

El uso de los modelos nacionales de valor añadido para la mejora de las escuelas británicas

The use of national value-added models for school improvement in English schools

Andrew Ray, Helen Evans, Tanya McCormack

Department for Children, Schools and Families, London, United Kingdom

Resumen

La mejora de la escuela es central en la política educativa en Inglaterra. Se anima a las escuelas a desarrollarse y mejorar dentro de un marco de autoevaluación, inspección, consejo y asesoramiento en el nivel local y nacional, y a tomar parte en distintas iniciativas y políticas. La recogida y uso de datos es crucial en este proceso. El análisis de datos tiene lugar en todos los niveles del sistema, desde la evaluación de políticas nacionales, hasta el nivel de alumnos individuales, donde las escuelas usan los datos para seguir el progreso académico y la evaluación de los aprendizajes. En las escuelas británicas se han desarrollado un abanico de iniciativas para tratar de elevar los estándares de rendimiento. Una pieza fundamental ha sido el desarrollo de pruebas que pueden vincularse a los logros previos de los estudiantes y usarse para generar estimaciones de la contribución de la escuela al progreso de sus alumnos: indicadores de valor añadido. En Inglaterra, la disponibilidad de datos nacionales de las características de los estudiantes individuales desde 2002 ha permitido el desarrollo de puntuaciones «contextuales de valor añadido» para escuelas y grupos de alumnos dentro de cada escuela. Estas características pueden analizarse con un paquete estadístico llamado RAISEonline, desarrollado conjuntamente por el gobierno central y la agencia independiente de inspección de las escuelas. Este artículo describe estos desarrollos y aporta ejemplos del tipo de análisis que RAISEonline puede realizar u ofrecer.

Palabras clave: modelos de valor añadido, sistema educativo inglés, sistemas de evaluación, sistemas de rendición de cuentas, modelos de crecimiento.

Abstract

School improvement is central to education policy in England. Schools are encouraged to develop and improve within a framework of self-evaluation, inspection, advice and guidance at local and national level, and by taking part in various initiatives and policies. The collection and use of data is crucial in this process. Analysis of data takes place at all levels of the system, ranging from the evaluation of national policies, down to the level of individual pupils, where schools use data in tracking progress and assessment for learning. A range of school improvement initiatives have been developed to try to raise standards in English schools. A crucial building block has been the development of test data which can be matched to pupils' prior attainment and used to generate estimates of schools' contribution to the progress of their pupils: value-added indicators. In England, the availability of national data on the characteristics of individual pupils since 2002 has allowed the development of «contextual value-added» scores for schools and groups of pupils within schools. These statistics can be analysed in a software package called RAISEonline, developed jointly by central government and the independent school inspection agency. The paper describes these developments and provides examples of the kind of analysis RAISEonline can now provide.

Key Words: Value-