiantes.

El primer hallazgo es que los centros presentan un VA más uniforme en matemáticas, siendo la proporción de centros con un valor añadido pequeño o no significativo sustancialmente mayor que en lengua (Gráfico 6); en cambio, la proporción de centros con valor añadido grande o muy grande es sustancialmente menor en matemáticas. Esto puede reflejar que la variabilidad en la enseñanza o el dominio del alumnado en competencia matemática sea menor o, al contrario, que las destrezas de lengua estén más ligadas a factores como el bagaje lector o el apoyo familiar, haciendo que ciertos centros destaquen más cuando el profesorado logra compensar esas diferencias iniciales.

Gráfico 6. Porcentaje de centros en función del tamaño de su valor añadido



Fuente: Elaboración propia con datos de la ACCUEE | EsadeEcPol. Los centros se clasifican con "ningún o pequeño", "mediano", "grande" o "muy grande" valor añadido cuando tienen valores absolutos en los rangos de 0 a 0,1, 0,1 a 0,2, 0,2 a 0,45 o mayores a 0,45; estos valores se miden en unidades de desviación estándar del rendimiento actual de los estudiantes.

¹⁰ La Education Endowment Foundation clasifica los efectos como: nulo o pequeño (de 0 a 0,1 SD), mediano (0,1 a 0,2SD), grande (0,2 a 0,45SD) y muy grande (>0,45SD).

En segundo lugar, como cabe esperar, la mayor corrección se produce al pasar del modelo “Raw” al modelo “VA”, especialmente en competencia matemática, donde la proporción de centros con VA muy alto o muy bajo se reduce notablemente al controlar las diferencias iniciales. Adicionalmente, los siguientes modelos apenas introducen cambios en el tamaño del VA. Esto contrasta con los resultados obtenidos por Leckie y Prior (2022), dado que, en el caso de Reino Unido (y en Secundaria) la proporción de centros con VA grande sufre una mayor reducción al controlar tanto por el rendimiento previo como por el resto de las variables.

Correlación entre modelos

Finalmente, se presenta información adicional de las correlaciones entre los cinco modelos de efectos escolares para aportar evidencia adicional sobre cómo los ajustes de variables influyen en el rendimiento relativo de los centros. Si bien es cierto que el porcentaje de centros con VA no cambia sustancialmente entre modelos, esto puede esconder un fuerte efecto composición: qué centros son clasificados. Por simplicidad, se comparan modelos consecutivos, es decir, que comparten un solo cambio, siendo estos donde se encuentran las correlaciones entre centros más elevadas.

Los resultados muestran que, aunque los efectos estimados de valor añadido entre modelos guardan correlaciones altas, los rankings de centros eficaces cambian sustancialmente, lo que revela que la clasificación de escuelas depende en gran medida de los ajustes incluidos. El Gráfico 7 a continuación muestra la correlación de los efectos de eficacia escolar entre los cinco modelos para ambas competencias. El Gráfico 9 muestra la correlación entre el ranking del VA de los centros, solo cuando el valor añadido es positivo y estadísticamente significativo. La Tabla 4, por su parte, compara los 100 primeros centros de ese ranking para cada uno de los modelos de VA.

La primera comparación corresponde a los modelos Raw y VA, siendo esta de 0,82 en el caso de matemáticas y 0,95 en el caso de lengua, sugiriendo que las escuelas con mejor rendimiento promedio siguen figurando como las más eficaces una vez ajustado el rendimiento previo de los estudiantes, aunque esta relación no es invariable. Adicionalmente, los resultados aportan aún más evidencia de la mayor relevancia del rendimiento previo en competencia matemática. Sin embargo, la correlación entre los rankings del Gráfico 8 es sustancialmente menor, especialmente en el caso de matemáticas (0,68): de los 287 centros eficaces (con $VA > 0$) en el modelo Raw, 41 dejan de serlo cuando se controla por el rendimiento previo y 75 nuevos centros (que no lo eran) pasan a ser eficaces —en la competencia de lengua, estas cifras descienden, siendo 213 a 13 y 31 nuevas, respectivamente—. Concretamente, y retomando la Tabla 4, de los 100 con mayor VA en Raw, solo 71 continúan siéndolo en VA en el caso de matemáticas y 78 en el caso de lengua.

Cabe destacar que, en el caso del modelo sin ajustes, el 69% de los centros eficaces en matemáticas eran de titularidad pública, mientras que, cuando se ajusta por el rendimiento previo, el porcentaje asciende al 78,5%; en línea con este resultado, los centros eficaces pasan a tener más alumnado con progenitores con estudios bajos y

ligeramente más alumnado extranjero; dicho de otro modo, mientras que, sin controlar por el rendimiento previo, el 20,5% de los centros eficaces en matemáticas tienen nivel socioeconómico bajo,¹¹ en el modelo VA el porcentaje asciende al 32,7% (en lengua estos porcentajes cambian al 30% y 35,9%). Así, centros que en el modelo Raw aparecían con peores resultados porque sus estudiantes partían de un punto más bajo, pueden revelarse como muy eficaces en términos de progreso real, lo que lleva a que más centros de nivel socioeconómico bajo pasen a estar clasificados como eficaces; es decir, una vez se elimina la desventaja inicial, hay un número significativo de centros con alumnado más vulnerable que logran avances notables, especialmente en matemáticas. Las correlaciones del valor añadido tienen magnitudes muy similares a las encontradas por Leckie y Prior (2022) y son ligeramente inferiores a Timmermans et al. (2011) o Marks (2021); sin embargo, las correlaciones en los rankings son sustancialmente más bajas, especialmente en el caso de matemáticas: el porcentaje de centros eficaces no cambia demasiado, pero sí que lo hace quiénes son los centros clasificados. De este modo, con el modelo sin ajustes, estaríamos infraestimando, de forma injusta, la eficacia para aquellos centros más complejos.

Gráfico 7. Correlaciones entre el valor añadido de los centros en matemáticas y lengua los cinco modelos principales, por pares

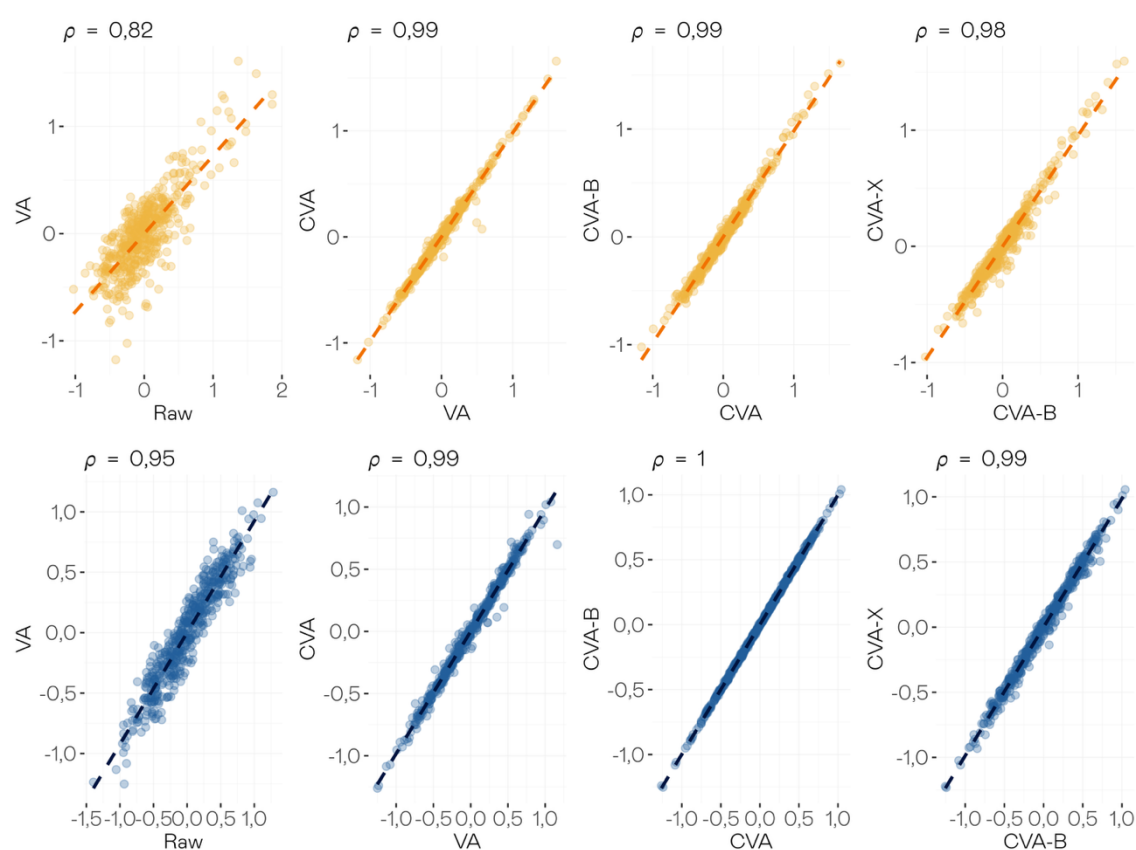


Gráfico 8. Correlaciones entre el ranking de los centros eficaces en matemáticas y lengua en función de su valor añadido en los cinco modelos principales

¹¹ Consideraremos centros con nivel socioeconómico bajo (o complejos) a los centros que se encuentren entre el 40% con menor nivel socioeconómico.

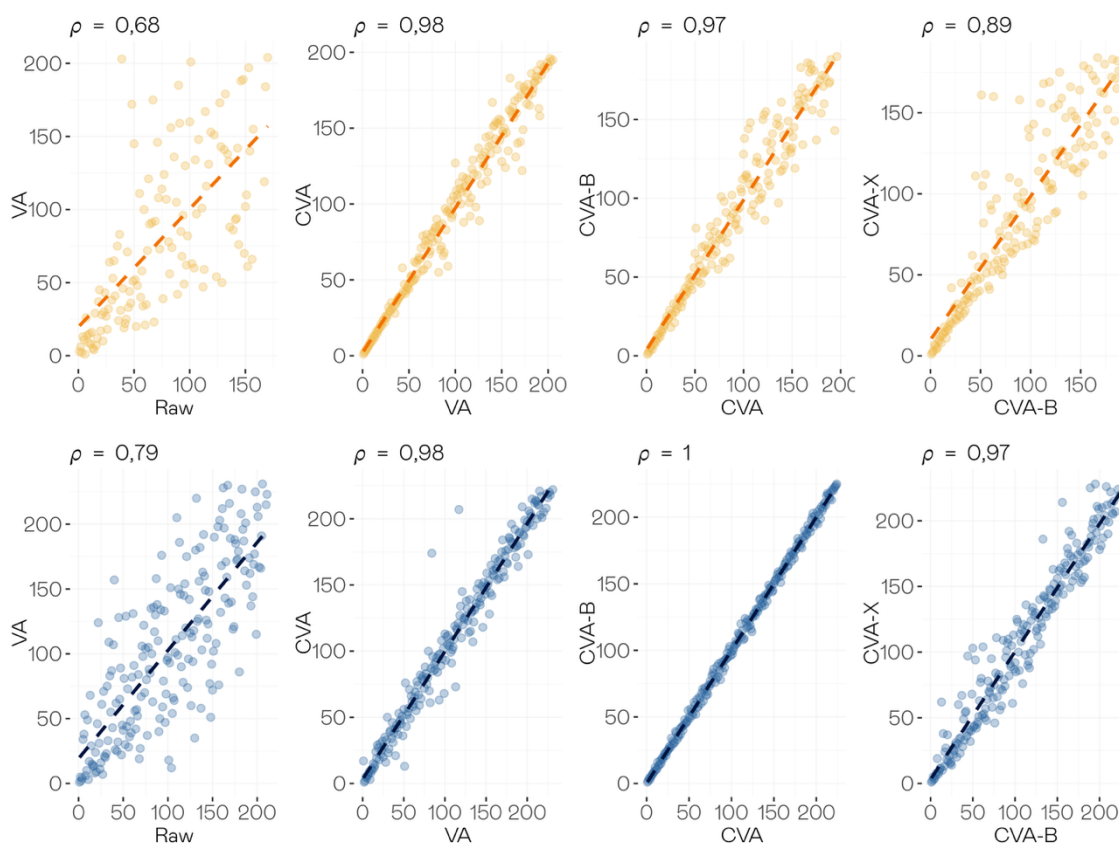


Tabla 4. Correlación entre modelos: 100 centros con mayor valor añadido

	Matemáticas					Lengua				
	RAW	VA	CVA	CVAB	CVAX	RAW	VA	CVA	CVAB	CVAX
RAW	100					100				
VA	71	100				78	100			
CVA	68	95	100			76	96	100		
CVAB	74	92	94	100		78	95	98	100	
CVAX	70	85	89	89	100	80	90	90	91	100

La segunda columna compara el modelo VA con CVA. Como cabe esperar dado el bajo impacto de las variables de contexto del estudiante en este análisis, la correlación es cercana a 1, tanto en el valor añadido como en el ranking de centros. De nuevo, en matemáticas la correlación del ranking es ligeramente menor, ya que de los 205 centros con VA positivo en el modelo VA, solo 11 pasan a tener un efecto no significativo al controlar por variables socioeconómicas y 2 nuevos adquieren significancia estadística; de hecho, entre los 100 más eficaces en matemáticas con VA, 95 continúan siéndolo al controlar por estas variables. Por otro lado, no existen diferencias relevantes entre los centros que dejan de ser eficaces al introducir variables de contexto. Las correlaciones

son sustancialmente mayores a las encontradas por otros autores para el contexto de Reino Unido, Países Bajos y Australia.¹²

Por último, la correlación entre modelos al incorporar las variables de contexto de los centros es elevada, especialmente con la inclusión del rendimiento previo a nivel de centro, que prácticamente no altera la clasificación relativa de las escuelas. La inclusión de variables socioeconómicas de centro sí altera ligeramente el ranking, especialmente en matemáticas, donde 19 centros dejan de ser eficaces y 14 centros pasan a serlo; de los 100 más eficaces, 89 se mantienen en el ranking. Además, como ocurría al pasar del modelo sin ajustes al VA, los centros con valor añadido positivo que pasan a tener un valor añadido nulo o negativo son mayoritariamente de titularidad privada (63%) y, por ende, minoritariamente complejos o de nivel socioeconómico bajo (5,3%), mientras que los centros que pasan a ser eficaces son todos públicos y, de hecho, el 71,4% tienen un nivel socioeconómico bajo. Así, al incluir en el modelo información socioeconómica agregada que representa características estructurales de las escuelas, se descuenta la desventaja que proviene del entorno global del centro, de forma que algunos centros que parecían muy eficaces simplemente porque atendían a un alumnado con un contexto favorable pueden ver reducido su VA, mientras que otros con entornos más complejos pueden “ascender” en el ranking una vez que se ajusta ese factor externo.

Sin embargo, hay que tener en cuenta que este modelo (CVA - X) está redefiniendo los efectos escolares comparando cada centro únicamente con otros de su mismo tipo, eliminando así diferencias en la efectividad escolar independientemente de que estas sean relevantes para los sistemas de evaluación que buscan identificar centros con dificultades. Por lo tanto, la utilización de este modelo (o la restricción del modelo a centros comparables desde el punto de vista socioeconómico) debe utilizarse, sobre todo, cuando quieran compararse centros con características similares.

Por último, como se intuye del análisis anterior, la correlación del valor añadido entre materias es quizá menor de lo que cabría esperar, siendo sustancialmente mayor el porcentaje de centros que consiguen que su alumnado tenga un rendimiento relativo positivo (y grande) en competencia lingüística que matemática. Concretamente, de los 100 centros con mayor VA en matemáticas, solamente hay 44 en lengua —comparando el modelo Raw—; si se compara con los modelos con ajustes, el número baja a entre 36 y 38 centros. Esto puede suceder porque ambas competencias, aun compartiendo ciertas destrezas transversales, no se benefician de las mismas estrategias docentes ni del mismo bagaje previo del alumnado. Un centro que logra un gran avance en lengua podría no ofrecer la misma ventaja en matemáticas (y viceversa) debido a diferencias en la preparación de los docentes, características estructurales de los centros e incluso características del alumnado y de sus familias. Esto justifica la separación de ambas competencias a lo largo del análisis.

¹² (Leckie y Prior, 2022; Marks, 2021; Leckie y Goldstein, 2019 y Timmermans et al., 2011).

Selección del modelo CVA y justificación

El modelo Raw, que no incluye ningún ajuste, se descarta porque refleja únicamente las diferencias brutas en los resultados finales. Así, los centros que reciben alumnado con un nivel previo más alto aparecen sistemáticamente mejor situados, mientras que los que atienden a estudiantes con más dificultades quedan penalizados. En la literatura sobre eficacia escolar se ha señalado de forma reiterada que este enfoque confunde nivel con progreso, por lo que no resulta adecuado como medida de valor añadido (Thomas & Mortimore, 1996; Marks, 2021).

El modelo VA, que incorpora el rendimiento previo del alumnado, constituye un avance notable al permitir estimar cuánto mejora cada estudiante respecto a su propio punto de partida. Este ajuste es fundamental para aproximarse al efecto real de los centros, ya que neutraliza en gran medida la ventaja inicial que algunos colegios presentan por la composición de su alumnado.

Por su parte, el modelo CVA - A enriquece la estimación al añadir variables individuales como el género, el país y el año de nacimiento (y variables socioeconómicas). Aunque estas variables no alteran de forma drástica los resultados, su inclusión permite ajustar el rendimiento de cada estudiante en función de sus características personales. De este modo, si un centro escolar concentra más alumnado con un perfil determinado, estas diferencias no se confunden con su valor añadido. En otras palabras, el CVA contribuye a que las comparaciones entre centros se basen en condiciones de alumnado más equiparables (Leckie & Prior, 2022).

El modelo CVA - X, en cambio, plantea una cuestión más de fondo. Su propósito es descontar las diferencias estructurales entre centros al introducir indicadores socioeconómicos agregados. Sin embargo, en este informe el interés no está en comparar todos los centros entre sí, sino en identificar cuáles logran buenos resultados dentro del grupo de escuelas que ya trabajan en condiciones de mayor complejidad. Si además de restringir la muestra a esos centros se incluyeran las variables agregadas que definen esa complejidad, se estaría eliminando la variación que permite distinguir a los resilientes. Dicho de otro modo, el riesgo es “ajustar en exceso” y perder aquello que se quiere analizar: cómo algunos centros complejos logran progresos destacados del alumnado pese a su contexto más adverso (Thomas, 2001; Muñoz-Chereau & Thomas, 2016).¹³

Por estas razones, el modelo CVA - A ofrece un equilibrio razonable: ajusta por el rendimiento previo y las características individuales básicas del alumnado, evita los sesgos más evidentes de Raw y mantiene la comparabilidad entre centros de distinta complejidad. Esto permite centrar el análisis en lo esencial: comprender qué hacen los centros complejos que logran un valor añadido positivo y qué enseñanzas pueden extraerse para orientar políticas de apoyo y mejora.

¹³ Cabe mencionar que, cuando se estima el modelo de eficacia escolar solamente en los centros complejos, los centros con VA estimado positivo son prácticamente los mismos que cuando se incluye en el modelo el control del nivel socioeconómico a nivel de centro, reforzando la robustez de los resultados.

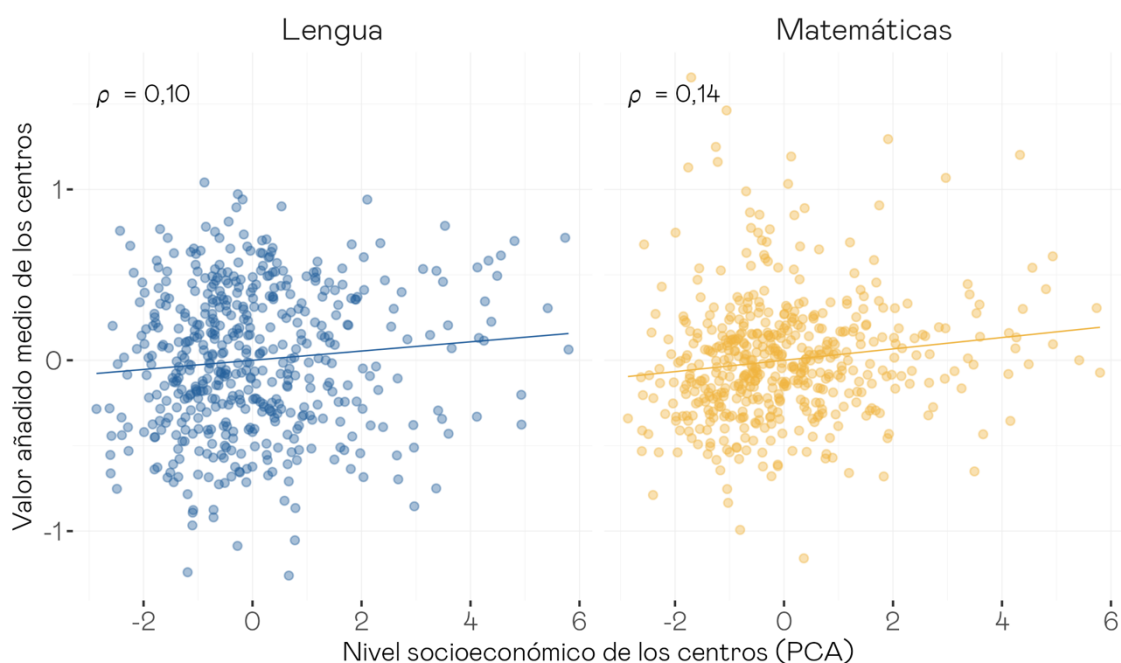
3.3. Relación entre eficacia y nivel socioeconómico de los centros de Primaria: centros *resilientes* en Canarias

Describiendo a los centros en función de la complejidad de su alumnado

El principal objetivo de este informe es identificar a los centros que, trabajando con alumnado con un nivel socioeconómico más bajo y, por ende, más vulnerable, obtienen una mejora en las competencias básicas evaluadas superior a lo que cabría esperar dada la vulnerabilidad del alumnado al que atienden.

La correlación entre el VAM de los centros y su nivel socioeconómico es baja, especialmente en la competencia lingüística, pudiendo esto tener implicaciones positivas en términos de equidad (Gráfico 9). Esto contrasta con la correlación observada entre el índice socioeconómico de los centros y los resultados académicos. Sin embargo, ambos resultados son compatibles. Un centro con alumnado de bajo perfil socioeconómico —que correlaciona con tener menores niveles en ambas competencias— pero que logra avances superiores a lo esperado dada su situación de partida, mostrará un valor añadido más alto. De hecho, que la relación entre complejidad y eficacia no sea elevada puede leerse como una señal positiva en términos de equidad: sugiere que el avance real que logran los estudiantes no está fuertemente determinado por los recursos o el contexto medio del centro. Es decir, incluso en entornos menos favorecidos, un volumen relevante de centros educativos consigue un incremento académico equiparable al de centros con un alumnado de mayor nivel socioeconómico. Así, cuando se estiman los modelos para diferentes grupos de complejidad, el porcentaje de centros que tienen un valor añadido positivo no cambia tanto como cabría esperar de antemano.

Gráfico 9. Correlación entre el nivel socioeconómico de los centros y su valor añadido



Fuente: Elaboración propia con datos de la Agencia Canaria de Calidad Universitaria y Evaluación Educativa |

Sin embargo, cabe pensar que existan diferencias sustanciales entre los centros con menor nivel socioeconómico y el resto que justifican estimar el VAM de forma diferenciada para identificar a las escuelas resilientes. En primer lugar, para clasificar a los centros entre complejos y no, se utiliza el umbral del 40%.¹⁴ Es decir, los centros entre el 40% con menor nivel socioeconómico son considerados complejos, habiendo, por lo tanto, 202 centros de los 504 clasificados como complejos.

En cuanto a las características diferenciadoras de estos centros, en primer lugar, la proporción de centros complejos es muy similar entre provincias: el 44,6% están ubicados en Tenerife y el 55,4% en Las Palmas, no siendo las diferencias estadísticamente significativas (el 47,3% y 52,7% del total de centros están ubicados en Tenerife y Las Palmas, respectivamente). Sin embargo, mientras que en Canarias hay un 25,2% de centros privado-concertados, solamente siete centros complejos son de titularidad privada (el 3,5%). En cuanto a las variables de nivel socioeconómico de los progenitores, en los centros “complejos”, el 18% de las madres tienen educación primaria o inferior, frente al 8% en los centros no complejos; sin embargo, solo el 21% tiene estudios superiores, frente al 46,7% en los centros no complejos. Con el nivel de ocupación ocurre algo similar: solo el 5% tienen una ocupación alta (técnicos o directivos) frente al 20% en los centros no complejos; por el contrario, el 35,4% de las madres y el 17,8% de los padres están desempleados, de media, en los centros complejos, frente a un 19,6% y 7,8% en los centros con mayor nivel socioeconómico. Las diferencias de origen son en su mayoría significativas, pero menos relevantes: en los centros complejos, en media, el 14,7% de los estudiantes tienen ambos progenitores nacidos en el extranjero, frente al 10% en los centros no complejos; las diferencias son mayores si tenemos en cuenta los progenitores nacidos fuera de la UE: un 12,3% frente al 6,8%. En el caso de los estudiantes, en los centros complejos la diferencia es de menos de 2 p.p., siendo mayor cuando se trata de estudiantes nacidos en países de renta media-baja.

Además de esto, los cuestionarios de contexto arrojan diferencias relevantes. A continuación, se describen aquellas diferencias estadísticamente significativas (ver Tablas 5.a y 5.b). Por el lado de los estudiantes, en los centros complejos la repetición de curso es sustancialmente mayor, hay un mayor absentismo, se observan menos días y tiempo dedicado a deberes y se da un menor sentimiento de pertenencia al centro. Además, la opinión del alumnado sobre el profesorado es peor y el alumnado tiene un autoconcepto menor en matemáticas, lengua, ciencias e inglés.

Por su parte, las familias también revelan patrones interesantes: están menos satisfechas con el profesorado y la dirección, con el trabajo en clase, el nivel de aprendizaje y el centro. No se encuentran diferencias significativas en la frecuencia con la que las familias visitan a los tutores y el centro, pero sí con las expectativas de que sus hijos alcancen educación superior y en el uso de dispositivos: en los centros complejos, el 40,5% de progenitores señalan que sus hijos utilizan el móvil o la tablet casi todos los días (frente al 36,3% en los

¹⁴ El umbral es seleccionado por proporcionar un número de centros suficiente para obtener poder estadístico. Los resultados son mayoritariamente robustos cuando se modifica el umbral al 30% y 50%.

centros no complejos); sin embargo, como cabe esperar al estar más correlacionado con la renta, el uso del ordenador es menor: 42,4% frente a 53,4%.

Tabla 5.a Diferencias entre centros complejos y no complejos: el alumnado y las familias

Porcentaje de <u>estudiantes</u> que declaran...	No complejo (ISEC>40%)	Complejo (ISEC<40%)	Ratio (Complejo / No complejo)
Ser bueno en matemáticas	63,2	54,5	0,86
Ser bueno en ciencias	69,1	60,1	0,87
Ser bueno en inglés	60,1	52,8	0,88
Ser bueno leyendo	77,2	69,6	0,90
Buen ambiente de trabajo	77,4	70,0	0,90
Dedicar >= 30' a deberes	63,6	57,7	0,91
Disfrutar la lectura	60,5	55,3	0,91
Disfrutar las ciencias	63,6	58,2	0,91
Disfrutar las matemáticas	58,1	54,3	0,93
Hacer deberes >= 4 días a la semana	76,7	72,0	0,94
Docentes revisan sus actividades	79,2	74,9	0,95
Docentes proponen actividades interesantes	78,8	75,8	0,96
Disfrutar el colegio	90,6	88,3	0,97
Estar orgulloso de su colegio	91,1	88,8	0,97
Sufrir bullying / sentirse apartado	15,8	17,4	1,11
Tener mayor dificultad que el resto en matemáticas	22,6	25,2	1,11
Tener mayor dificultad que el resto en ciencias	13,0	14,7	1,13
Desear cambiar de centro	10,8	13,2	1,22
Faltar a clase >= 1/mes	16,3	24,3	1,50
Haber repetido curso	9,5	14,6	1,53

Porcentaje de <u>familias</u> que declaran...	No complejo (ISEC>40%)	Complejo (ISEC<40%)	Ratio (Complejo / No complejo)
Que sus hijos usan ordenador o tablet casi todos los días	53,37	42,24	0,79
Estar muy satisfechos con el nivel de aprendizaje	71,09	63,37	0,89
Estar muy satisfechos en general con el centro	73,93	67,07	0,91
Recomendar el centro	84,64	77,45	0,91
Crear que su hijo/a estudiará FPS o universidad	70,42	65,09	0,92
Estar muy satisfechos con el profesorado	74,42	69,09	0,93
Estar muy satisfechos con el trabajo en clase	73,33	68,57	0,94
Estar muy satisfechos con el tutor	81,66	76,63	0,94
Estar muy satisfechos con el uso de recursos	66,98	62,93	0,94
Estar muy satisfechos con la dirección	69,54	65,92	0,95
Que sus hijos usan el móvil casi todos los días	36,33	40,52	1,12

(*) *Muy satisfecho* indica que los progenitores respondieron “bastante o muy satisfechos”

Fuente: Elaboración propia a partir de los cuestionarios de contexto de la ACCUEE | EsadeEcPol

Pero no solo los estudiantes tienen una opinión menos positiva hacia sus docentes, sino que los tutores (docentes) también son más pesimistas respecto a su alumnado en los centros complejos (Tabla 5.b): los asocian con más problemas disciplinarios, menor interés y menor participación en el aula. Además, los tutores identifican en menor medida buenas relaciones con el alumnado y con las familias. Su percepción hacia el centro en general, y la dirección en particular, también es peor: dudan más de su buen funcionamiento, lo perciben como un centro de referencia en menor medida y están menos de acuerdo con que exista un buen ambiente de trabajo. No sorprende, por tanto, que la antigüedad de los tutores en el centro sea de seis años de media en los centros

complejos frente a nueve en los no complejos. En línea con este resultado, solo el 18,2% tienen una rotación del profesorado anual nula frente al 33,7% en los centros no complejos.

Los directores también muestran percepciones más negativas en los centros complejos: resaltan una menor participación y colaboración de las familias como limitación: así, solo el 67% están satisfechos con las familias (frente al 83% en los centros no complejos). Asimismo, destacan la menor colaboración docente como limitación, pero están más satisfechos con el equipo de orientación y las asesorías que en los centros de menor complejidad.

Tabla 5.b Diferencias entre centros complejos y no complejos: profesorado y dirección

Porcentaje de <u>tutores</u> que declaran...	No complejo (ISEC>40%)	Complejo (ISEC<40%)	Ratio (Complejo / No complejo)
Estar en un centro de referencia	39,0	23,9	0,61
Tener muy buena relación con el alumnado	29,7	18,4	0,62
Años de antigüedad en el centro	9,9	6,6	0,67
Que existe muy buena relación entre profesorado y familias	35,1	23,4	0,67
Estar muy de acuerdo en recomendar este centro	58,0	41,5	0,72
Estar muy de acuerdo con que el centro funciona bien	47,8	37,5	0,78
Que las ventajas de la profesión > desventajas	40,7	32,8	0,81
Que los estudiantes tienen buena actitud	84,2	69,5	0,82
Que hay muy buena relación entre docentes y alumnado	49,0	40,9	0,83
Que la dirección promueve el trabajo en equipo	49,9	41,9	0,84
Estar muy de acuerdo con recomendar el centro	55,4	47,5	0,86
Estar muy de acuerdo con que hay buen ambiente de trabajo	54,3	46,7	0,86
Que el alumnado pregunta cuando no entiende algo	81,0	70,5	0,87
Que el alumnado cumple las normas	87,0	76,3	0,88
Los estudiantes tienen interés por la clase y el área	87,5	77,2	0,88
Los estudiantes participan activamente en clase	91,1	82,9	0,91
Los estudiantes realizan las actividades propuestas	91,5	85,7	0,94
Los estudiantes escuchan la explicación del profesor/a	91,2	86,3	0,95
Creer que la profesión docente está valorada por la sociedad	98,2	95,6	0,97
Porcentaje de <u>directores</u> que declaran...	No complejo (ISEC>40%)	Complejo (ISEC<40%)	Ratio (Complejo / No complejo)
Se agrupa al alumnado por elección de materia	13,9	6,9	0,50
Falta plan de formación en el centro	21,8	11,1	0,51
Rotación anual del profesorado nula	33,7	18,2	0,54
Satisfacción con las familias	82,9	67,0	0,81
Satisfacción con el equipo de orientación (EOEP)	82,4	91,1	1,10
Satisfacción con las asesorías del CEP	73,9	85,5	1,16
Profesorado que cambia de centro cada curso	208,5	243,1	1,17
Falta de participación y apoyo de las familias	53,2	73,3	1,38
Falta de colaboración docente	11,1	16,7	1,51
Falta de colaboración de las familias	36,9	60,0	1,63
Rotación del profesorado entre 25% y 50%	22,3	39,2	1,76

Estos patrones apuntan en una clara dirección: los centros complejos (es decir, con alumnado de nivel socioeconómico más bajo) tienen características sustancialmente diferentes a los centros que concentran alumnado de más nivel socioeconómico. Por ello, parece lógico pensar que, para identificar qué diferencia a los centros que atienden

alumnado más vulnerable, pero tienen un valor añadido positivo de los centros que, ateniendo a una población similar, no consiguen esa mejora, debemos comparar a los centros dentro de un grupo de complejidad similar.

Centros educativos “resilientes”

Estimación de la eficacia escolar en centros complejos de Educación Primaria de Canarias

El análisis a continuación se circunscribe a los 202 centros educativos que atienden a alumnado vulnerable y realiza una estimación de su capacidad para mejorar los resultados del alumnado. Para ello, como se mencionaba anteriormente, las estimaciones se llevarán a cabo teniendo en cuenta los centros que se encuentran entre el 40% de centros con menor nivel socioeconómico según el índice socioeconómico: 202 centros y 5.373 estudiantes que realizan las evaluaciones tanto en 6º de EP como en 3º de EP.

La Tabla 6 a continuación presenta los coeficientes del modelo CVA así como la varianza de los resultados explicada por diferencias entre el alumnado (*Var-e*) y por las diferencias entre centros (*Var-uO*) y el porcentaje de centros con valor añadido positivo en las dos competencias analizadas. Las columnas (1) y (3) se corresponden con las estimaciones del modelo CVA - A que incluye únicamente variables de registro administrativo: género del estudiante, origen y edad; por su parte, las columnas (2) y (4) se incluyen como robustez, añadiendo variables socioeconómicas de los progenitores (ambos progenitores con educación superior; al menos un progenitor con ocupación alta).¹⁵

En primer lugar, se observa que los resultados previos tienen un impacto mayor en matemáticas que en lengua castellana y que las niñas consiguen más progreso que los niños en lengua, y similar en matemáticas. Cabe mencionar que los resultados son robustos a la inclusión de no linealidades en los resultados de 3º de EP. Por su parte, el coeficiente asociado a ser niña no tiene un impacto significativo en el caso de competencia matemática pero sí (0,19 d.e) en lengua. Aunque el primer resultado pueda parecer contraintuitivo, es consistente: si no se incluye el rendimiento previo, el coeficiente es negativo —aunque sustancialmente inferior que en otros contextos, como ya apuntaban Cobreros y Gortazar (2023) para el caso de Secundaria en Canarias—, y solo se torna no significativo al incluir el resultado en 3º de EP. Así, mientras que sin controlar por la nota previa las alumnas obtienen un menor rendimiento en matemáticas, cuando se comparan los avances teniendo en cuenta el punto de partida, esta diferencia de género desaparece, sugiriendo que los alumnos y las alumnas mejoran, en promedio, de forma similar en matemáticas, mientras que las alumnas mejoran más en lengua.

¹⁵ Se incluyen estas variables por ser las que tienen un mayor poder explicativo, pero la especificación es robusta a la inclusión de variables similares como el nivel educativo de la madre o el nivel de ocupación de ambos progenitores. Como se comentaba con anterioridad, muchos estudiantes no tienen información de variables socioeconómicas y, además, la información ausente no es aleatoria. Aun así, dado que la relación de estas variables con los resultados es estadísticamente significativa, se incluyen para descartar que no haya cambios relevantes.

El coeficiente de origen extranjero es positivo y significativo, especialmente en competencia en comunicación lingüística. En primer lugar, hay que tener en cuenta que el modelo se estima para una muestra más homogénea de estudiantes en cuanto a su origen; sin embargo, incluso cuando se incluye toda la población de estudiantes y se considera el origen como única variable explicativa de los resultados, el coeficiente no es estadísticamente significativo, en línea, de nuevo, con los hallazgos de Cobreros y Gortazar (2023) para el caso canario. Cabría pensar que esto se debe a un efecto composición del país de origen; por ello, se incluye como variable explicativa una dicotómica que indica si el estudiante ha nacido en un país de renta baja según la clasificación del Banco Mundial: los resultados se mantienen estables. Para profundizar más en este fenómeno, se analiza la distribución de estudiantes nacidos en el extranjero en los 202 centros complejos: el 17,8% han nacido en Venezuela, el 15,7% en Italia, el 11,2% en Cuba y el 6,6% en Marruecos. Profundizando en el nivel socioeconómico de este alumnado, se observa que el 64% de los nacidos en Venezuela, el 37,5% en Italia y el 44,4% en Cuba que acuden a centros complejos tienen al menos un progenitor con formación universitaria (frente al 28% para padres nacidos España); sin embargo, solo entre el 6 y 8% tienen un progenitor con ocupación alta —similar al porcentaje para padres nacidos en España para estos centros—. Así, los niños y niñas nacidas en el extranjero en Canarias que acuden a centros considerados complejos tienen progenitores con formación incluso superior a los nativos, lo cual puede explicar que no existan diferencias estadísticamente significativas cuando únicamente se tiene en cuenta el origen.

Más allá de estas diferencias, se observa que el coeficiente es positivo y estadísticamente significativo cuando se incluye el rendimiento previo y, sobre todo, información socioeconómica de los progenitores. Así, se revela que, en los centros complejos, el alumnado de origen migrante, en promedio, progresa más de lo esperado dado su nivel inicial y la formación y ocupación de sus progenitores. En otras palabras, hay algo en la experiencia o las características de los estudiantes nacidos en el extranjero (que podría ser una mayor motivación, capacidad de adaptación, expectativas, etc.) que, una vez descontadas las diferencias socioeconómicas, se traduce en una ganancia adicional en su rendimiento respecto a sus compañeros nativos.

Respecto al número de centros resilientes —eficaces en contextos complejos— el porcentaje es mayor en competencia lingüística que en matemáticas: 44,3% frente al 38,3%. Comparando ambos modelos, al tener en cuenta información socioeconómica se pierden dos centros educativos en matemáticas y cuatro en lengua, no habiendo diferencias estadísticamente significativas entre el tipo de centros que cambian su estatus. Por ello, siguiendo la aproximación más conservadora, el análisis continúa con el modelo de la especificación (1) y (3), aunque los resultados que siguen son prácticamente idénticos a los de (2) y (4) —dado el bajo nivel de relevancia que tienen las variables de contexto—.

En cuanto al efecto de centro, se observa que la varianza del intercepto aleatorio es de magnitud considerable tanto en matemáticas como en lengua, lo que indica que una parte importante de la variabilidad en los resultados se explica por diferencias entre

centros. De hecho, la correlación intraclase indica que entre el 23% y 27% de la variación de los resultados se atribuye a diferencias entre centros. Estos valores pueden considerarse relativamente altos, en particular en comparación con la literatura internacional donde, en estudios de Primaria, la ICC suele encontrarse entre el 0,10 y el 0,20 (Teddlie & Reynolds, 2000; Marks, 2015; OECD, 2016). Además, la ICC resulta algo mayor en lengua que en matemáticas, lo que sugiere que las diferencias entre centros explican una fracción más amplia del rendimiento en lengua. En conjunto, los resultados confirman que, en el caso de Canarias, el "efecto centro" en Primaria es elevado, y que las características y prácticas escolares desempeñan un papel relevante en la explicación de los aprendizajes del alumnado.

Tabla 6. Resultados de los modelos de valor añadido para los centros con un nivel socioeconómico bajo en Canarias en Primaria: "escuelas resilientes"

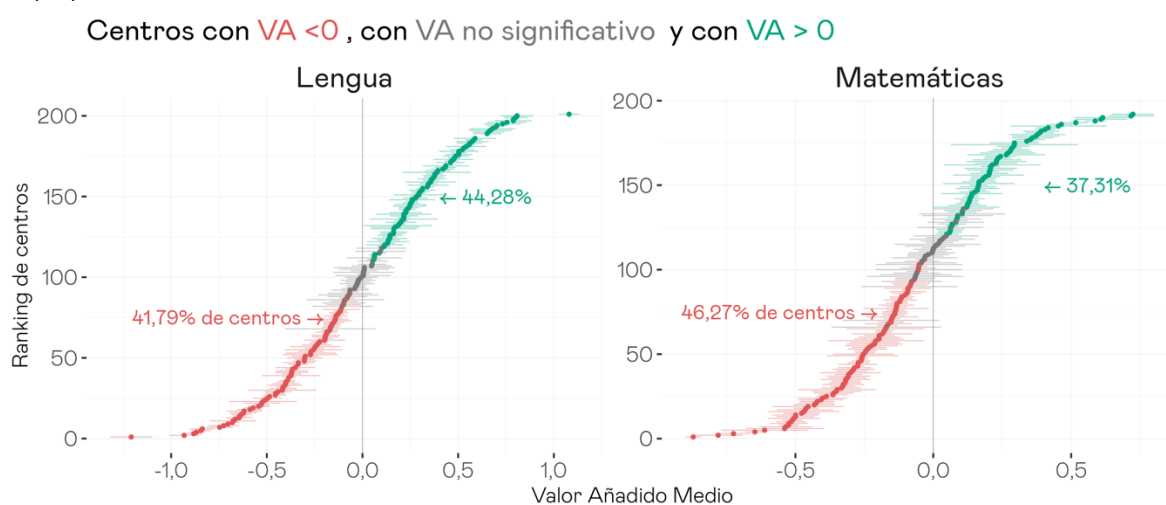
	Matemáticas		Lengua castellana	
Resultados 6º de Primaria	(1)	(2)	(3)	(4)
Resultados 3º de Primaria	0,462*** (0,012)	0,455*** (0,014)	0,356*** (0,012)	0,359*** (0,013)
Género (chica)	-0,001 (0,022)	0,004 (0,025)	0,196*** (0,022)	0,207*** (0,025)
Nacido extranjero	0,049 (0,058)	0,135* (0,070)	0,188** (0,059)	0,123* (0,070)
Año nacimiento	-0,272*** (0,046)	-0,278*** (0,056)	-0,246*** (0,045)	-0,283*** (0,056)
Progenitores estudios superiores		0,161*** (0,058)		0,153** (0,050)
Progenitores ocupación alta		0,145*** (0,047)		0,041 (0,047)
R2 Ajustado	0,205	0,207	0,175	0,176
Var(u0)	0,178	0,174	0,196	0,200
Var(e)	0,573	0,567	0,575	0,555
ICC	0,237	0,235	0,255	0,265
% centros VA>0	38,31	36,32	44,28	41,79
N estudiantes	4965	3847	4936	3821

*Nota: En esta tabla se muestran los coeficientes de regresión del modelo de efectos mixtos con intercepto aleatorio y, entre paréntesis, sus errores estándar. ***p < 0,01; ** p < 0,05; * p < 0,1. Var (u0) representa la varianza del intercepto aleatorio —entre centros—; Var(e) representa la varianza de los residuos —entre estudiantes dentro de cada centro—; ICC es la proporción de la varianza total explicada por las diferencias entre centros; Centros VA>0 (*) indica el número de centros con un valor añadido positivo y

estadísticamente significativo y en la fila inferior, el porcentaje que suponen sobre el total de centros complejos.

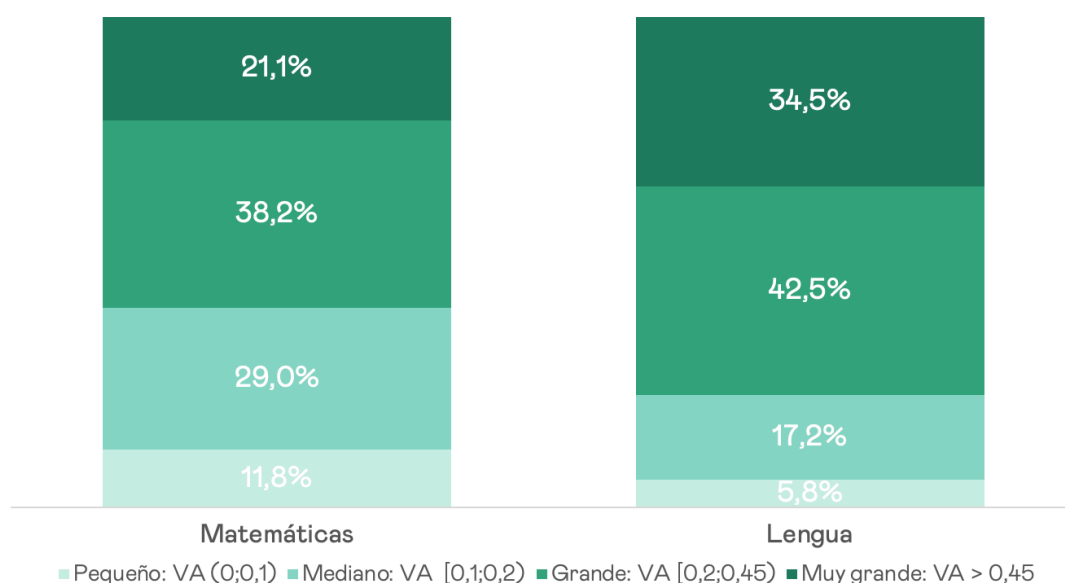
Por último, el gráfico en forma de "oruga" muestra la dispersión de los efectos de valor añadido estimados para cada centro, ordenados de menor a mayor. Se observa, en primer lugar, que muy pocos centros tienen valores extremos —tanto negativos como positivos— y que, como se comentaba con anterioridad, hay muchos más centros con valores muy altos en lengua que en matemáticas. El gráfico 11 complementa al anterior representando, para los centros resilientes —con VA > 0 y significativo—, qué porcentaje consiguen efectos bajos, medios, altos y muy altos. De nuevo, se observa que lengua concentra más efectos grandes o muy grandes: cerca de tres cuartas partes consiguen aportar mucho valor (>0.2 d.e.).

Gráfico 10. Estimación y clasificación del VAM de los centros con nivel socioeconómico bajo, por materias



Fuente: Elaboración propia con datos de la Agencia Canaria de Calidad Universitaria y Evaluación Educativa | EsadeEcpol.
Nota: Los puntos representan el valor añadido estimado para cada centro, y las líneas horizontales indican el intervalo de confianza al 95%.

Gráfico 11. Porcentaje de centros resilientes de Primaria (Canarias) con valor añadido pequeño, mediano, grande y muy grande



Fuente: elaboración propia con datos de la ACCUEE | EsadeEcPol

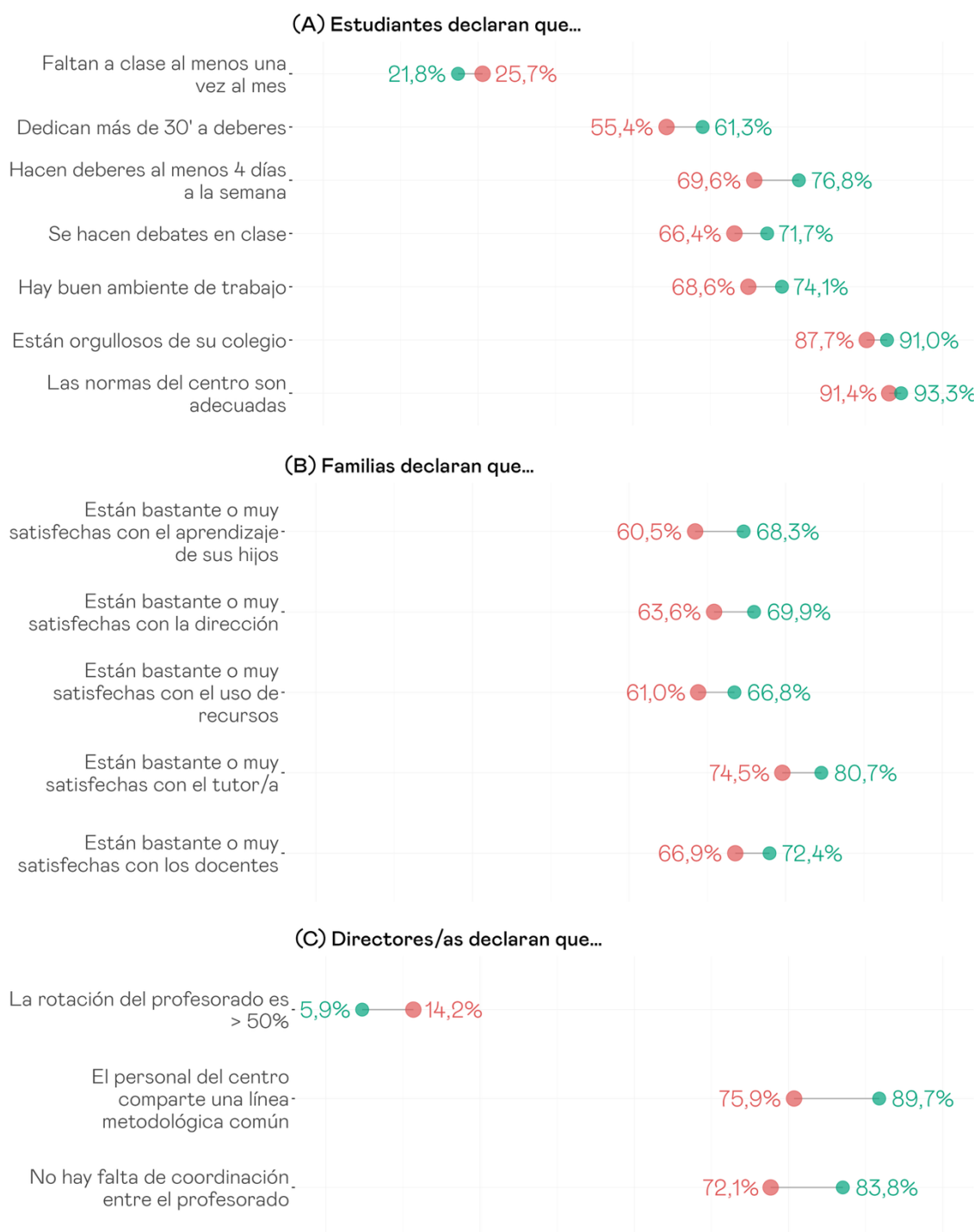
3.4. Caracterización de los centros de Educación Primaria *resilientes* en Canarias

En este apartado se analiza, utilizando los cuestionarios de contexto de estudiantes, familias, directores y tutores, qué caracteriza a los centros complejos con valor añadido positivo en lengua o matemáticas respecto al resto de centros educativos que atienden a alumnado de nivel socioeconómico más bajo. En los gráficos 13 y 14 se presentan todas las diferencias estadísticamente significativas. Aunque hay factores que son comunes a ambas competencias, como se mencionaba con anterioridad, un centro puede presentar un valor añadido positivo en matemáticas y no en lengua, o viceversa, debido a diferencias en la especialización y experiencia del profesorado, las estrategias de enseñanza aplicadas o la disponibilidad de recursos específicos para cada materia. Esto pone de relieve la necesidad de contemplar las particularidades de cada competencia, reconociendo que los procesos de enseñanza y aprendizaje pueden diferir sustancialmente según la asignatura y el contexto en el que se desarrolle. Por último, cabe destacar que estas relaciones reflejan diferencias entre tipos de centros, pero no relaciones causales: resulta complejo diferenciar si los centros son eficaces en conseguir un avance de su alumnado por encima de la media debido los factores observados, o los factores observados son consecuencia de que estos centros sean eficaces.

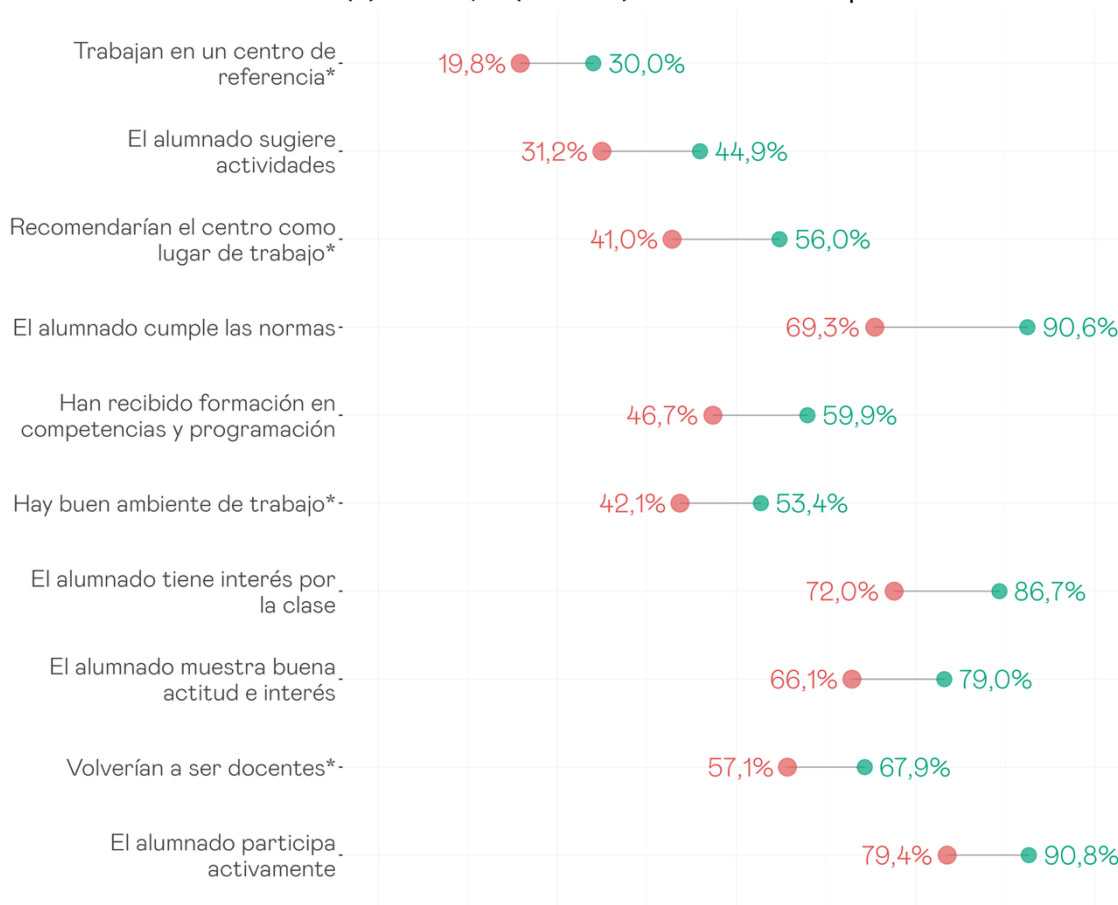
En primer lugar, hay elementos comunes que caracterizan a los centros resilientes en ambas competencias: un ambiente de trabajo más positivo, mayor implicación y participación del alumnado y una relación más sólida entre docentes, familias y dirección, factores que derivan en una mayor satisfacción de todos los agentes con el centro, el aprendizaje y la profesión docente. Estos factores apuntan a que la cohesión interna y la estabilidad organizativa son condiciones compartidas por los centros complejos que logran un valor añadido positivo, independientemente de la materia.

Sin embargo, también hay factores diferenciales en ambas materias. En matemáticas (Gráfico 12), los centros con valor añadido positivo destacan por una mayor implicación del alumnado y del profesorado. En estos centros los estudiantes consideran que hay buen ambiente de trabajo, dedican más tiempo a los deberes y los realizan con mayor frecuencia a lo largo de la semana, participan más en debates en clase y son menos absentistas: un 21,8% faltan al menos una vez al mes frente al 25,7% en centros no resilientes. En paralelo, las familias de estos centros muestran una satisfacción más alta con el aprendizaje de sus hijos, la labor de la dirección y la implicación del profesorado. Por su parte, los tutores señalan más interés y participación del alumnado, cumplimiento de las normas y un ambiente de trabajo más positivo. Así, no es de extrañar que declaren en mayor medida que están muy de acuerdo con trabajar en un centro de referencia en el que volverían a ser docentes. Los directores, por su parte, apuntan a una mayor coordinación metodológica entre el profesorado y una rotación docente más baja, factores que refuerzan la estabilidad organizativa. En conjunto, estas diferencias sugieren que en los centros con VA positivo en matemáticas se combina una mayor motivación y compromiso del alumnado con un clima escolar más estructurado y cohesionado.

Gráfico 12. Porcentaje de estudiantes, familias, directores y tutores que opinan que (...) en los cuestionarios: diferencias entre centros con $VA > 0$ y $VA \leq 0$ en **matemáticas**
Ordenado por tamaño relativo de la diferencia



(D) Tutores/as (docentes) de 6º EP declaran que...



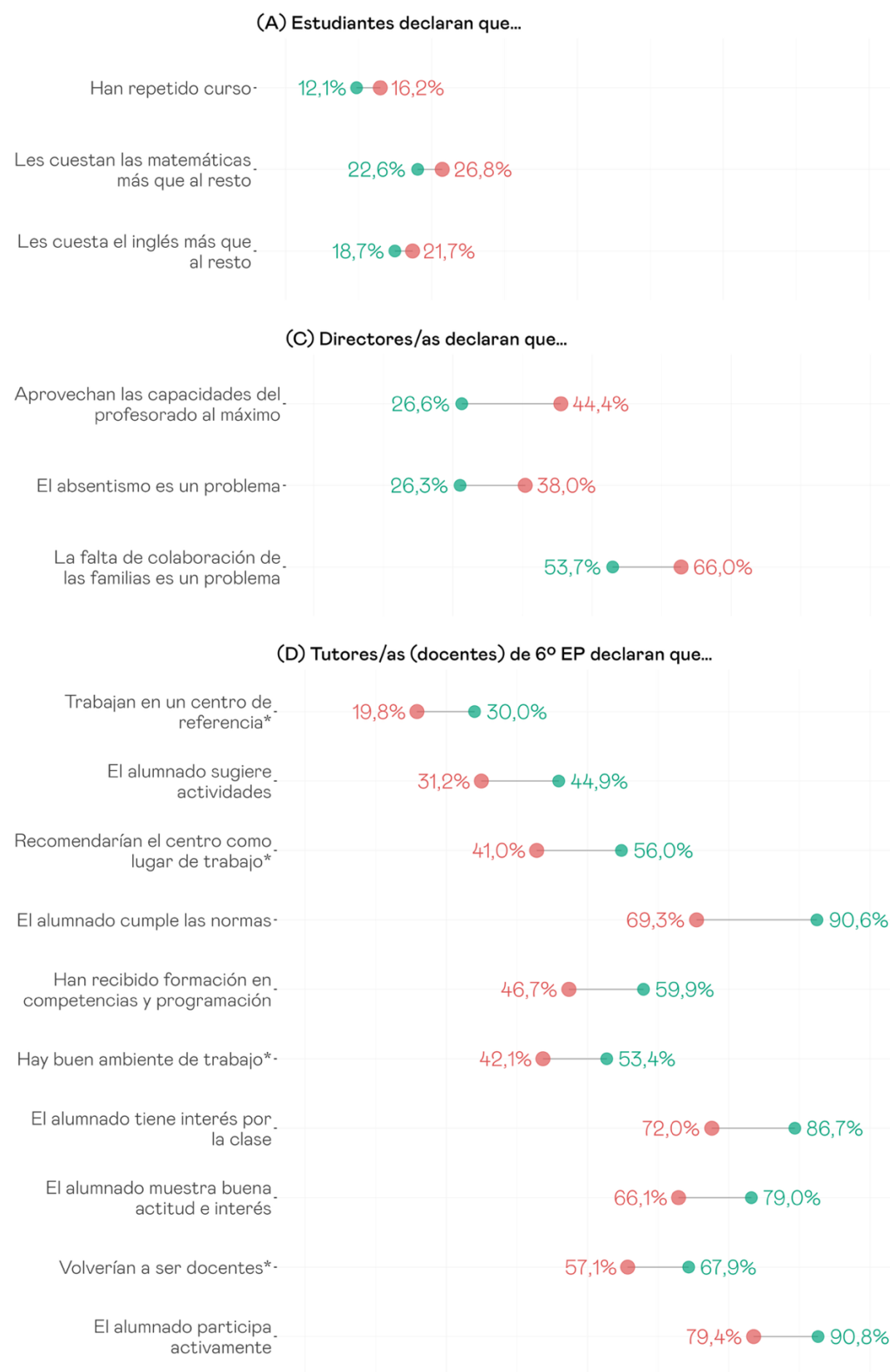
Fuente: Elaboración propia a partir de cuestionarios de contexto de la ACCUEE | EsadeEcPol

(*) Los tutores/as están "muy de acuerdo con que..."; Nota: Todas las diferencias son estadísticamente significativas

Por su parte, en lengua (Gráfico 13), los estudiantes de centros complejos con VA positivo reportan en menor medida haber repetido curso y muestran mayor autoconfianza. Además, los directores de estos centros señalan que el profesorado aprovecha mejor sus capacidades, que el absentismo es menos problemático y, de nuevo, que hay menos problemas de colaboración con las familias. Por su parte, los tutores refuerzan esta visión, destacando una mayor participación activa del alumnado, una mejor actitud e interés en las clases y, de nuevo, muestran mayor satisfacción con el centro y la profesión docente en general.

En resumen, la comunidad educativa en los centros con valor añadido positivo refleja como estos dan más importancia a tener una cultura compartida, un liderazgo más claro por parte del equipo directivo, una mayor variedad metodológica y una mayor colaboración entre equipo directivo, docentes y familias. A su vez, eso se refleja en una mayor satisfacción entre docentes, así como familias y estudiantes, quienes, de alguna manera, perciben el impacto académico de su centro a través de variables relacionadas con el clima escolar. Todos estos aspectos han sido recogidos como clave en la revisión de la literatura y vienen a confirmar la importancia de la cultura escolar, el liderazgo, la innovación y la colaboración.

Gráfico 13. Porcentaje de estudiantes, familias, directores y tutores que opinan que (...) en los cuestionarios: diferencias entre centros con $VA > 0$ y $VA \leq 0$ en lengua
Ordenado por tamaño relativo de la diferencia



Fuente: Elaboración propia a partir de cuestionarios de contexto de la ACCUEE | EsadeEcPol

(*) Los tutores/as están "muy de acuerdo con que..."; Nota: Todas las diferencias son estadísticamente significativas

Finalmente, al imponer la condición de que el centro sea resiliente en ambas competencias, la muestra de centros resilientes se reduce a 40 (Tabla 7). Este tamaño de muestra vuelve difícil la comparativa de variables de los cuestionarios de contexto dada la clara diferencia en la especialización y orientación de los centros y la menor muestra de centros seleccionados, que disminuye el poder estadístico. Además, se está comparando a centros que son eficaces en ambas competencias con centros que pueden no serlo en ninguna o serlo en alguna de las dos; es decir, el grupo de control es muy heterogéneo. Así pues, una parte no menor de las diferencias anteriores se disipan.

Tabla 7. Número (%) de centros complejos en función de su Valor Añadido

	Matemáticas	Lengua	Matemáticas y lengua
VA<0	92 (45,8%)	86 (42,8%)	161 (80,1%)
VA=0	33 (16,4%)	28 (13,4%)	
VA>0	76 (37,8%)	87 (43,3%)	40 (19,9%)
	201	201	201

A continuación, se identifican las diferencias estadísticamente significativas de los centros complejos que tienen un VAM positivo y significativo en ambas competencias en comparación con los que solo son resilientes en una de las dos o, en ninguna.

- ⇒ La tasa de repetidores es mucho menor en estos centros (9,8%) con respecto al resto de centros complejos (15,6%). Los estudiantes tienen más autoconfianza y muestran mayor afección por las materias y el centro. Los docentes también declaran que la motivación y participación del alumnado es mayor, declarando en mayor medida que: tienen buena actitud en clase (81,5% vs 68,3%), cumplen las normas (91,4% vs 73,7%), muestra interés por la clase (87,8% vs 74,7%), toman de notas y/o resúmenes (67,1% vs 48,7%) y sugieren más actividades.
- ⇒ La formación docente también se revela como un factor clave: los tutores en los centros resilientes consideran en mayor medida que la dirección promueve el Plan de Formación del Centro y, concretamente, el 66,2% dicen haber recibido formación en competencias y programación didáctica (frente al 48,1% en centros no resilientes).
- ⇒ Por último, la colaboración docente vuelve a mostrarse como un factor clave: el 21,1% de los directores en los centros no resilientes consideran que la falta de colaboración docente es un problema, frente al 2,9% en los centros resilientes. Igualmente, consideran en mayor medida que la falta de coordinación entre miembros del equipo es un problema. Por su parte, los tutores señalan una mejor relación entre agentes: entre miembros del equipo directivo y familias (41,7% vs. 27,3%), entre equipo directivo y el profesorado (61,1% vs 45,6%) y entre miembros del equipo directivo (74,1% vs 53,1%). De este modo, no sorprende que su concepción del centro sea mejor y su satisfacción no solo con el centro, sino con la profesión, mayor: están más de acuerdo en que el centro funciona correctamente, que hay un buen ambiente de trabajo (61,1% vs 42,7%), el 39,8% creen que es un centro de referencia (frente al 19,7% en centros no resilientes) y el 53,7% de los tutores en centros resilientes lo recomendarían (vs 38,1%); de hecho, el 70,9% (frente al 58,7%) declara que, con mucha probabilidad, volvería a ser docente.

4. Escuelas Resilientes en Educación Secundaria: un caso de estudio para Catalunya

4.1. Datos, variables y complejidad de las escuelas en Catalunya

Fuentes de datos y variables

El análisis del caso de estudio para Catalunya se ha basado en información de estudiantes y centros proporcionada por el Departament d'Educació i Formació Professional y bases de datos abiertas del Instituto de Estadística de Catalunya (IDESCAT) y el Instituto Nacional de Estadística (INE). La principal fuente de información ha sido el Departament d'Educació, que, bajo estrictos acuerdos de confidencialidad, compartió microdatos a nivel de estudiante y datos agregados a nivel de centro entre los cursos 2017/2018 y 2021/2022. Concretamente, en este estudio se utiliza un panel de los estudiantes que se examinaron de las Pruebas de Competencias Básicas¹⁶ en 2017/2018 en 6º de Educación Primaria y en 2021/2022 en 4º de ESO, cursos en los que se llevan a cabo las pruebas, así como información del centro para dichos cursos. Asimismo, se recurre a información pública sobre los centros de Secundaria de Catalunya publicada igualmente por el Departament. Por otro lado, se incluye información a nivel de sección censal procedente de IDESCAT y el INE.

Información a nivel de estudiante del Departament d'Educació

- ⇒ **Resultados de aprendizaje censales del alumnado en las evaluaciones diagnósticas** en 6º de Primaria (2017/2018) y 4º de ESO (2021/2022). Dichas evaluaciones se conocen, en el caso de Catalunya, como las Pruebas de Competencias Básicas. Concretamente, se dispone de información del nivel de competencia en matemáticas, lengua castellana, lengua catalana y ciencias para ambos cursos. Los resultados de 6º de Primaria actuarán como punto de partida y variable explicativa de los principales modelos de esta investigación, mientras los resultados de 4º de ESO estandarizadas serán la variable dependiente. En 4º de ESO, se cuenta con información de 70.914 estudiantes (entre 68.319 y 69.456 con resultados dependiendo de la competencia); sin embargo, de estos, solo entre 59.345 y 60.363 —en función de la competencia— tienen información en 6º de Primaria, siendo estos los estudiantes objeto de análisis, por tener información en ambos cursos. Las calificaciones se estandarizarán en el informe —con media 0 y varianza 1— para facilitar las comparaciones.
- ⇒ **Identificación del alumnado con Necesidades Específicas de Apoyo Educativo (NESE, por sus siglas en catalán)** para los centros del Departament. Se proporciona una variable dicotómica que indica si el estudiante tiene o no un diagnóstico de NESE

¹⁶ Estas pruebas se aplican para comprobar el grado de adquisición de las competencias clave establecidas en el currículo (lingüística en catalán y castellano, matemática y científica), con la finalidad de permitir a la administración educativa disponer de información objetiva sobre el grado de consecución de estas competencias, ofrecer transparencia y rendición de cuentas y proveer de información a la comunidad educativa.

y la clasificación de la tipología de necesidad de apoyo educativo identificada, así como la existencia o no de adaptación curricular. El Departament proporciona 26 categorías en función del tipo de necesidad específica del alumnado, que, por simplicidad, se agruparán de la siguiente forma a lo largo del análisis: (i) alumnado con altas capacidades; (ii) alumnado de incorporación tardía al sistema educativo; (iii) alumnado en situación socioeconómica desfavorecida; y (iv) alumnado con diversidad funcional.

- ⇒ **Información de matrícula para todos los estudiantes del centro**, incluyendo el género, fecha de nacimiento, país de nacimiento, nacionalidad, tipo de documento identificador proporcionado por los progenitores en el momento de la matrícula como proxy de la nacionalidad (DNI, NIE o pasaporte) y código postal de la residencia familiar.

Información a nivel de centro educativo

- ⇒ **Información de tipología, oferta y ubicación** del centro educativo procedente del portal de datos abiertos de la Generalitat (*Directori de Centres Docents Anual*; Concert Educatiu: dades), que incluye la titularidad del centro, la ubicación geográfica y los estudios ofertados.
- ⇒ **Horario y dotaciones del centro educativo**, información proporcionada por el Departament junto a datos de inspección. Incluye el horario de apertura, modalidad de jornada escolar y movilidad del profesorado, así como recursos tecnológicos, número de docentes, tipología de contrato de los docentes, y lenguas ofertadas en el centro.
- ⇒ **Complejidad del centro**. Por un lado, la Generalitat publica de forma abierta una variable dicotómica que indica si los centros son de *màxima complexitat*.¹⁷ Adicionalmente, se dispone de información proporcionada por el Departament — originalmente de Idescat— del nivel educativo y de ocupación de los progenitores, así como una variable de aproximación al origen migrante. Concretamente se facilita información sobre: (i) el número de madres y padres con educación inferior a primaria, educación primaria, educación secundaria y educación superior; (ii) el número de mujeres y hombres con ocupación baja (oficiales y especialistas), media (soporte, administración y otros puestos de oficina) y alta (técnicos y directivos); y (iii) el número de hombres y mujeres que proporcionan DNI, NIE o PASS, cruzado a su vez con el nivel de educación y ocupación.
- ⇒ **Información sobre el bienestar emocional del alumnado a nivel de centro**. La *Enquesta de benestar emocional de l'alumnat* que elaboró el Departament d'Educació en 2022 recoge información anónima y de forma voluntaria sobre el bienestar emocional del alumnado del centro, las relaciones familiares y sociales, el entorno escolar, violencia, abusos y dependencia.

¹⁷ Idescat | Dades obertes: Plantilles del personal docente dels centres i serveis educatius i forma d'ocupació.

Información a nivel de sección censal

Por último, para identificar la complejidad de los centros educativos, se baraja la inclusión de información del nivel socioeconómico de la población que reside en la sección censal en la que se encuentran los centros. Adicionalmente, estas variables se utilizan para completar el perfil socioeconómico del alumnado que realiza las pruebas de competencias; al conocer el Código Postal de residencia, es posible cruzar esta información a nivel geográfico con la información socioeconómica disponible a nivel de sección censal.¹⁸ Concretamente, se han contemplado, para estos fines, las siguientes bases de datos y variables a nivel de sección censal en Catalunya:

- ⇒ **Índice Socioeconómico Territorial (IST)** creado por Idescat, que recoge información de las siguientes variables: población ocupada, trabajadores de baja cualificación, población con nivel de estudios bajos, jóvenes sin estudios secundarios postobligatorios, población extranjera de países de renta media-baja y renta media.
- ⇒ **Información de ingresos, situación laboral, nivel educativo, ocupación y origen migrante** del Atlas de Renta y del Censo de Población (2021) del INE, estando algunas de las variables del Censo cruzadas y desagregadas por grupos de edad.

Construcción del índice de complejidad de los centros

Para identificar a los centros que atienden a alumnado con menor nivel socioeconómico se construye un índice de vulnerabilidad socioeconómica de los centros basado en el **Índex de Complexitat**. Catalunya cuenta con una sólida tradición en la creación de indicadores de vulnerabilidad socioeconómica en el ámbito educativo.¹⁹ Concretamente, el Índex de Complexitat, calculado a nivel de centro educativo, recopila información sobre el nivel educativo y ocupacional de los progenitores, así como sobre la nacionalidad y el país de origen tanto de estos como del alumnado. Además, contempla el porcentaje de estudiantes con Necesidades Específicas de Apoyo Educativo²⁰ (NESE) y aquellos de incorporación tardía al sistema educativo. Específicamente, incluye: (i) el nivel educativo promedio de los progenitores, junto con la proporción con niveles educativos bajos y altos (según el Registro de Población de Catalunya y el Censo de Población); (ii) el nivel ocupacional promedio de los progenitores, y la proporción con niveles bajos y altos (datos

¹⁸ Los indicadores socioeconómicos se asignaron a los códigos postales mediante interpolación areal: a cada código se le imputó el valor promedio de las secciones censales que lo componen, ponderando por la proporción del área de cada sección incluida en el polígono del código postal. Este procedimiento supone una distribución homogénea de la característica dentro de cada sección censal.

¹⁹ La preocupación por reducir las desigualdades escolares, mejorar los resultados del alumnado con mayor desventaja social y combatir la segregación escolar ha impulsado el desarrollo de políticas que buscan garantizar la equidad en el sistema educativo catalán. Tras diversas versiones de indicadores desde 2004, en 2017 se desarrolló un índice más avanzado basado exclusivamente en variables socioeconómicas y, en 2019, Grimaldo-Moreno y López-Iñesta (2021) actualizaron metodológicamente el conocido como Índex de Complexitat, incorporando tanto variables administrativas del alumnado como socioeconómicas y consolidándolo como una herramienta robusta para medir la complejidad de los centros educativos catalanes.

²⁰ En Catalunya denominado alumnado con Necessitats Específiques de Suport Educatiu (NESE), que incluye: (i) con NEE (discapacidad intelectual, auditiva, visual, motora, trastornos autistas, trastornos graves de conducta, otras dificultades en el desarrollo); (ii) con desventaja educativa (por incorporación tardía y situación socioeconómica desfavorecida); y (iii) con otros trastornos de aprendizaje y altas capacidades.

de la Tesorería General de la Seguridad Social); (iii) las proporciones de progenitores sin DNI, nacidos fuera de España o provenientes de países en desarrollo, además de la proporción de alumnado con nacionalidad extranjera (según el Registro de Población de Catalunya); y (iv) las tasas de alumnado NESE y de nueva incorporación al sistema.

Aunque se trata de un índice continuo, ha sido categorizado en cinco niveles que delimitan el grado de complejidad de los centros. El Departament publica el listado de centros públicos categorizados como de máxima complejidad que incluyó 352 centros en el último curso. Sin embargo, ni el valor del índice para cada centro ni la clasificación de los centros restantes se comparten con los equipos de investigación, lo que representa una limitación para el presente estudio.

Dadas las limitaciones de acceso a la información relativa al Índex de Complexitat, se calcula un índice alternativo. Se exploran varias definiciones de vulnerabilidad socioeconómica basadas en los datos proporcionados. Además, la categorización pública de los centros de máxima complejidad permite garantizar una elevada correspondencia entre los cálculos propios y los oficiales, a pesar de posibles diferencias en las variables incluidas. Este enfoque aumenta la fiabilidad del análisis y facilita una evaluación rigurosa de las condiciones de vulnerabilidad en los centros educativos catalanes.

La elección de las variables que se incluyen en el índice se fundamenta, como se comentaba con anterioridad, en su capacidad para explicar la mayor proporción de la varianza, la claridad interpretativa de las cargas asociadas al primer componente principal y su relación con los resultados académicos del alumnado en 4º de ESO. Adicionalmente, se ha utilizado el método LASSO²¹ para asegurar que las variables incluidas en el índice maximizan la relevancia predictiva y minimizan la redundancia. En primer lugar, el Gráfico 1 presenta las variables consideradas por el modelo LASSO, así como su relación con el resultado medio a nivel de centro en las cuatro competencias.

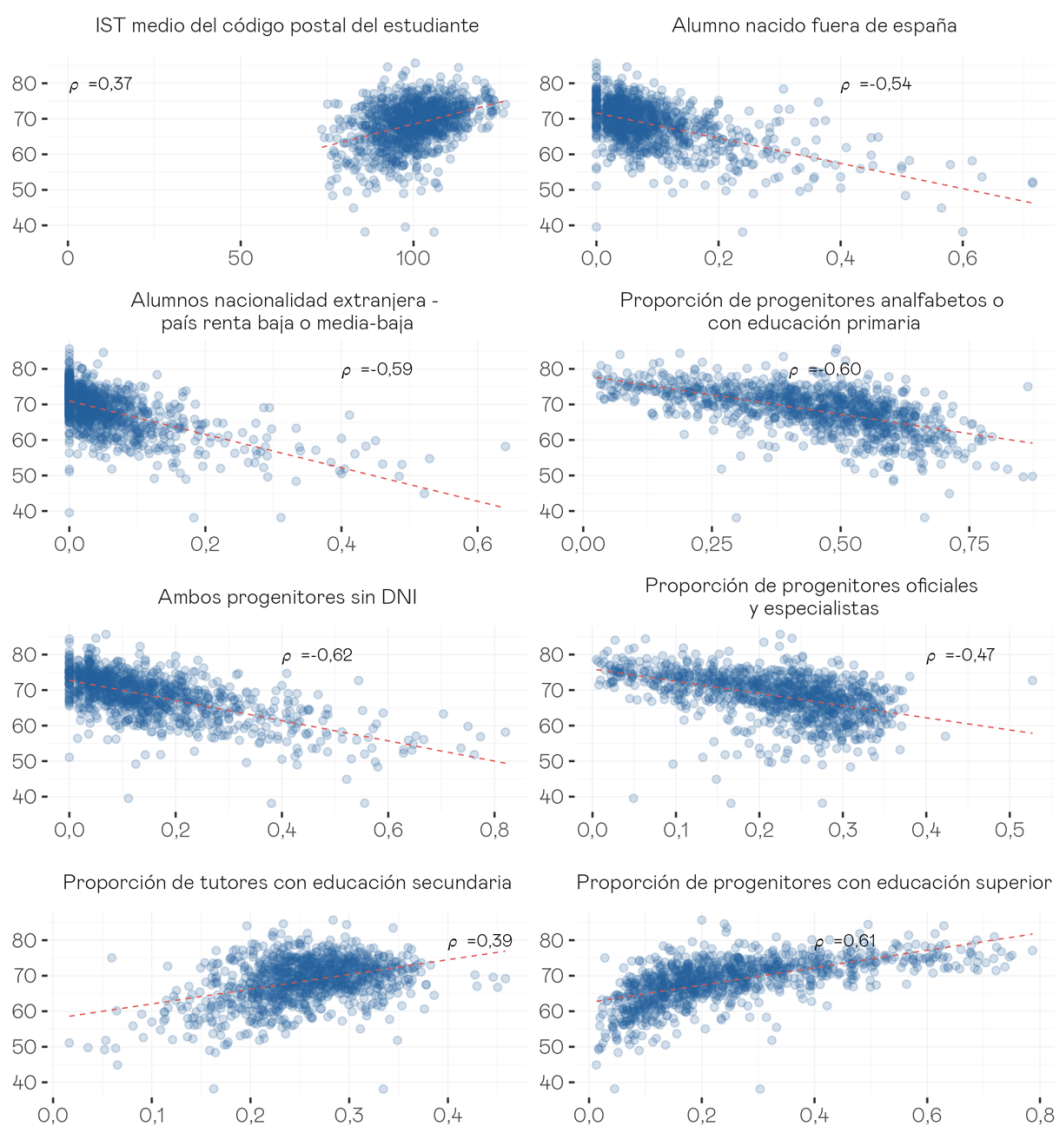
Se observa que existe una fuerte correlación con el porcentaje de estudiantes nacidos en el extranjero o con nacionalidad de un país de renta media-baja, así como con el nivel de ocupación y educación de los progenitores. En el caso del origen, la correlación más elevada con los resultados se da al usar el porcentaje de progenitores sin DNI. Respecto al nivel educativo de los progenitores, y en línea con los hallazgos de la literatura, se observa una correlación de -0,6 con el porcentaje de estudiantes cuyos progenitores tienen como máximo estudios primarios y la calificación media y de +0,6 cuando se trata del porcentaje de progenitores con educación universitaria. Lo mismo ocurre, aunque en menor medida, con la relación entre el porcentaje de progenitores con un nivel de ocupación bajo y los resultados académicos (correlación de -0,47); sin embargo, cabe

²¹ El *Least Absolute Shrinkage and Selection Operator* es un método de regresión penalizada que, al añadir un término de penalización sobre los coeficientes, fuerza a que algunos se reduzcan hasta cero y selecciona así las variables más relevantes. Esto resulta especialmente valioso cuando se cuenta con numerosas variables y se busca un modelo parsimonioso y fácil de interpretar, como sucede al estudiar el efecto de factores socioeconómicos en los resultados de evaluaciones.

recordar que esta variable contiene muchos valores ausentes. Por último, el índice socioeconómico del código postal en el que se encuentran los estudiantes del centro presenta la menor correlación con la calificación media, ligeramente inferior a 0,4.

En segundo lugar, en la Tabla 1 se presentan las cuatro combinaciones de variables consideradas para construir el indicador final y el porcentaje de información capturada con el primer componente de cada PCA. Como puede observarse, en el caso del PCA 2, el primer componente captura el 61% de la información total de las variables consideradas (la máxima variación), si bien las diferencias son sutiles. En dicho indicador, las variables asociadas con un nivel socioeconómico elevado (educación secundaria y terciaria y ocupación alta) tienen una contribución negativa, mientras que las variables de origen migrante, educación y ocupación baja tienen una contribución positiva, siendo un indicador por lo tanto de complejidad, que será utilizado en valor absoluto para representar un índice de nivel socioeconómico.

Gráfico 1. Relación entre la media de los resultados en 4º de ESO y las variables clave para la construcción del índice de nivel socioeconómico de los complejos



Fuente: Elaboración propia con datos del Departament d'Educació i Formació Professional (2021/2022) | EsadeEcPol.

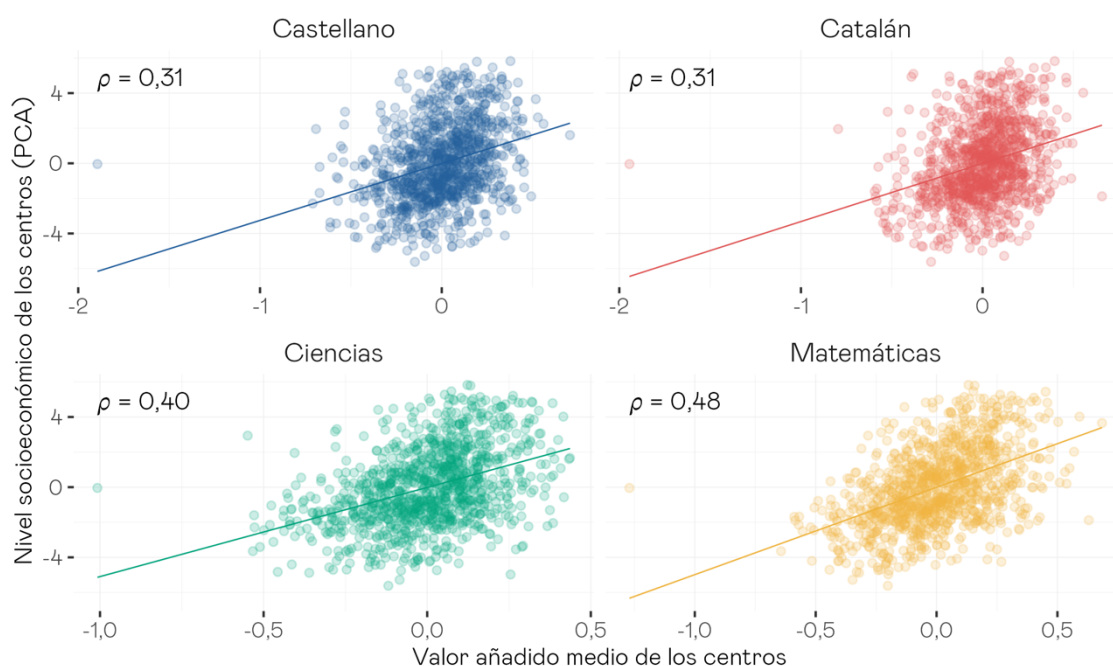
Tabla 1. Variables incluidas en los índices de nivel socioeconómico de los centros y porcentaje de la varianza explicada por el primer componente

	Variables incluidas		% 1er componente
PCA 1	Edu + Inmig	Progenitores con educación primaria; secundaria; superior Progenitores sin DNI; estudiantes nacidos en el extranjero; con nacionalidad de un país de renta media-baja	58,2%
PCA 2	Edu + Inmig + Ocu	PCA 1 + Progenitores con ocupación baja; alta.	60,9%
PCA 3	Edu + Inmig+ IST	PCA 1 + Índice Socioeconómico del CP de residencia del estudiante	57,4%
PCA 4	Edu + Inmig + Ocu + IST	PCA 2 + Índice Socioeconómico del CP de residencia del estudiante	59,1%

El último paso en la toma de decisión del índice es su comparación con los centros públicos clasificados por el Departament como centros de máxima complejidad (CMC). Si bien todos los índices estimados presentan una elevada correlación con la clasificación de CMC, cuando se utiliza el PCA 2, casi la totalidad de los CMC recaen en los dos primeros deciles de PCA 2 —excepto dos que caen en el tercero—, apuntando claramente a la validez del índice como indicador socioeconómico de los centros.

El Gráfico 2 muestra la correlación del índice de nivel socioeconómico de los centros con los resultados de las cuatro competencias en 4º de ESO —el resultado medio en los centros—. La correlación es elevada y mayor para matemáticas y ciencias.

Gráfico 2. Relación entre los resultados a nivel de centro en las cuatro competencias (4º de ESO) y el índice de nivel socioeconómico de los centros



Fuente: Elaboración propia con datos del Departament d'Educació i Formació Professional (2021/2022) | EsadeEcPol.

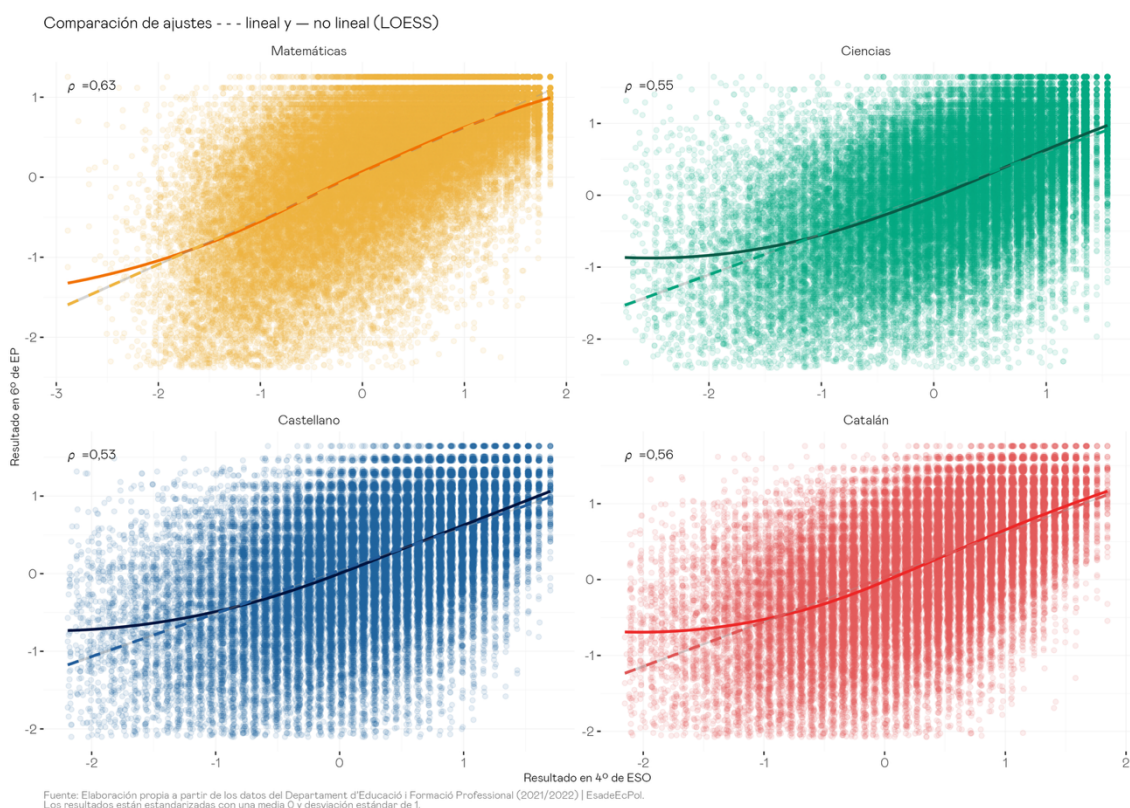
4.2. ¿Cómo estimamos la eficacia escolar en Secundaria?

Comparación entre modelos de Valor Añadido en Catalunya

En este apartado, se lleva a cabo la comparación entre los cinco modelos de eficacia escolar, siguiendo el análisis realizado por Leckie y Prior (2022). Para este fin, se estiman los modelos para las cuatro competencias (es decir, un total de 20 modelos), donde las variables dependientes son los resultados estandarizados en las *Proves d'Avaluació de Competències Bàsiques*. De esta forma, se dispondrá de un conjunto de estimaciones (efectos de escuela) que permitirán evaluar cómo varía la clasificación de los centros según el nivel de ajuste estadístico incluido.

En primer lugar, el Gráfico 3 presenta la relación entre los resultados de aprendizaje en las pruebas de competencias del alumnado entre los cursos de 6º de Primaria y 4º de ESO. Como puede observarse, las correlaciones son elevadas, especialmente en matemáticas (0,63). Esta correlación alta obliga a considerar el uso de un modelo de valor añadido; es decir, a controlar por el aprendizaje previo.

Gráfico 3. Correlación entre los resultados de las competencias en 6º de EP y 4º de ESO



Como se mencionaba con anterioridad, de cara a mejorar la comparabilidad de los modelos, en el caso del modelo sin ajustes se usa la muestra de estudiantes que realizan las pruebas tanto en 6º de Primaria como en 4º de ESO. De este modo:

- ⇒ En el modelo general de valor añadido (VA), se añade el rendimiento previo de 6º de EP para controlar por el rendimiento previo.

- ⇒ En el modelo CVA, además de los resultados en 6º de EP, se añade el género del estudiante, si es de origen migrante de primera y/o segunda generación, si tiene nacionalidad de un país de renta media-baja o baja según la clasificación del Banco Mundial, su año de nacimiento, si es clasificado NESE por desventaja económica y si es clasificado NESE por tener alguna diversidad funcional, siendo todas estas variables individuales y de carácter administrativo.
- ⇒ En el modelo CVA - B se incluye todo lo anterior (CVA) y se añade el rendimiento medio del alumnado en 6º de EP, para capturar el efecto par de rendimiento.
- ⇒ En el modelo CVA - X, se incluye lo anterior y se añade información socioeconómica de los progenitores a nivel de centro: concretamente el nivel educativo y origen, por ser las variables estadísticamente significativas y con un menor porcentaje de valores ausentes.

En la Tabla 2 se presentan los resultados de los distintos modelos para las cuatro competencias: el coeficiente asociado al rendimiento previo, la varianza del intercepto aleatorio (Var-u0 o varianza entre centros), la varianza de los residuos (Var-e), la correlación intraclass (ICC) y el porcentaje de centros con un valor añadido positivo.

En primer lugar, se observa que los resultados tienen un coeficiente estimado elevado —expresado en desviaciones estándar y siempre estadísticamente significativo al 99% de confianza—, siendo especialmente alto en el caso de matemáticas. Además, son estos resultados previos los que explican la mayor parte de la variabilidad en los resultados de 4º de ESO, como muestra el R^2 ajustado en ambas competencias, especialmente en competencia matemática. Cabe mencionar que, además, estos resultados son robustos a la inclusión de no linealidades en el rendimiento previo.

Al añadir variables de contexto del estudiante (CVA), aunque su relación es estadísticamente significativa, la varianza de los resultados explicada por el modelo solo aumenta ligeramente; de hecho, la diferencia es sustancialmente menor que la existente en la literatura para otros contextos (Leckie y Prior, 2022). En el modelo CVA - B, en el que se trata de capturar el efecto de los pares mediante la inclusión de la media de los resultados previos a nivel de centro, se encuentra que el R^2 ajustado también aumenta solo ligeramente, en línea con los resultados obtenidos por Thomas y Mortimore (1996). Sin embargo, esta variable es estadísticamente significativa y con un coeficiente elevado (entre 0,2 y 0.25 desviaciones estándar, dependiendo de la competencia). Por último, en CVA - X, el aumento de la varianza explicada por las variables de nivel socioeconómico es ligeramente superior, siendo únicamente estadísticamente significativa el porcentaje de progenitores con educación superior.

Tabla 2. Resumen de los estadísticos de los principales modelos de valor añadido para las cuatro competencias evaluadas en Catalunya

	Raw	VA	CVA	CVA - B	CVA - X
Resultados individuales 6º de EP		X			
Género, origen, nacimiento, NESE		X	X		
Resultados centro 6º de EP (peer)		X	X	X	
Educación y origen padres - centro		X	X	X	X

Matemáticas					
Coef. Mate 6º de EP		0,626	0,589	0,586	0,587
R ² Ajustado	0,000	0,442	0,461	0,467	0,477
Var(u0)	0,107	0,057	0,054	0,048	0,037
Var(e)	0,838	0,471	0,456	0,456	0,456
ICC	0,113	0,107	0,106	0,096	0,076
% centros VA>0	47,00	44,93	44,79	43,00	42,92
N estudiantes	60363	60363	60296	60296	60256
Ciencias					
Coef. Ciencias 6º de EP		0,569	0,539	0,535	0,536
R ² Ajustado	0,000	0,382	0,394	0,399	0,404
Var(u0)	0,086	0,041	0,040	0,035	0,030
Var(e)	0,816	0,516	0,505	0,505	0,505
ICC	0,095	0,074	0,073	0,064	0,056
% centros VA>0	47,56	47,65	47,51	44,21	43,38
N estudiantes	59345	59345	59280	59280	59240
Castellano					
Coef. Castellano 6º de EP		0,541	0,490	0,487	0,487
R ² Ajustado	0	0,373	0,397	0,402	0,405
Var(u0)	0,096	0,054	0,054	0,049	0,046
Var(e)	0,739	0,468	0,448	0,448	0,448
ICC	0,115	0,104	0,108	0,099	0,094
% centros VA>0	48,17	46,53	46,85	46,20	45,94
N estudiantes	60292	60292	60224	60224	60184
Catalán					
Coef. Català 6º de EP		0,564	0,522	0,519	0,519
R ² Ajustado	0	0,406	0,421	0,425	0,428
Var(u0)	0,095	0,050	0,051	0,047	0,045
Var(e)	0,735	0,443	0,429	0,429	0,429
ICC	0,115	0,101	0,106	0,098	0,094
% centros VA>0	48,87	46,34	47,23	46,29	45,75
N estudiantes	59978	59978	59916	59916	59877

La correlación ICC, que mide qué parte de la variabilidad de las notas de 4º de ESO se debe a diferencias entre centros muestra que, en el modelo sin ajustes, alrededor del 11 % de la variación en matemáticas se explica por el centro, con cifras similares en lengua y algo menores en ciencias. Al incluir el rendimiento previo, la ICC baja ligeramente, ya que este explica mejor las diferencias entre centros que dentro de ellos. Finalmente, añadir variables de contexto socioeconómico a nivel de centro reduce aún más la varianza atribuible a los centros, especialmente en matemáticas.

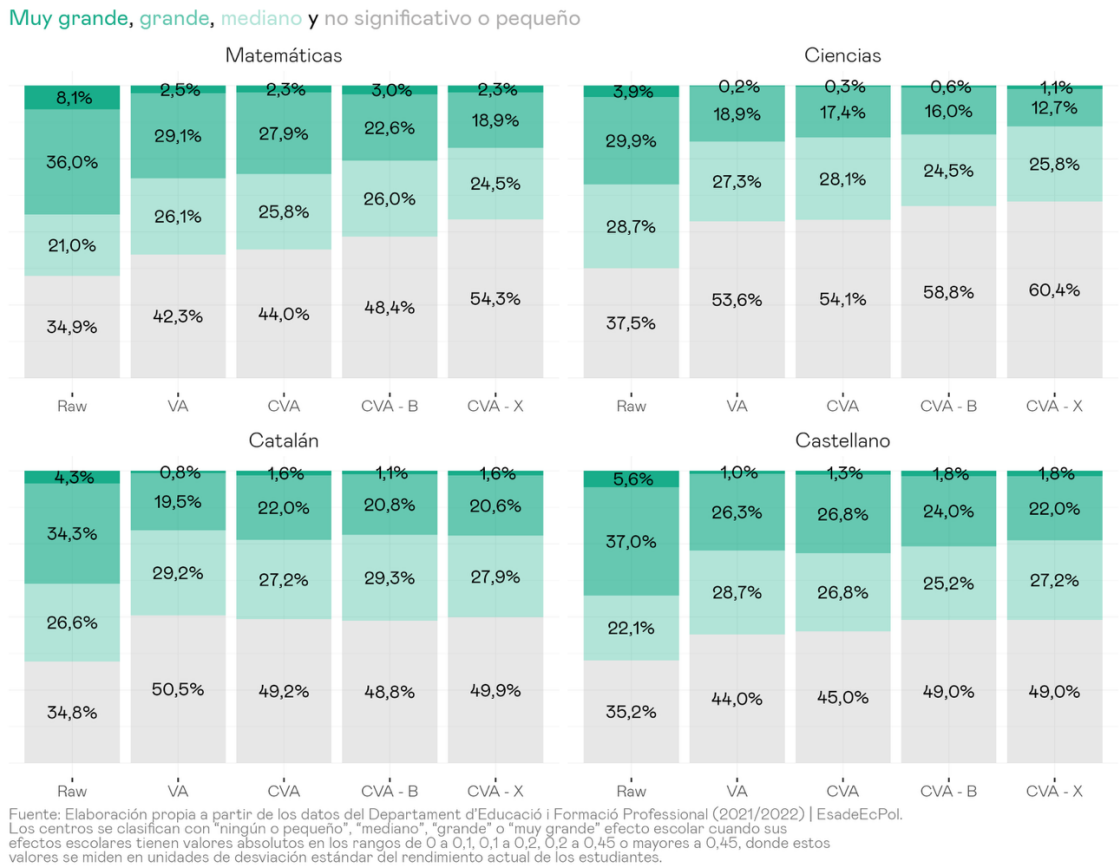
Por último, como cabe esperar, al controlar las diferencias de partida del alumnado, cambia el porcentaje de centros con valor añadido (VA) positivo y estadísticamente significativo —excepto en competencia científica—. Entre materias, el porcentaje de centros con un valor añadido positivo y significativo varía solo entre el 47 y 49%, no habiendo por lo tanto evidencia de que haya competencias en las que, en media y en el caso de Catalunya, los centros logren impulsar el rendimiento del alumnado por encima

de lo esperado dadas sus condiciones iniciales más que en otras competencias. Al controlar por el nivel socioeconómico del centro mediante variables de origen y estudios de los progenitores, de nuevo, disminuye el porcentaje de centros eficaces, pero lo hace solo ligeramente —y de nuevo especialmente en el caso de matemáticas—. Así las diferencias del porcentaje de centros clasificados como eficaces entre modelos al introducir variables de control son inferiores a las encontradas por otros autores.

Tamaño del efecto de los centros

En este apartado, se complementa el análisis de la comparativa entre modelos mediante una comparación de la magnitud de los efectos de los centros educativos para las cuatro competencias, clasificando a los centros con un valor añadido mayor o igual a cero en cuatro categorías según el valor absoluto de sus efectos estimados. Siguiendo las definiciones que utilizan Leckie y Prior (2022) y la *Education Endowment Foundation* para evaluar intervenciones educativas, los cuatro grupos se clasifican en función de la magnitud del VA expresado en desviaciones estándar de los resultados del alumnado en 4º de ESO en cada competencia.²² Como se aprecia en el Gráfico 4, la mayor corrección se produce al pasar del modelo sin ajustes al modelo que ajusta por el rendimiento previo y, por lo tanto, estima el avance real. Entre el resto de los modelos, las diferencias son menores.

Gráfico 4. Porcentaje de centros eficaces en función del tamaño de su valor añadido



²² Nulo o pequeño (de 0 a 0,1 d.e.), mediano (0,1 a 0,2 d.e.), grande (0,2 a 0,45 d.e.) y muy grande (> 0,45)

Llama la atención que el porcentaje de centros con un valor añadido grande o muy grande (superior a 0,2 d.e) es en general bajo, aunque superior en matemáticas y catalán y mucho más bajo en ciencias.

Correlación entre modelos

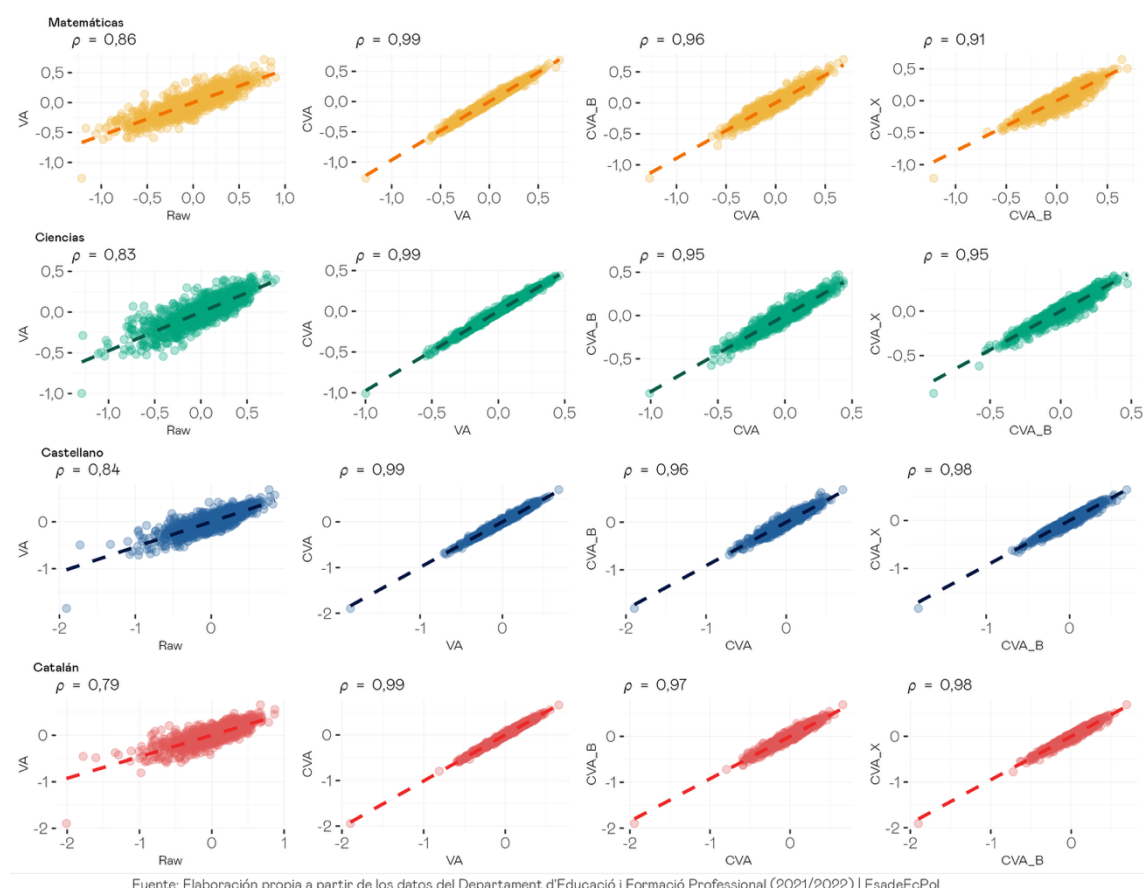
Para terminar, se presenta información de las correlaciones entre el valor añadido de los cinco modelos, así como del ranking de centros. Para facilitar las interpretaciones, se comparan modelos consecutivos, es decir, que comparten un solo cambio en la especificación, siendo además estos donde se encuentran las correlaciones más elevadas (ej. Raw con VA; VA con CVA etc.). Mientras que en el Gráfico 5 se muestra la correlación entre el valor añadido de los centros —independientemente de que este sea negativo, nulo o positivo—, en el Gráfico 6 se pone el foco en la correlación entre aquellos centros con valor añadido positivo, utilizando el ranking de centros.

Lo primero que se observa es que la menor correlación se da entre el modelo sin ajustes y el modelo que ajusta por el rendimiento previo, en línea con los resultados anteriores, siendo esta correlación menor en el caso de catalán. Esto apunta a que la inclusión de la variable de rendimiento previo modifica de manera significativa la estimación del valor añadido de los centros, reordenándolos especialmente en la competencia de catalán. En otras palabras, la medición sin control del rendimiento previo es menos precisa, y al incorporar esta variable se obtiene una evaluación más ajustada de la eficacia de cada centro, lo que refuerza la importancia de considerar las habilidades iniciales del alumnado. Además, en catalán, incluir el nivel inicial del alumnado modifica de forma más notoria la estimación del valor añadido, potencialmente porque en catalán puede existir una mayor heterogeneidad de partida del alumnado —de modo que, si no se tienen en cuenta las diferencias, el valor añadido de los centros queda menos ajustado al contexto real—.

La correlación entre la posición en el ranking de los centros con un valor añadido positivo y significativo es incluso menor, de nuevo, especialmente en el caso de catalán. En competencia matemática, de los 501 centros clasificados como eficaces con el modelo sin ajustes, 82 dejan de ser eficaces cuando se controla por el rendimiento previo y 60 pasan a ser eficaces. Visto de otra forma, de los 100 centros con un mayor valor añadido en matemáticas, 66 pasan a dejar de estar en este top 100 al controlar por el rendimiento previo. En catalán, de los 515 centros eficaces en el modelo sin ajustes, 112 dejan de serlo al incluir el rendimiento previo y 86 nuevos entran en la clasificación. Quizá el resultado más interesante sea el siguiente: al ajustar por el rendimiento previo, los centros “nuevos” que entran a ser clasificados como eficaces son un 72% de titularidad pública —frente al 37% que había cuando no se añadían ajustes—. Entran también más centros con un elevado porcentaje de progenitores con educación baja y de origen migrante. En otras palabras, casi la mitad de los centros que antes no eran eficaces y pasan a serlo son complejos —entendiendo complejos como aquellos que, según el índice anteriormente calculado, se encuentran en los tres primeros deciles de la distribución de nivel socioeconómico—. Esto resultados confirman la evidencia anterior: sin ajustar por el rendimiento previo, los centros que atienden a más estudiantes con un nivel bajo —usualmente asociados a entornos socioeconómicos más complejos— pueden ver

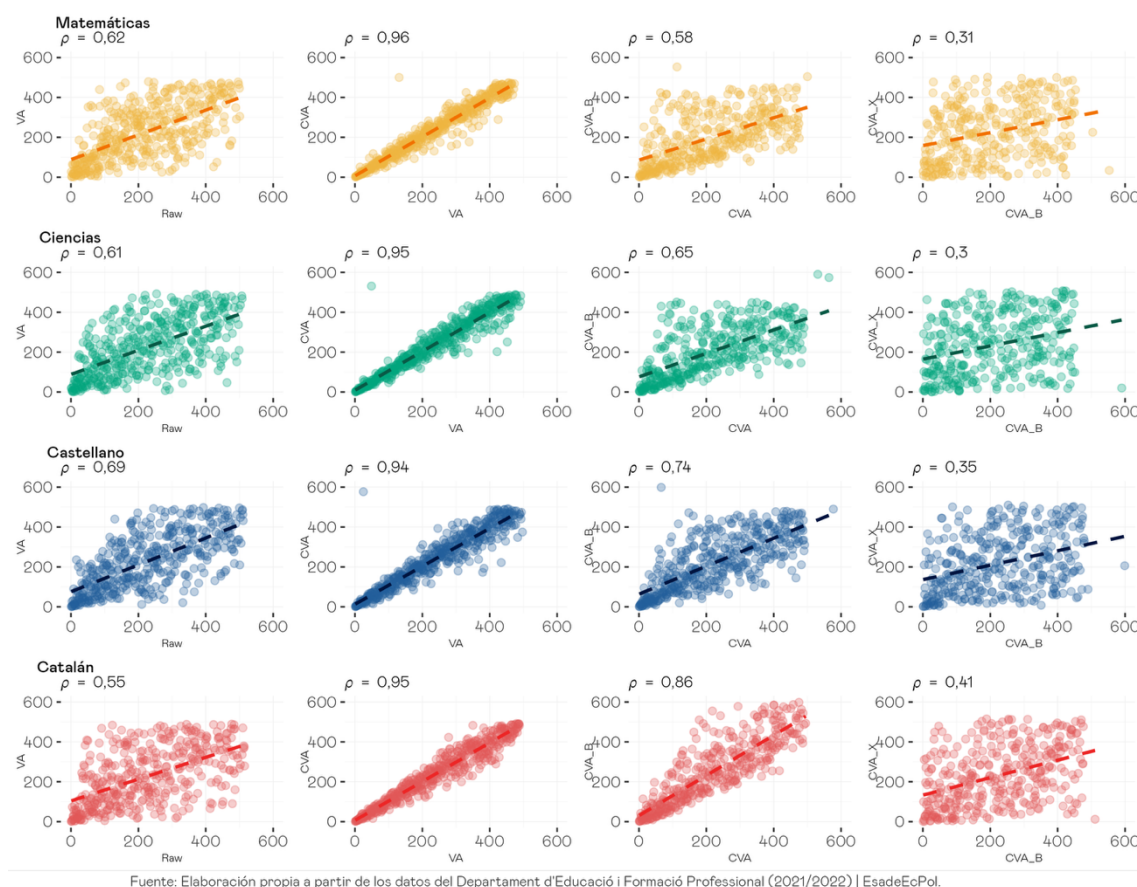
infravalorada su contribución real al progreso del alumnado; por ello, ajustar por el rendimiento previo es esencial para reconocer la labor de los centros que, atendiendo a un alumnado con mayores desventajas, consiguen resultados por encima de lo esperado.

Gráfico 5. Correlaciones entre el valor añadido en **matemáticas**, **ciencias**, **castellano** y **catalán** de los centros en los cinco modelos principales, por pares



En segundo lugar, se compara el modelo VA con el modelo que ajusta por características de los estudiantes: concretamente su género, origen y edad. Como cabía esperar dadas las escasas diferencias en la tabla resumen, la correlación entre el valor añadido en estos modelos es prácticamente total, e incluso cuando se comparan los rankings, las correlaciones son muy elevadas. En el caso de matemáticas, de los 520 centros considerados eficaces, 21 dejan de serlo en CVA y 20 pasan a serlo; de los 100 mejores, 91 continúan siéndolo al controlar por características del alumnado, siendo los números prácticamente idénticos en todas las materias. Aun así, de nuevo, los centros que pasan a ser clasificados como eficaces son en mayor medida complejos, por lo que añadir ajustes de variables socioeconómicas también contribuye a reconocer mejor la labor de aquellos centros que, atendiendo a un alumnado con condiciones iniciales más desfavorables, logra un mayor progreso en el aprendizaje. De esta forma se refuerza la idea de que, cuanto más detallada sea la información sobre las circunstancias de partida del alumnado a nivel individual, más justa y precisa será la valoración del valor añadido de cada centro.

Gráfico 6. Correlaciones entre el ranking de los centros eficaces en **matemáticas**, **ciencias**, **castellano** y **catalán** en los cinco modelos principales, por pares



Por último, aunque la correlación entre el VA de los modelos continúa siendo muy elevada al controlar por variables de centro, la correlación en los rankings de centros eficaces disminuye sustancialmente, especialmente al incorporar la educación y origen de los progenitores a nivel de centro. Concretamente, en el caso de competencia matemática, de los 461 centros clasificados como eficaces en el modelo CVA - B, 134 dejan de serlo y 174 pasan a serlo; de los 100 más eficaces, solo 43 se mantienen en el top —ascendiendo este número a 65 si se compara con CVA en vez de con CVA - B—. En el resto de las competencias, los números son muy similares. Además, incorporar estas variables en el modelo tiene un claro efecto en la composición de los centros seleccionados como eficaces, especialmente en aquellos que lo son en las competencias lingüísticas: disminuye sustancialmente el porcentaje de centros públicos (pasando a un 36%) y cae el porcentaje de centros complejos con nivel socioeconómico bajo.

Selección del modelo CVA y justificación

A la luz de los resultados obtenidos en Catalunya y siguiendo la lógica propuesta por la literatura de eficacia escolar (Thomas & Mortimore, 1996; Timmermans et al., 2011; Leckie & Prior, 2022; Marks, 2021), se justifica la elección del modelo CVA como la especificación elegida en el contexto de este informe. A continuación, se exponen los motivos principales detrás de esta elección que, de nuevo, no es absoluta:

Necesidad de controlar la disparidad de partida: El modelo Raw, que no incluye ningún ajuste, subestima de forma sistemática la eficacia de los centros que escolarizan a un alumnado con niveles de rendimiento previos más bajos y, por extensión, con condiciones socioeconómicas menos favorables. Esto pone de manifiesto que comparar los resultados “en bruto” es impreciso, pues confunde la posible calidad de la labor docente con la composición inicial de los estudiantes y obvia el progreso que hacen los centros.

Ventaja del modelo VA y razón para enriquecerlo con variables de contexto: Incluir el rendimiento previo supone un gran avance frente al modelo bruto, ya que permite aislar buena parte de la contribución del centro al progreso de cada estudiante, tomando en cuenta su punto de partida. Además, incluir variables de contexto del alumnado también hace que aumente el porcentaje de centros con mayor complejidad que son clasificados como eficaces. Es decir, el modelo CVA, al incorporar variables individuales relevantes como el género, el origen migrante o las necesidades específicas de apoyo educativo, brinda un ajuste adicional y reduce posibles sesgos en la comparación de centros con perfiles demográficos sistemáticamente distintos.

Equilibrio entre ajuste y reconocimiento de la complejidad: Frente a la tentación de añadir múltiples capas de variables socioeconómicas a nivel de centro (modelo CVA - X), el modelo CVA mantiene un punto de equilibrio. Por un lado, hace visibles las diferencias individuales más obvias (como el rendimiento previo y el origen), pero sin “descontar” completamente la complejidad socioeconómica del entorno escolar, que en Catalunya puede ser muy relevante para entender la eficacia real de los centros. Ajustar en exceso por variables agregadas a nivel de centro puede llevar a infravalorar el logro de escuelas que, a pesar de tener un alumnado con grandes desventajas, alcanzan niveles de progreso iguales o superiores a los de centros con contextos más favorables.

Identificación de centros complejos con alta eficacia: Como se mencionaba, uno de los hallazgos más interesantes al pasar de VA a CVA en Catalunya es que, al añadir las variables de contexto individual, aumenta la visibilidad de los centros con nivel socioeconómico bajo que logran un alto valor añadido. En términos prácticos, esto significa que, cuando se controla tanto por la nota previa como por las características básicas del alumnado, emergen escuelas de titularidad pública y con mayor presencia de estudiantes de origen migrante o con menor capital cultural, que en un modelo sin ese ajuste habrían quedado ocultas. Precisamente este punto es esencial para las políticas de mejora y equidad: el CVA ayuda a identificar centros que, en circunstancias objetivamente más complicadas, están impulsando el rendimiento de su alumnado por encima de lo esperado.

El modelo CVA, por tanto, cumple dos objetivos a la vez: (i) ajusta por las diferencias individuales más relevantes y (ii) permite seguir comparando centros con distintos niveles de complejidad social. En un escenario como el de Catalunya, en el que coexisten centros con perfiles de alumnado extremadamente diversos, este equilibrio resulta fundamental.

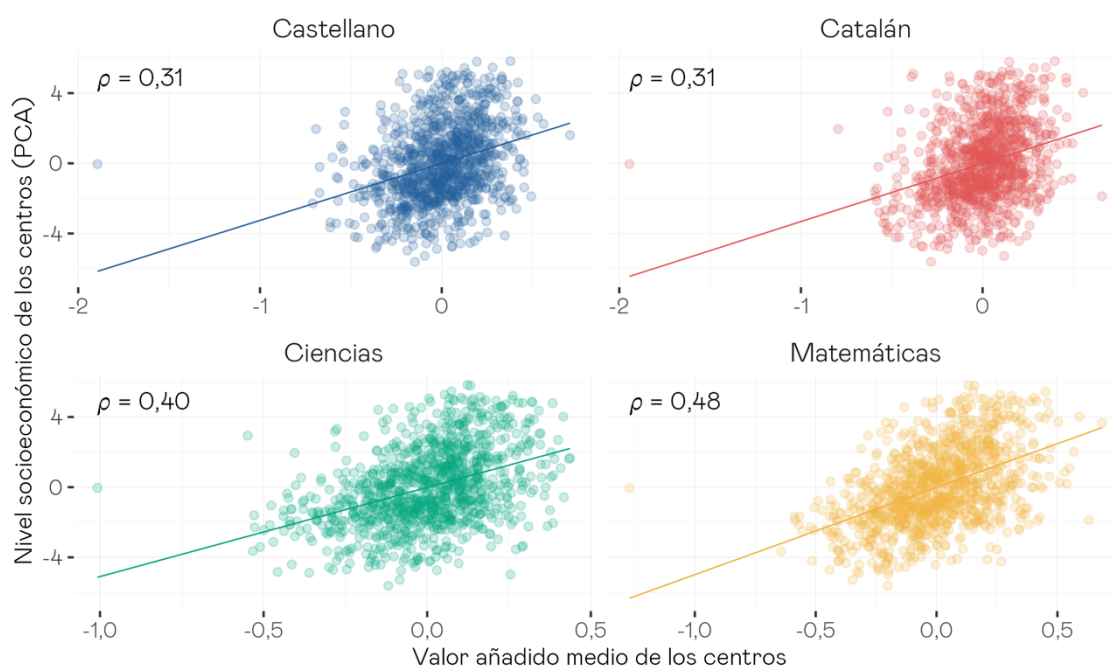
4.3. Relación entre eficacia y nivel socioeconómico de los centros de Secundaria: centros resilientes en Catalunya

Describiendo a los centros en función de la complejidad de su alumnado

El objetivo principal de este informe no es identificar a los centros eficaces, sino a aquellos que, trabajando en circunstancias de mayor complejidad, obtienen una mejora en competencias superior a lo que cabría esperar. Por ello, en este apartado se establecen relaciones entre el nivel socioeconómico de los centros y su valor añadido.

En primer lugar, se observa una correlación positiva entre el valor añadido de los centros —estimado mediante CVA— y su nivel socioeconómico —calculado con el índice de socioeconómico de los centros—. Concretamente esta correlación varía entre 0,3 y 0,5, siendo sustancialmente mayor para competencia matemática y científica que para las competencias lingüísticas. Esto quiere decir que, a medida que aumenta el porcentaje de progenitores con educación superior, ocupación alta o nacionalidad española, tiende a mejorar también el valor añadido de los centros. Así, los centros que acogen a un alumnado con mejores condiciones familiares en Catalunya, suelen obtener, en promedio, un resultado más alto en términos de progreso del alumnado —teniendo esto implicaciones negativas en términos de equidad—. Que la correlación sea más elevada en matemáticas y ciencias puede deberse a que el apoyo familiar o la posibilidad de acceder a recursos extraescolares tienen un impacto más directo en el rendimiento, siendo posible que el alumnado de mayor nivel socioeconómico obtenga más apoyo en estas competencias; sin embargo, en competencias lingüísticas, esta relación puede diluirse algo más.

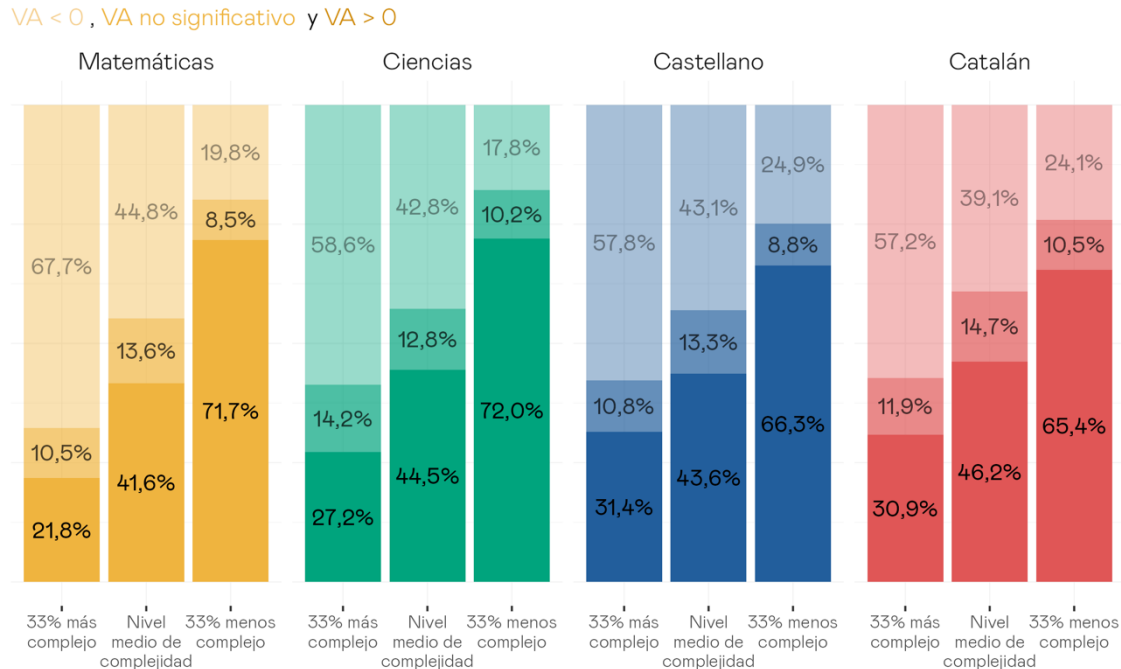
Gráfico 7. Correlación entre el nivel socioeconómico de los centros y su valor añadido



Fuente: Elaboración propia con datos del Departament d'Educació i Formació Professional (2021/2022) | EsadeEcPol.

Al clasificar a los centros por terciles de nivel socioeconómico se observa que, entre los centros más complejos, solamente entre el 20% y 30% son eficaces, frente a cerca del 70 % entre los centros de mayor nivel socioeconómico. En el Gráfico 8 se clasifican los centros en función de su nivel socioeconómico en terciles: el tercil 1 representa al 33% de centros con un nivel socioeconómico más bajo y, por lo tanto, más complejos. Se observa que el porcentaje de centros eficaces entre los centros más complejos es sustancialmente menor que entre los centros con mayor nivel socioeconómico: por ejemplo, en matemáticas, solo el 21,8% de los centros en el tercil 1 son eficaces mientras que en el tercil 3 lo son el 71,7%. En ciencias las cifras son similares, siendo las diferencias menores para las competencias lingüísticas, dada la menor correlación entre nivel socioeconómico y valor añadido: por ejemplo, en catalán, el 30,9% de los centros del tercil 1 son eficaces mientras que el 65,4% de los centros del tercil 3 lo son.

Gráfico 8. Porcentaje de centros con valor añadido negativo, no significativo y positivo en función de su nivel socioeconómico, clasificados por terciles



Fuente: Elaboración propia a partir de los datos del Departament d'Educació i Formació Professional (2021/2022) | EsadeEcPol.
 Nota: El nivel de complejidad se determina utilizando una variable creada a partir de un análisis de componentes principales, que agrupa a las escuelas en terciles: bajo, medio y alto.
 Los terciles reflejan las características socioeconómicas de las escuelas.

Por último, resulta necesario reportar las diferencias existentes entre los centros con menor nivel socioeconómico y el resto. Para ello, en lo que resta de informe, se considerará como centros complejos aquellos que se encuentren entre el 30% de centros con menor nivel socioeconómico (medido con el índice): es decir, 318 centros. Se ha utilizado el umbral del 30% por ser el que delimita con total precisión que todos los centros denominados de *màxima complexitat* (indicador oficial de la Generalitat de Catalunya) se incluyen dentro de este grupo. La otra razón principal para fijar el umbral del 30% es que esto facilita un volumen de centros educativos suficiente para llevar a cabo el posterior trabajo de campo, que requiere una muestra suficiente.

En la Tabla 3 pueden observarse las diferencias entre los centros complejos y aquellos que atienden a alumnado con un nivel socioeconómico mayor. Más allá de las diferencias obvias derivadas de la propia construcción del índice (60,4% de progenitores con educación primaria o inferior frente al 36,8%; 11,1% de progenitores con educación superior vs. 30,6%; 6,5% de progenitores técnicos o directivos frente a 19,4%; y 28,8% de tutores sin DNI frente a 12,6%), los centros con un nivel socioeconómico más bajo son más pequeños (387 alumnos matriculados de media, frente a 470), tienen una mayor rotación del profesorado (22% frente al 15%), los estudiantes viven en barrios con un menor índice socioeconómico, hay más estudiantes con necesidades especiales de apoyo educativo (especialmente económicas) y, como es de esperar, obtienen resultados sustancialmente inferiores en las pruebas de competencias, siendo las diferencias mayores en Secundaria que en Primaria y en matemáticas que en otras materias.

También hay diferencias clave en cuanto a la titularidad de los centros: el 79,3% de los centros complejos son de titularidad pública, mientras que solo el 44,5% de los centros no complejos lo son; o, a la inversa, solo el 20,7% de los centros entre el 30% más complejos son de titularidad privada —concretamente, concertados—. Muy relacionado, se encuentra que, de los centros complejos, el 72,4% tienen jornada continua frente al 39,2% en los centros que no se encuentran entre el 30% de mayor complejidad.

Toda esta evidencia dibuja las mayores dificultades que tienen los centros educativos que atienden alumnado de perfil socioeconómico más bajo para contribuir a la mejora de su aprendizaje de manera significativa como lo hace el resto de los centros. Por ello, su trabajo es doblemente meritorio: son capaces de lograr mejoras de aprendizaje como hace el resto, pero partiendo de unas condiciones sociales que dificultan más el aprendizaje. Así, parece recomendable, en el modelo final, comparar a centros con niveles de complejidad similares, evitando así penalizar injustamente a estos centros que tienen una mayor concentración de alumnado vulnerable. Sin embargo, esta opción es preferible a la introducción de la complejidad socioeconómica del centro como variable de control dentro del modelo, ya que absorbería prácticamente la totalidad de la varianza y se podría pasar por alto el mérito de ciertos centros.

Tabla 3. Caracterización de los centros con un nivel socioeconómico bajo (entre el 30% con menor índice) frente al resto

ISEC - Índice socioeconómico de los centros		> 30%	< 30% "Complejos"
	Número de centros	743 (70,1%)	318 (29,9%)
ISEC de los progenitores (%)	Analfabetos o con educación < primaria	36,8 (0,140)	60,4 (0,079)
	Educación secundaria	28,0 (0,043)	20,0 (0,044)
	Educación superior	30,6 (0,146)	11,1 (0,049)
	Ocupación baja - oficiales y especialistas	18,8 (0,082)	28,4 (0,047)
	Ocupación media - administrativo	24,2 (0,046)	17,2 (0,044)

	Ocupación alta - técnicos y directivos	19,4 (0,094)	6,5 (0,030)
	Sin DNI (en matrícula)	12,6 (0,071)	28,8 (0,130)
Provincia del centro (%)	Barcelona	80,1 (0,400)	59,2 (0,492)
	Girona	7,7 (0,266)	13,8 (0,345)
	Lleida	5,1 (0,221)	9,7 (0,297)
	Tarragona	7,1 (0,258)	17,2 (0,378)
Titularidad (%)	Titularidad pública	44,5 (0,497)	79,3 (0,406)
	Titularidad privada	55,5 (0,497)	20,7 (0,406)
Información del centro	Alumnado matriculado (2021)	470,17 (216,324)	387,37 (126,381)
	Movilidad del profesorado (%)	15,17 (15,440)	22,1 (14,225)
	Jornada continua (%)	39,2 (0,489)	72,4 (0,448)
Información del alumnado	Nacidos fuera España (%)	7,0 (0,052)	14,7 (0,092)
	Nacionalidad no española (%)	9,8 (0,067)	23,1 (0,114)
	Nacidos en país de renta media-baja (%)	38,7 (0,199)	58,1 (0,167)
	ISEC del CP del hogar	103,23 (8,210)	91,8 (7,186)
Alumnado con NESE (4º ESO)	Altas capacidades (%)	0,7 (0,017)	0,4 (0,010)
	Económica (%)	4,6 (0,057)	11,3 (0,119)
	Diversidad funcional (%)	10,1 (0,074)	8,2 (0,063)
Resultados en las pruebas de competencias	Matemáticas (4º ESO)	64,2 (6,653)	55,0 (7,974)
	Matemáticas (6º PRI)	81,6 (3,765)	77,7 (4,802)
	Ciencias (4º ESO)	66,2 (5,116)	58,6 (6,239)
	Ciencias (6º PRI)	73,2 (4,009)	67,6 (5,684)
	Catalán (4º ESO)	76,1 (4,312)	70,2 (5,858)
	Catalán (6º PRI)	75,2 (3,445)	70,6 (4,711)
	Castellano (4º ESO)	76,9 (4,493)	71,7 (5,248)
	Catalán (6º PRI)	79,0 (3,086)	75,7 (3,817)

Fuente: Elaboración propia a partir de datos del Departament d'Educació y Formació Professional (2021/2022) e Idescat.
Nota (1): Las desviaciones estándar se encuentran entre paréntesis; Nota (2): Todas las diferencias entre ambas columnas son estadísticamente significativas al 99% de confianza.

Centros educativos “resilientes”

Estimación de eficacia escolar considerando el nivel socioeconómico

A continuación, se identificarán a los centros complejos —es decir, los 318 centros educativos entre el 30% con menor nivel socioeconómico— con un valor añadido positivo. Esto corresponde a 21.676 estudiantes de un total de 70.914; concretamente, si se tiene en cuenta que los estudiantes tengan resultados de las evaluaciones tanto en 6º de Primaria como en 4º de ESO, la muestra se reduce a entre 16.800 y 17.112.

Cabe destacar que los modelos de valor añadido contextualizado se estimarán para cada una de las cuatro competencias. Así, se considera a un centro resiliente en una competencia concreta siempre que, estando bajo el umbral del 30% de centros con menor

nivel socioeconómico, obtenga un valor añadido medio estimado positivo y significativo en dicha competencia.²³ En el apartado de correlaciones entre modelos se insistió en la relación entre diferentes modelos dentro de una misma materia, pero las diferencias entre materias dentro de un mismo modelo son incluso mayores: la correlación entre el valor añadido en matemáticas y ciencias es de 0,7; con castellano y catalán disminuye a 0,55 y 0,58, respectivamente. Visto de otro modo, de los 100 centros más eficaces en matemáticas, solamente 38 están entre los 100 más eficaces en ciencias, 25 en castellano y 28 en catalán. Por ello, es fundamental estimar la eficacia escolar para cada competencia. Cada una de ellas mide dimensiones específicas del aprendizaje —razonamiento matemático, habilidades científicas, competencias lingüísticas— y puede verse afectada de manera distinta por el rendimiento previo, las características del alumnado y el contexto del centro.

La Tabla 4 presenta los resultados de las estimaciones de valor añadido (CVA) para las cuatro competencias en Secundaria en Catalunya. En matemáticas, todas las variables de contexto resultan significativas, siendo el rendimiento previo el predictor de mayor peso. En este modelo, una vez controlado por el nivel alcanzado en 6º de Primaria, las diferencias observadas reflejan progreso. Así, las chicas muestran un avance relativo menor, mientras que el alumnado de origen migrante registra resultados solo ligeramente inferiores. Las brechas son más acusadas en quienes tienen nacionalidad de países de renta media-baja, así como en los estudiantes repetidores, con nivel socioeconómico bajo y con diversidad funcional.

El bajo coeficiente asociado al origen confirma que buena parte de la desventaja inicial asociada al origen migrante se explica por diferencias de partida y condiciones socioeconómicas. Al estimar el mismo modelo con el origen migrante como único regresor, el coeficiente alcanza -0,39 d.e.; al introducir el rendimiento previo se reduce a la mitad y se reduce a 0,06 d.e. al añadir el resto de las características. Esto indica que, una vez se controlan los factores iniciales, el progreso académico del alumnado de origen extranjero tiende a asemejarse al de sus compañeros en matemáticas. En ciencias se observan patrones similares: el efecto del origen migrante desaparece por completo al controlar por el nivel previo y otras covariables, pese a que como único regresor presenta un coeficiente de -0,32. También las diferencias de género y nivel socioeconómico resultan menos acusadas que en matemáticas,

En las competencias lingüísticas, los resultados difieren. Las chicas no solo obtienen mejores calificaciones finales —por ejemplo, 0,33 d.e. en castellano—, sino que, tras considerar el nivel previo y otras características, muestran un progreso

²³ En este trabajo se ha optado por definir la resiliencia de manera específica para cada competencia: un centro se considera resiliente en matemáticas, lengua castellana, lengua catalana o ciencias si, siendo complejo, presenta un valor añadido positivo en la respectiva materia. Esta aproximación permite identificar hasta cuatro formas de resiliencia, de modo que un mismo centro puede serlo en una competencia y no en otra. La elección no es única: cabría, por ejemplo, agrupar bajo un único indicador de resiliencia científico-matemática y otro lingüístico, o bien considerar como resilientes únicamente a los centros eficaces en varias competencias a la vez (2, 3 o todas). La definición seleccionada, sin embargo, se considera la más adecuada para este análisis y la más conservadora en términos de interpretación.

significativamente mayor (0,15 en catalán y 0,19 en castellano). De nuevo, el origen migrante vuelve a mostrar un efecto elevado cuando no se tiene en cuenta ningún factor (-0,37 en castellano y -0,30 en catalán), pero se reduce sustancialmente al introducir el rendimiento previo y, en el caso del catalán, deja de ser estadísticamente significativo. Sin embargo, en estas competencias el alumnado con nivel socioeconómico bajo, con diversidad funcional y repetidor, obtienen un progreso inferior que sus pares.

Tabla 4. Resultados de los modelos de valor añadido para los centros con un nivel socioeconómico bajo en Catalunya en Secundaria: “escuelas resilientes”

<i>Resultados 4º ESO</i>	Matemáticas	Ciencias	Castellano	Catalán
Resultados en 6º de EP	0,56** (0,00529)	0,56** (0,00591)	0,50** (,00575)	0,54** (0,00557)
Género (chica)	-0,18** (0,0108)	-0,11** (0,0117)	0,19** (0,0114)	0,15** (0,0109)
Migrante de 1ª o 2ª generación	-0,06** (0,0163)	0 (0,0179)	-0,07** (0,0174)	0,01 (0,0168)
Nacionalidad - país renta media-baja	-0,09** (0,0250)	-0,02 (0,0275)	0,01 (0,0266)	0,03 (0,0257)
Año de nacimiento <=2005	-0,26** (0,0252)	-0,31** (0,0276)	-0,36** (0,0269)	-0,29** (0,0257)
NESE - Diversidad funcional	-0,10** (0,0216)	-0,17** (0,0235)	-0,25** (0,0233)	-0,21** (0,0223)
NESE - Económico	-0,13** (0,0210)	-0,13** (0,0230)	-0,19** (0,0224)	-0,15** (0,0216)
R² Ajustado	0,46	0,41	0,41	0,43
Var(u0)	0,05	0,04	0,05	0,05
Var(e)	0,47	0,56	0,54	0,49
ICC	0,10	0,06	0,08	0,10
Centros VA > 0 (*)	138	136	133	144
(%) VA > 0	43,40	42,77	41,82	45,28
N estudiantes	17096	16778	17090	16956

*Nota: En esta tabla se muestran los coeficientes de regresión del modelo de efectos mixtos con intercepto aleatorio y, entre paréntesis, sus errores estándar. ***p < 0,01; ** p < 0,05; * p < 0,1. Var (u0) representa la varianza del intercepto aleatorio —entre centros—; Var(e) representa la varianza de los residuos —entre estudiantes dentro de cada centro—; ICC es la proporción de la varianza total explicada por las diferencias entre centros; Centros VA>0 (*) indica el número de centros con un valor añadido positivo y estadísticamente significativo y en la fila inferior, el porcentaje que suponen sobre el total de centros complejos.

En la segunda parte de la tabla se presenta además del R² ajustado —resultado de un modelo de regresión tradicional en el que se agrupan los errores a nivel de centro—, la varianza del intercepto aleatorio —o la varianza de los resultados explicada por los centros—, la varianza de los residuos —varianza entre estudiantes dentro de cada centro—, la correlación intraclase —proporción de la varianza total explicada por la varianza entre centros— y el porcentaje de centros que son eficaces en cada competencia; es decir, que tienen un valor añadido positivo y estadísticamente significativo.

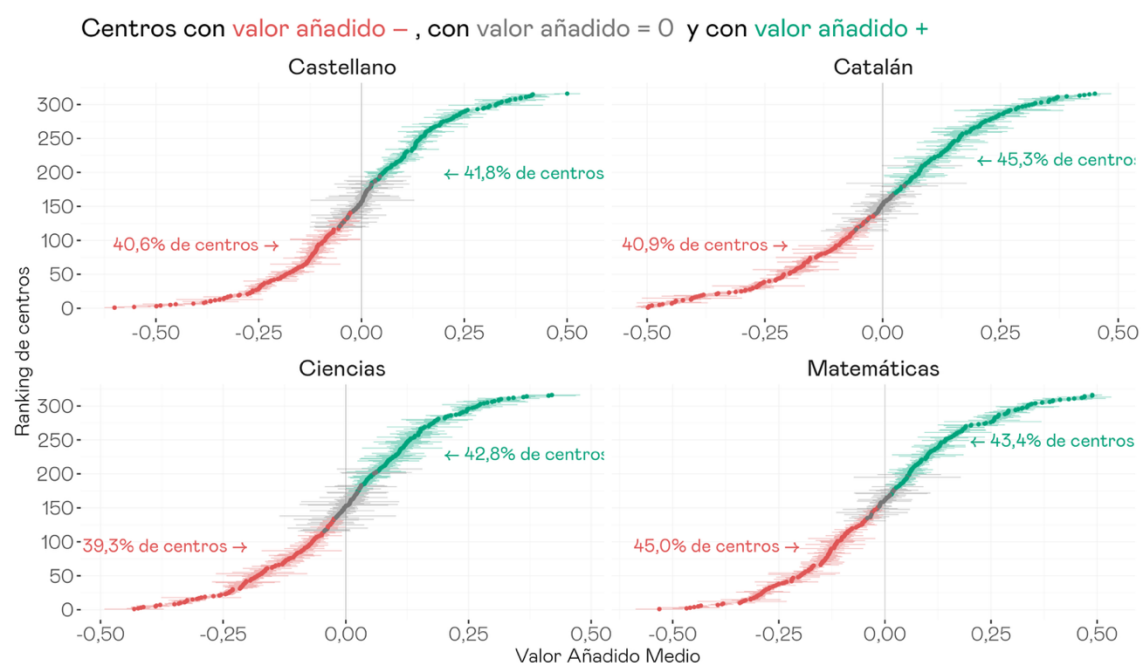
Respecto a la varianza entre centros, esta se sitúa en valores similares y ciertamente bajos (0,04-0,05) en matemáticas, castellano y catalán, siendo ligeramente menor en ciencias; esto es lógico si se tiene en cuenta que se están comparando centros que tienen

un nivel socioeconómico similar. De hecho, si se controla por el nivel socioeconómico de los progenitores en la regresión, esta varianza tiende a cero, ya que queda muy poco margen de variación atribuible al efecto centro, dominado por las características socioeconómicas. Por otro lado, la variabilidad entre estudiantes dentro de cada centro es sustancialmente mayor —siendo de mayor magnitud en ciencias (0,56) y menor en matemáticas (0,47)— lo que indica que las diferencias individuales en ciencias tienden a ser algo más marcadas y que, en general, como se observa con la correlación intraclase, la proporción de la varianza explicada por las diferencias entre centros es baja, siendo ligeramente mayor en ciencias que en otras competencias.

Por último, el porcentaje de centros que tienen un valor añadido positivo y que, por lo tanto, son eficaces, es sustancialmente mayor al observado en el Gráfico 8, al estar comparando a centros con un nivel socioeconómico similar. Asimismo, varía el porcentaje de centros eficaces entre competencias, siendo menor en castellano (41,8%) y mayor en catalán (45,3%).

Los resultados de valor añadido se presentan también en el gráfico de “oruga” (Gráfico 9), que ordena los centros por su valor añadido estimado (de menor a mayor) incluyendo el intervalo de confianza, que depende fundamentalmente del tamaño del centro. Así pues, los centros educativos con un valor añadido estimado no significativo están dibujados en gris; los centros con valor añadido positivo y significativo aparecen en verde; y aquellos con valor añadido negativo y significativo aparecen en rojo. Se aprecia una elevada distorsión en los extremos y muy pocos centros con un valor añadido muy elevado, especialmente en ciencias.

Gráfico 9. Estimación y clasificación del VAM de los centros con nivel socioeconómico bajo, por materias

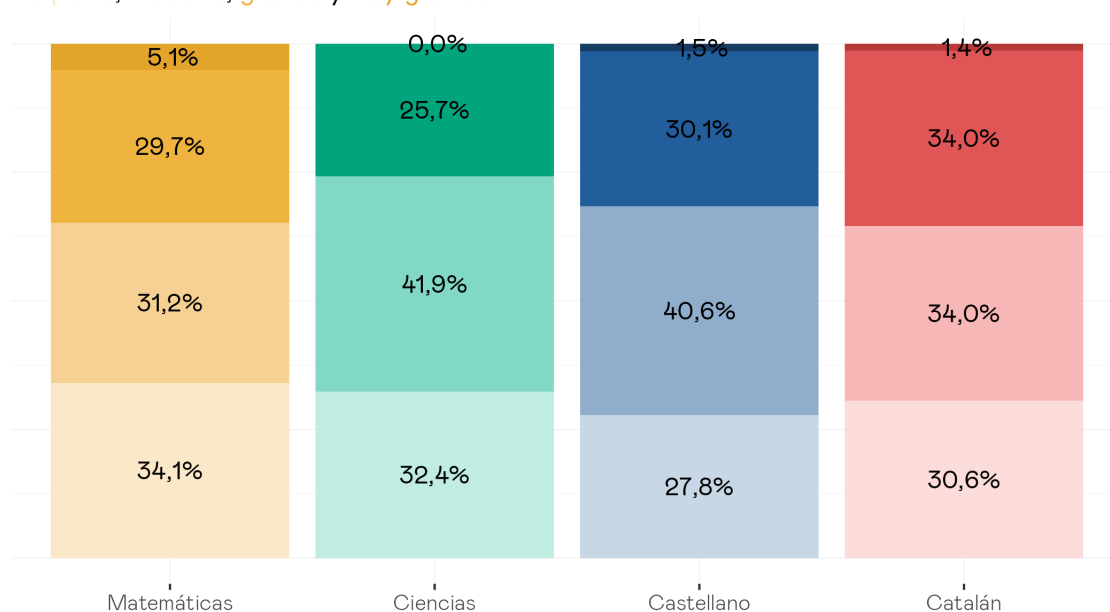


Fuente: Elaboración propia a partir de los datos del Departament d'Educació i Formació Professional (2021/2022) | EsadeEcpol
Nota: Los puntos representan el valor añadido estimado para cada centro, y las líneas verticales indican el intervalo de confianza al 95%.

Para profundizar en esta idea, el Gráfico 10 presenta qué porcentaje de centros resilientes —con VA > 0 y significativo— tienen un valor añadido pequeño, mediano, grande y muy grande. Se observa que el número de centros con valor añadido grande o muy grande es mayor en castellano y bajo en ciencias. Concretamente, prácticamente no hay centros con valor añadido muy grande (más de 0,45 d.e.): solo siete en matemáticas, dos en competencias lingüísticas y ninguno en ciencias. Esta escasez contrasta con lo observado en Primaria en Canarias, donde la proporción de centros complejos con valor añadido muy grande era sustancialmente mayor en todas las materias. Esto va en línea con los hallazgos de la literatura (OCDE, 2016; Hattie, 2009), que encuentran que el efecto del centro (y profesor) se diluye en Secundaria y otros factores externos —como la familia y los pares— ganan peso.

Gráfico 10. Porcentaje de centros resilientes en función del tamaño de su valor añadido, por competencias

Pequeño, mediano, grande y muy grande



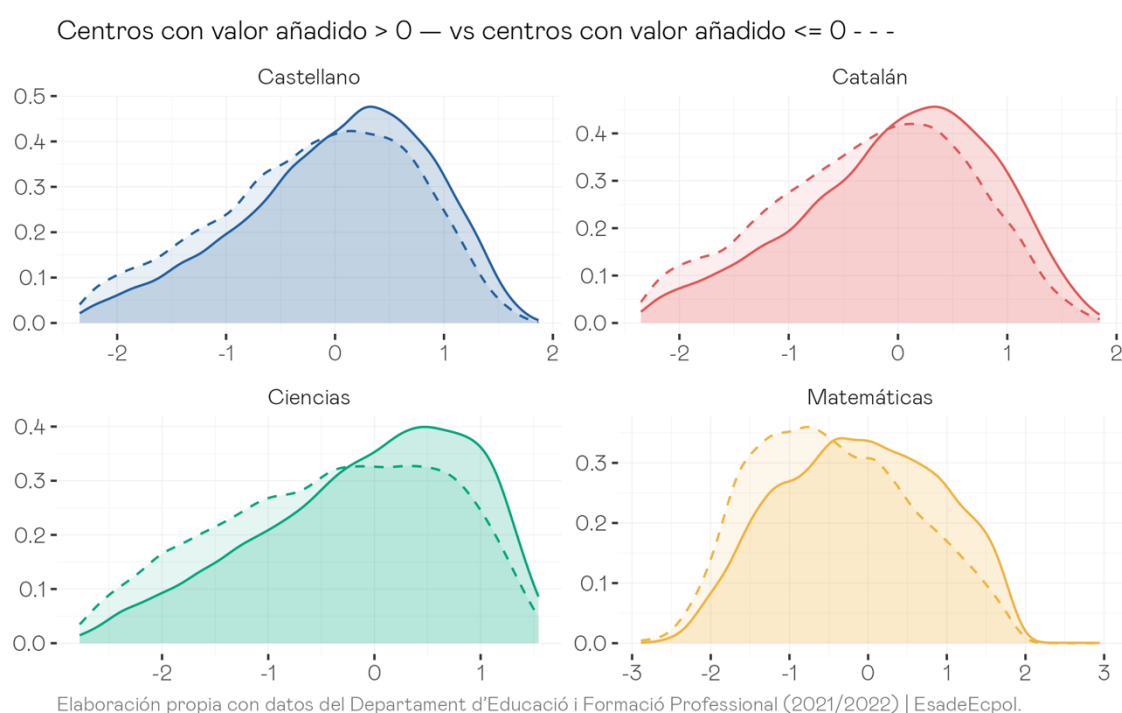
Fuente: Elaboración propia a partir de los datos del Departament d'Educació i Formació Professional (2021/2022) | EsadeEcPol. Los centros se clasifican con "pequeño", "mediano", "grande o muy grande" efecto escolar cuando sus efectos escolares tienen valores absolutos en los rangos (0;0,1], (0,1;0,2], (0,2;0,45], >0,45 desviaciones estándar

4.4. Caracterización de los centros de Educación Secundaria *resilientes* en Catalunya

A continuación, se presenta una caracterización de los centros resilientes a partir de datos de centro y del cuestionario de bienestar del alumnado (2022). Aunque existen factores comunes de éxito escolar, un centro puede ser resiliente en una competencia y no en otra. Por ello, los factores asociados a la resiliencia se analizan por separado en cada competencia. Cabe destacar que las diferencias presentadas en esta sección representan asociaciones entre tipos de centros, pero no permiten establecer causalidad.

En primer lugar, en el Gráfico 11, se observa que los centros que tienen valor añadido positivo en cada una de las competencias tienen resultados en 4º de ESO superiores a los centros con nivel socioeconómico bajo que no obtienen un valor añadido positivo, siendo esta diferencia más acusada en ciencias que en el resto de las competencias. Es decir, no solo tienen un mayor progreso, sino que tienen también mejores resultados.

Gráfico 11. Distribución de los resultados por competencias: centros con valor añadido positivo frente a cero o negativo

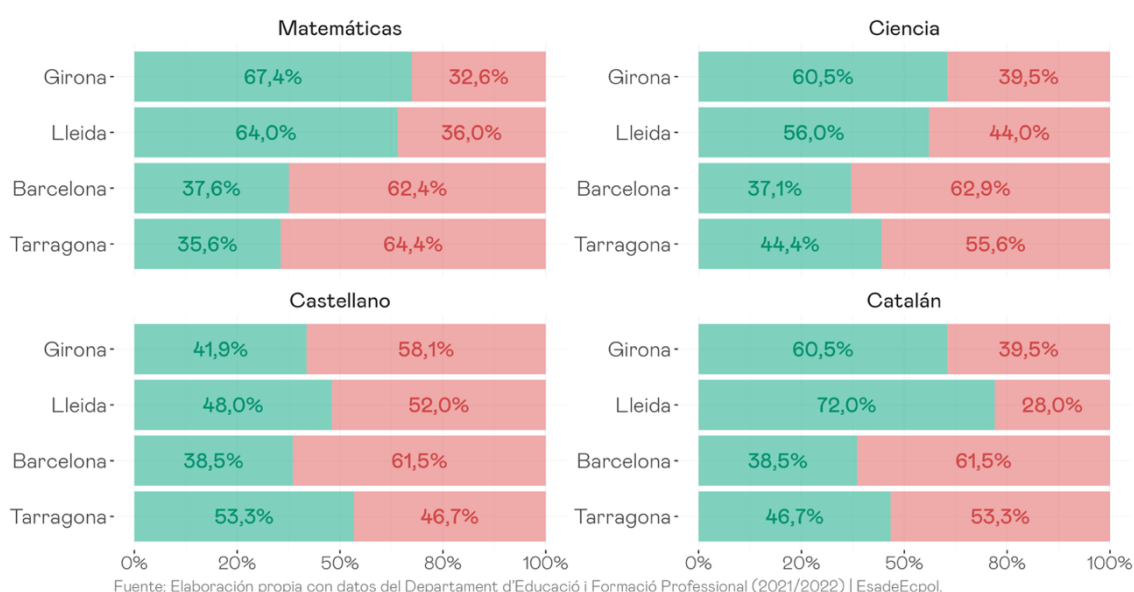


La caracterización territorial de los centros resilientes en Secundaria muestra contrastes muy marcados entre provincias (Gráfico 12). Girona y Lleida concentran una proporción significativamente mayor de centros con valor añadido positivo en competencias de matemáticas y ciencias, mientras que Tarragona, y sobre todo Barcelona, tienen un porcentaje claramente menor de centros resilientes. Esta desventaja se acentúa en el caso de la competencia de catalán, donde apenas un tercio de los centros de Barcelona logra un valor añadido positivo, frente a más de la mitad en Lleida y Tarragona. El análisis de regresión logística confirma estas disparidades: un centro ubicado en Barcelona tiene una probabilidad de ser eficaz en matemáticas o ciencias más de 16 puntos porcentuales menor que uno situado en cualquier otra provincia, y en catalán la brecha asciende a 18,9 puntos. En términos relativos, ello implica las provincias catalanas que no son Barcelona tienen, en media, 1,5 veces más probabilidades de ser eficaces en esta competencia.

Aunque puede resultar tentador asignar estas diferencias a que Barcelona concentra la mayoría de los centros complejos (59,2%), dentro de la provincia el porcentaje de estos centros es menor que en el resto de la comunidad autónoma (24,1% frente a más del 40% en Girona, Lleida y Tarragona). Además, Barcelona concentra también una proporción

mucho mayor de centros de nivel socioeconómico alto (35% en el tercil superior del ISEC, frente al 10,1% en Lleida). No solo el valor añadido, sino que los resultados de los centros complejos en Barcelona son inferiores a los de sus homólogos en otras provincias — especialmente en matemáticas—, mientras que entre los centros no complejos las diferencias territoriales prácticamente desaparecen. En este sentido, la menor presencia de centros resilientes en Barcelona parece vinculada menos a la distribución de la vulnerabilidad y más a la interacción entre nivel socioeconómico, composición lingüística y eficacia escolar en contextos complejos.

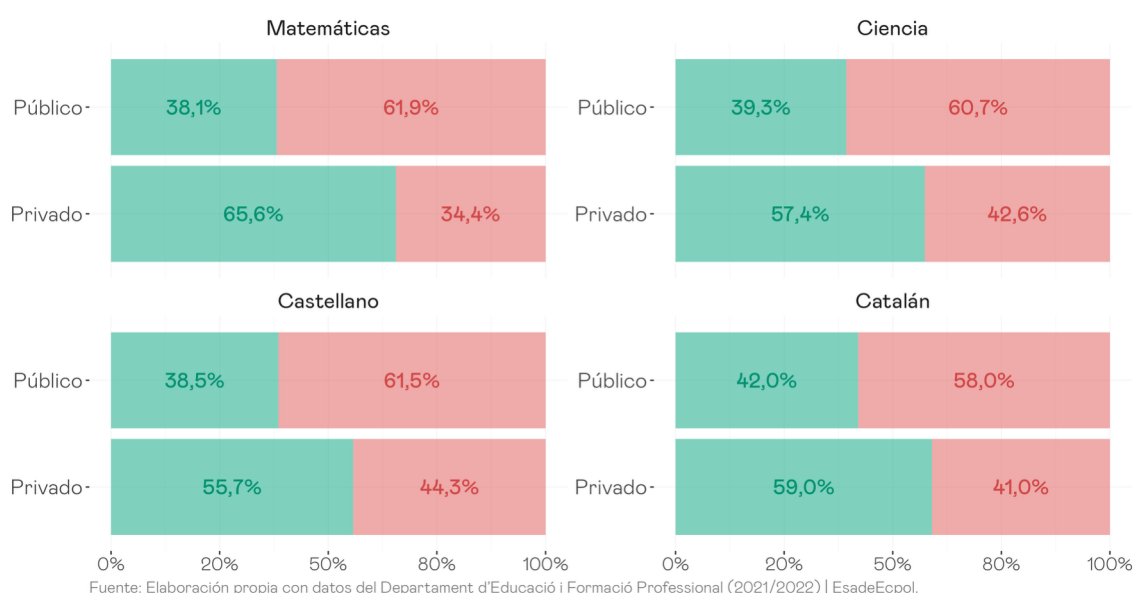
Gráfico 12. Porcentaje de centros con un nivel socioeconómico bajo con VA > 0 (y significativo) y VA < 0 o no significativo, por provincia y competencia



En relación con la titularidad, la red pública concentra la mayoría de los centros resilientes porque es mayoritaria en los contextos vulnerables: ocho de cada diez centros complejos son públicos (Gráfico 13). Ahora bien, cuando se comparan centros de similar nivel socioeconómico, los centros concertados muestran una probabilidad mayor de ser eficaces. Entre los centros situados en el 30 % más vulnerable, un 65,6 % de los concertados son resilientes en matemáticas, frente al 38,1 % de los públicos; en ciencias, castellano y catalán la ventaja es también significativa. De nuevo, el análisis de regresión logística confirma estas disparidades: sin tener en cuenta ningún otro factor, un centro concertado tiene una probabilidad de ser eficaz en matemáticas 23,4 puntos porcentuales más que uno público, en ciencias 17,5 p.p. y en catalán y castellano 16,7 p.p. más. Cabe destacar que la distribución por deciles de complejidad muestra que los centros concertados tienen una presencia algo menor en los contextos de máxima vulnerabilidad: solo un 24 % de los concertados complejos se ubican en el primer decil, frente a un 35 % en el caso de los públicos. Sin embargo, incluso cuando se tiene en cuenta es la regresiones el índice de complejidad de los centros, la ventaja de los centros concertados en la resiliencia permanece prácticamente constante.

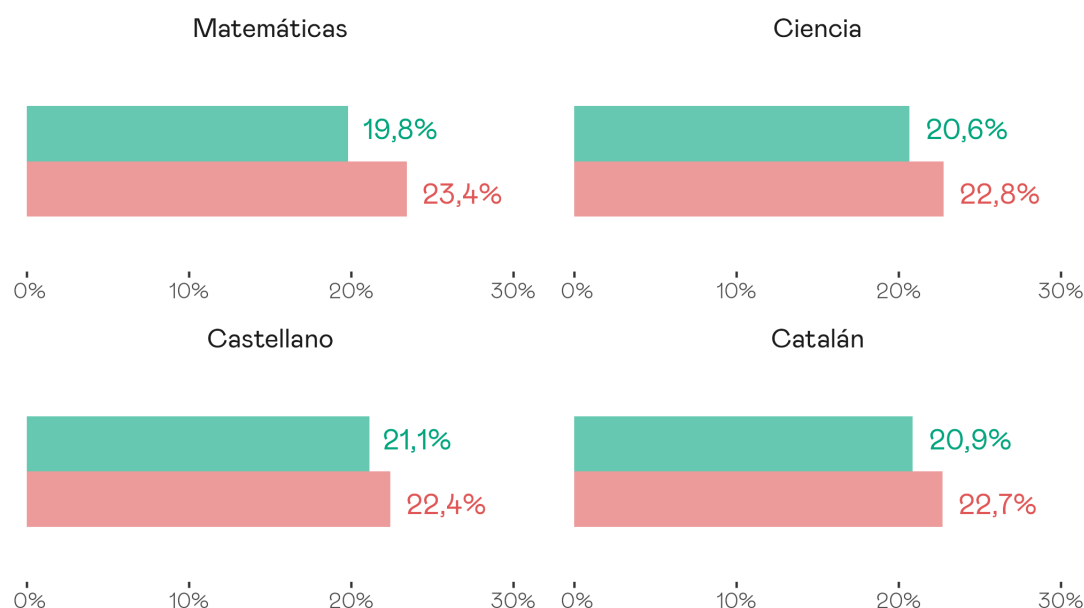
Por otro lado, los centros resilientes tienden a presentar menor rotación del profesorado (20 % frente al 23 % anual) y menor presencia de jornada continua (65 % frente al 80 %). En este sentido, conviene subrayar que las diferencias observadas en eficacia entre redes no se deben necesariamente a las decisiones de la dirección escolar, dado que la autonomía de los centros públicos para seleccionar docentes es muy limitada. Los resultados reflejan tanto características del propio centro como del profesorado que en ese momento forma parte de él. En el caso de la red concertada, que dispone de mayor margen para contratar y retener a sus docentes, la capacidad de configurar equipos más estables y alineados podría estar detrás de parte de su ventaja relativa, a pesar de contar con menos recursos públicos por alumno.

Gráfico 13. Porcentaje de centros con un nivel socioeconómico bajo con $VA > 0$ (y significativo) y $VA < 0$ o no significativo, por titularidad y competencia



En cuanto a la rotación del profesorado, los centros resilientes tienden a contar con claustros más estables (Gráfico 14). La evidencia sobre la rotación docente es consistente: no solo empeora los resultados de los estudiantes expuestos de forma directa a la rotación, sino del resto, ya que la inestabilidad erosiona la organización del centro (Ronfeldt et al., 2013); además, disminuye la eficacia agregada del profesorado, especialmente en contextos vulnerables (Hanushek & Rivkin, 2010) y tiene importantes efectos socioemocionales en el alumnado (Sorensen & Ladd, 2020). En el caso de Secundaria en Catalunya, se confirma que la rotación está muy asociada a la vulnerabilidad, pasando del 21,1% anual en el decil 1 de centros con menor nivel socioeconómico al 6,6% en el decil más alto. Así, no sorprende que, entre los centros más complejos, los resilientes tengan menor rotación: en promedio, la rotación anual de docentes se sitúa en torno al 20 % en estos centros, frente a un 23 % en los no resilientes, una diferencia de 2–3 puntos que resulta estadísticamente significativa en todas las competencias. El patrón es especialmente claro en matemáticas, lo que apunta a que la estabilidad docente puede constituir un factor diferencial de eficacia en esta materia.

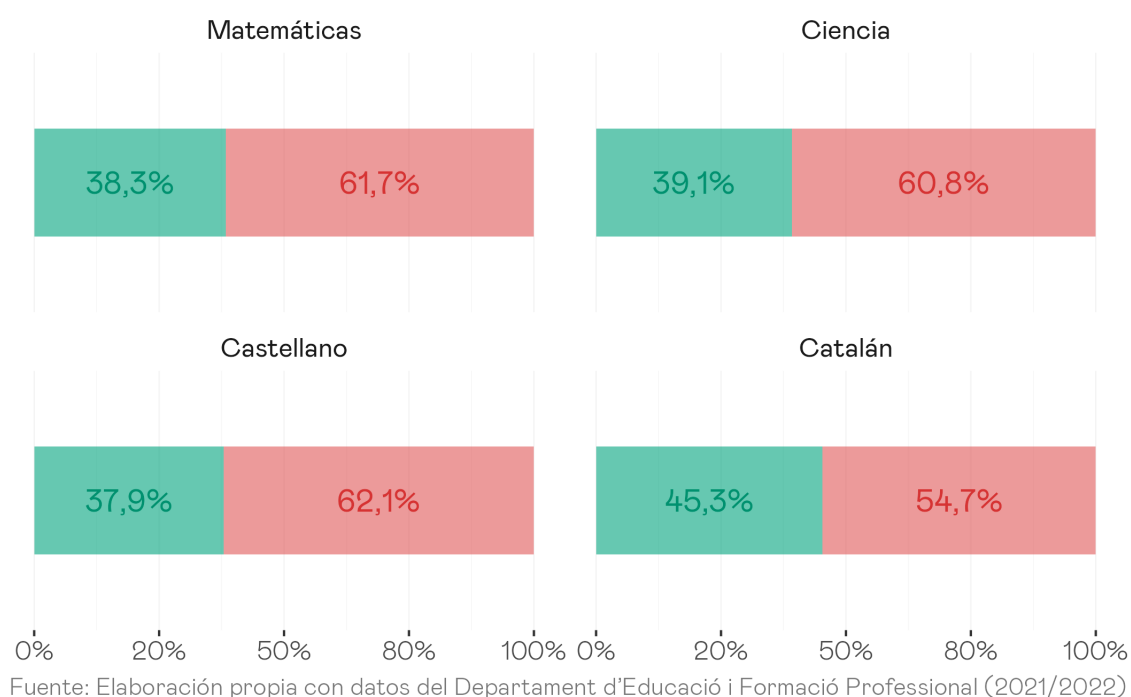
Gráfico 14. Movilidad del profesorado en centros con un nivel socioeconómico bajo en función de su valor añadido: $VA > 0$ (y significativo) y $VA \leq 0$



Fuente: Elaboración propia con datos del Departament d'Educació i Formació Professional (2021/2022) | EsadeEcPol.

El tiempo en la escuela también correlaciona con la resiliencia académica de los centros, teniendo los centros con jornada partida una mayor probabilidad de ser resilientes (Gráfico 15). La evidencia sobre el tipo de jornada (continua frente a partida) es limitada y poco concluyente (Ferrero et al., 2022), aunque en España suele vincularse la partida con un mejor rendimiento (Caride, 1993; Morales-Yago et al., 2017). En cambio, la evidencia respecto al tiempo en la escuela es clara: ampliar la presencialidad con comedor y actividades de tarde mejora el aprendizaje, la salud y el bienestar socioemocional, con especial impacto positivo en alumnado de menor renta (Bellei, 2009; Hincapié, 2016; Padilla-Romo, 2022). De nuevo, el tipo de jornada guarda una fuerte correlación con el nivel de complejidad: los centros complejos tienen 1,85 veces más probabilidad de contar con jornada continua que los no complejos. Así, de nuevo, no sorprende encontrar que, entre los centros con jornada partida, el 57,8% son resilientes en matemáticas (frente al 38,3% de los centros con jornada continua). La dinámica es la misma en todas las competencias, aunque el porcentaje de centros resilientes entre los que tienen jornada partida ronda el 53 %. Visto de otra manera, en el caso de matemáticas, los centros con jornada continua presentan una probabilidad casi 20 puntos porcentuales menor de que su alumnado sea resiliente en matemáticas en comparación con los centros de jornada partida (p -valor=0,002).

Gráfico 15. Porcentaje de centros con jornada continua en función de su valor añadido VA > 0 (y significativo) y VA ≤ 0



Estos tres factores (titularidad – movilidad del profesorado – jornada escolar) no son independientes: la combinación de menor movilidad docente y de menor presencia de jornada continua contribuye a explicar la mayor probabilidad de resiliencia en la red concertada. Aunque los concertados son minoritarios en los contextos de mayor complejidad, muestran un perfil organizativo más favorable: rotación media del 7 % frente al 24 % en la pública y un uso residual de la jornada continua frente a cerca del 90 % en la pública. Estos factores organizativos, asociados de forma consistente con mejores resultados de progreso, pueden ayudar a entender por qué los concertados tienen más probabilidad de ser resilientes en entornos vulnerables. De hecho, en análisis de regresión logística, una vez se controla por la titularidad del centro, la rotación docente y la movilidad del profesorado pierden su relevancia estadística.

Las estimaciones logísticas multivariantes confirman la robustez de la relación entre resiliencia y red concertada en Secundaria en Catalunya en el caso de competencia matemática, pero no en el resto de las competencias. Mientras que, en matemáticas, cuando se tiene en cuenta la provincia, movilidad docente, tipo de jornada y nivel socioeconómico del centro los centros concertados mantienen una probabilidad de 22,9 puntos mayor de ser resilientes que los públicos, en el resto de las competencias esta probabilidad deja de ser estadísticamente significativa, aunque en catalán y castellano mantiene un coeficiente elevado. Cuando se analiza directamente la relación entre titularidad y valor añadido en matemáticas, el coeficiente asociado a la concertada es de 0,13 d.e., un tamaño relevante.

En cualquier caso, estos resultados deben interpretarse teniendo en cuenta que existen factores inobservables que pueden afectar de forma diferencial a los centros públicos y

concertados relacionados con la autonomía para seleccionar y retener docentes o con dinámicas específicas de los equipos presentes en cada centro en un momento dado. En este sentido, este efecto podría estar reflejando la capacidad de configurar claustros más estables y cohesionados.

Tabla 5. Efecto marginal de asistir a un centro concertado en la resiliencia por competencia (efectos marginales de modelos Logit)

	Efecto marginal	Error estándar	p-valor
Matemáticas	0,230	0,121	0,058
Ciencias	0,161	0,125	0,196
Castellano	0,066	0,126	0,602
Catalán	0,158	0,126	0,212

Notas: Resultados de estimación de modelos logit multivariantes controlando, además de titularidad del centro, provincia, índice de nivel socioeconómico, tipo de jornada y movilidad de los profesores.

Cabe destacar que estas diferencias observadas se deben fundamentalmente a centros con un valor añadido elevado. De hecho, cuando el análisis se ciñe a estos centros (con un valor añadido superior a 0,2 desviaciones estándar) se encuentran diferencias relevantes en el bienestar físico y emocional del alumnado²⁴ respecto a los centros con un valor añadido nulo o negativo en una determinada competencia —aunque la muestra, y por ende el poder estadístico, es menor—. La diferencia más relevante se da en el porcentaje de alumnos que reportan pasar mucho tiempo ante dispositivos digitales, con un 57 % en los centros con valor añadido nulo o negativo en Matemáticas frente al 50 % en los centros con valor añadido positivo y alto, siendo la diferencia ligeramente menor en el resto de las competencias.

Por último, de cara a seleccionar los centros para el trabajo de campo, se impone la condición de que el centro sea resiliente en todas competencias, de forma que la muestra de centros se reduce a 44 (Tabla 6). Se aprecia que los 44 centros resilientes presentan, en promedio, un valor añadido mayor en las cuatro competencias, tal como cabría esperar al restringir la definición a un rendimiento superior en todos los frentes. Además, cuando se hace la comparación todavía más estricta (por ejemplo, enfrentando estos 44 centros solo contra los 86 que no tienen VA positivo en ninguna de las competencias), la diferencia en rendimiento es todavía más marcada.

La diferencia estadísticamente significativa más notable surge de nuevo en la titularidad: el 10,9 % de los centros de nivel socioeconómico bajo públicos son resilientes en todas las competencias frente al 24,6 % de los concertados: es decir, la probabilidad de que un centro concertado complejo sea resiliente en todas las competencias es 2,8 veces mayor de la que lo sea un público. También se acrecientan las diferencias provinciales: la

²⁴ Hay que tener en cuenta que hay un 29% de centros complejos (entre el 30% de menor nivel socioeconómico según el índice) que no presentan información de bienestar por no haber contestado a los cuestionarios o no tener un número suficiente de respuestas.

probabilidad de que un centro de Barcelona sea resiliente en todas las materias es del 10,7%, mientras que, en el caso del resto de provincias, de media, esta probabilidad asciende al 18,5%.

Tabla 6. Número (%) de centros complejos en función de su Valor Añadido

	Matemáticas	Ciencias	Castellano	Catalán	Todas
VA<0	143 (45%)	125 (39,3%)	129 (40,6%)	130 (40,9%)	274 (13,8%)
VA=0	37 (11,6%)	57 (17,9%)	56 (17,6%)	44 (13,8%)	
VA>0	138 (43,4%)	136 (42,8%)	133 (41,8%)	144 (45,3%)	44 (19,9%)
	318	318	318	318	318

5. Referencias

- Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society: Series A (General)*, 149(1), 1–43. <https://doi.org/10.2307/2981882>
- American Statistical Association (2014). *ASA statement on using value-added models for educational assessment*.
- Amrein-Beardsley, A. (2014). *Rethinking value-added models in education: Critical perspectives on tests and assessment-based accountability*. Routledge.
- Angrist, J., Hull, P., Pathak, P. A., & Walters, C. (2021). Credible school value-added with undersubscribed school lotteries. *The Review of Economics and Statistics*, 1-46.
- Åslund, O., Edin, P. A., Fredriksson, P., & Grönqvist, H. (2011). Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy. *American Economic Journal: Applied Economics*, 3(2), 67-95.
- Australian Curriculum, Assessment and Reporting Authority. (2021). *My school*. <https://www.myschool.edu.au/>
- Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37–65. <https://doi.org/10.3102/10769986029001037>
- Bellei, C., Muñoz, G., Pérez, L. M., & Raczyński, D. (2004). Escuelas efectivas en sectores de pobreza? Quién dijo que no se puede
- Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review*, 28(5), 629-640.
- Beuermann, D. W., Jackson, C. K., Navarro-Sola, L., & 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. <https://doi.org/10.1093/restud/rdac025>
- Braun, H. I., & Wainer, H. (2006). Value-added modelling. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics* (pp. 867–892). [https://doi.org/10.1016/S0169-7161\(06\)26027-9](https://doi.org/10.1016/S0169-7161(06)26027-9)
- Britton, J., Clark, D., Lee, I. (2023). *Unveiling school effectiveness: Progress 8, parental choices and closing the achievement gap*. Published by the Institute for Fiscal Studies.
- Calero, J., Choi, Á., & Waisgrais, S. (2010). *Determinantes del riesgo de fracaso escolar en España: una aproximación a través de un análisis logístico multinivel aplicado a PISA-2006*. <https://hdl.handle.net/20.500.12799/1203>
- Castellano, K. E., & Ho, A. D. (2013). *A practitioner's guide to growth models*. Council of Chief State School Officers.

- Carneiro, P., Meghir, C., & Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, 11, 123-160. <https://doi.org/10.1111/j.1542-4774.2012.01096.x>
- Caride, J. (1993). *A xornada escolar de sesión única en Galicia. Estudio avaliativo: Conclusións xerais e criterios de actuación*. Santiago de Compostela, Xunta de Galicia.
- Castellano, K. E., Rabe-Hesketh, S., & Skrondal, A. (2014). Composition, context, and endogeneity in school and teacher comparisons. *Journal of Educational and Behavioral Statistics*, 39(5), 333–367. <https://doi.org/10.3102/1076998614547576>
- Chetty, R., Hendren, N., Kline, P., & 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. <https://doi.org/10.1093/qje/qju022>
- Chiswick, B. R., & DebBurman, N. (2004). Educational attainment: analysis by immigrant generation. *Economics of Education Review*, 23(4), 361-379. <https://doi.org/10.1016/j.econedurev.2003.09.002>
- Cobrerros, L., & Gortázar, L. (2023). *Todo lo que debes saber de PISA 2022 sobre equidad*. EsadeEcPol (ESADE Center for Economic Policy) y Save the Children. Recuperado de: <https://www.esade.edu/ecpol/es/publicaciones/todo-lo-que-debes-saber-de-pisa-2022-sobre-equidad>.
- Coleman, J. S. (1968). Equality of educational opportunity. *Integrated education*, 6(5), 19-28.
- Deming, D. J. (2014). Using school choice lotteries to test measures of school effectiveness. *American Economic Review*, 104(5), 406-11.
- Department for Education. (2020). *Secondary accountability measures: Guide for maintained secondary schools; academies, and free schools*.
- Dustmann, C., Frattini, T., & Lanzara, G. (2012). Educational achievement of second-generation immigrants: an international comparison. *Economic Policy*, 27(69), 143-185 <https://doi.org/10.1111/j.1468-0327.2011.00275.x>
- Ferrero, M., Gortazar, L., & Martínez-Jorge, A. (2022). *Jornada escolar continua: cómo la pandemia está acelerando un modelo social y educativo regresivo*. EsadeEcPol Brief #24, Mayo de 2022.
- Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement*, 8(4), 369–395. <https://doi.org/10.1080/0924345970080401>
- Grimaldo-Moreno, F. & López-Iñesta, E. (2021). *Tipologia dels centres educatius segons el nivell de complexitat*. Generalitat de Catalunya, Consell Superior d'Avaluació del Sistema Educatiu.
- Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge.

Hincapié, D. (2016). Do longer school days improve student achievement? Evidence from Colombia (No. IDB-WP-679). *IDB Working paper series*.

Jackson, C. K., Porter, S. C., Easton, J. Q., Blanchard, A., & Kiguel, S. (2020). School effects on socioemotional development, school-based arrests, and educational attainment. *American Economic Review: Insights*, 2 (4), 491-508. DOI: [10.1257/aeri.20200029](https://doi.org/10.1257/aeri.20200029)

Keeves, J. P., Hungi, N., & Afrassa, T. (2005). Measuring value added effects across schools: Should schools be compared in performance? *Studies in Educational Evaluation*, 31(2-3), 247-266. <https://doi.org/10.1016/j.stueduc.2005.05.012>

Koretz, D. (2017). *The testing charade: Pretending to make schools better*. University of Chicago Press.

Lamb, S., Jackson, J., Walstab, A., & Huo, S. (2015). *Educational opportunity in Australia 2015: Who succeeds and who misses out*. Centre for International Research on Education Systems, Victoria University. <https://www.vu.edu.au/centre-for-international-research-on-education-systems-cires/research-projects/educational-opportunity-in-australia>

Leckie, G., & Goldstein, H. (2017). The evolution of school league tables in England 1992-2016: "Contextual value-added", "expected progress" and "progress 8". *British Educational Research Journal*, 43(2), 193-212. <https://doi.org/10.1002/berj.3264>

Leckie, G., & Goldstein, H. (2019). The importance of adjusting for pupil background in school value added models: A study of Progress 8 and school accountability in England. *British Educational Research Journal*, 45(3), 518-537. <https://doi.org/10.1002/berj.3511>

Leckie, G., & Prior, L. (2022). *A comparison of Value-Added Models for School Accountability*. Centre for Multilevel Modelling and School of Education, University of Bristol. DOI: [10.1257/000282806776157641](https://doi.org/10.1257/000282806776157641)

Levy, J., Brunner, M., Keller, U., & Fischbach, A. (2023). How sensitive are the evaluations of a school's effectiveness to the selection of covariates in the applied value-added model?. *Educational Assessment, Evaluation and Accountability*, 35(1), 129-164.

Lizasoain, L., Azpillaga V., & Joaristi, L. M. (2014). Detección y caracterización de los centros escolares de alta eficacia de la Comunidad Autónoma del País Vasco mediante Modelos

Transversales Contextualizados y Modelos Jerárquicos Lineales. Estudios sobre educación

Lizasoain, H.L., & Angulo, V.A., (2014). *Buenas prácticas de escuelas eficaces del país vasco: metodología y primeros resultados*. Participación educativa.

Lizasoain, L., Angulo, A., Azpillaga, V., Bartau, I. Damborena, M. D., Del-Frago, R., Etxeberria, F., Intxausti, N., Joaristi, L., Méndez, Y., Núñez, C., Valadez, C. (2015). *La eficacia escolar en los centros del País Vasco*. Informe final 2011-2015. ISEI-IVEI.

- Loeb, S., Kalogrides, D., & Bételle, T. (2012). Effective schools: Teacher hiring, assignment, development, and retention. *Education Finance and Policy*, 7(3), 269-304.
- Loviglio, A. (2023). School Quality Beyond Test Scores: The Role of Schools in Shaping Educational Outcomes. *IZA Discussion Paper*. Retrieved from: <https://www.iza.org/publications/dp/16111/school-quality-beyond-test-scores-the-role-of-schools-in-shaping-educational-outcomes>
- Marks, G. N. (2017). Is adjusting for prior achievement sufficient for school effectiveness studies? *Educational Research and Evaluation*, 23 (5-6), 148-162. <https://doi.org/10.1080/13803611.2017.1455287>
- Marks, G. N. (2021). Should value-added school effects models include student- and school-level covariates? Evidence from Australian population assessment data. *British Educational Research Journal*, 47(1), 181-204. <https://doi.org/10.1002/berj.3684>
- Mbekeani, P. P., Papay, J., Mantil, A., & Murnane, R. J. (2023). Understanding High Schools' Effects on Longer-Term Outcomes. *EdWorkingPaper* No. 23 – 723. Annenberg Brown University. Retrieved from: <https://edworkingpapers.com/sites/default/files/ai23-729.pdf>
- Meunier, M. (2011). Immigration and student achievement: Evidence from Switzerland. *Economics of education review*, 30(1), 16-38. <https://doi.org/10.1016/j.econedurev.2010.06.017>
- Morales Yago, F. J., Galán, A., & Pérez, R. (2017). Jornada escolar partida y continua. ¿Existen evidencias que motiven el cambio en la gestión del tiempo escolar en España? *Revista Complutense de Educación*, 28, 965-984.
- Muñoz-Chereau, B., Anwandter, A., & Thomas, S. (2020). Value-added indicators for a fairer Chilean school accountability system: A pending subject. *Journal of Education Policy*, 35(5), 602-622. <https://doi.org/10.1080/02680939.2019.1598584>
- OECD (2016). *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*. Paris: OECD Publishing.
- Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review*, 96(1), 152-175.
- Pinskaya, M., Khavenson, T., Kosaretsky, S., Zvyaginsev, R., Mikhailova, A., & Chirkina, T. (2018). Above barriers: A survey of resilient schools. *Educational Studies Moscow*, 2, 198-227.
- Pinskaya, M., Kosaretsky, S., Zvyagintsev, R., & Derbishire, N. (2019). Building resilient schools in Russia: Effective policy strategies. *School Leadership & Management*, 39(2), 127-144.
- Prior, L., Jerrim J., Thomson, D., & Leckie, G. (2021). A review and evaluation of secondary school accountability in England: Statistical strengths, weaknesses and

challenges for "Progress 8" raised by COVID-19. *Review of Education*, 9(3), Article e3299. <https://doi.org/10.1002/rev3.3299>

Raudenbush, S. W., & Willms, J. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307–335. <https://doi.org/10.3102/10769986020004307>

Reardon, S. F., & Raudenbush, S. W. (2009). Assumptions of value-added models for estimating school effects. *Education Finance and Policy*, 4(4), 492–519

Reynolds, D., Sammons, P., Stoll, L., Barber, M., & Hillman, J. (1996). School effectiveness and school improvement in the United Kingdom. *School effectiveness and school improvement*, 7(2), 133–158.

Saunders, L. (1999). A brief history of educational "value added": How did we get to where we are? *School Effectiveness and School Improvement*, 10 (2), 233–256. <https://doi.org/10.1076/sesi.10.2.233.3507>

Scheerens, J. (2004). *Review of school and instructional effectiveness research*

Thomas, S., & Mortimore, P. (1996). Comparison of value-added models for secondary-school effectiveness. *Research Papers in Education*, 11(1), 5–33. <https://doi.org/10.1080/0267152960110103>

Thomas, S. (2001). Dimensions of secondary school effectiveness: Comparative analyses across regions. *School Effectiveness and School Improvement*, 12(3), 285–322. <https://doi.org/10.1076/sesi.12.3.285.3448>

Timmermans, A. C., Doolaard, S., & de Wolf, I. (2011). Conceptual and empirical differences among various value-added models for accountability. *School Effectiveness and School Improvement*, 22 (4), 393–413. <https://doi.org/10.1080/09243453.2011.590704>

Timmermans, A. C., & Thomas, S. M. (2015). The impact of student composition on schools' value added performance: A comparison of seven empirical studies. *School Effectiveness and School Improvement*, 26(3), 487–498. <https://doi.org/10.1080/09243453.2014.957328>

Wainer, H. (2004). Introduction to a special issue of the Journal of Educational and Behavioral

Statistics on value-added assessment. *Journal of Educational and Behavioral Statistics*, 29(1), 1– 2. <https://doi.org/10.3102/10769986029001001>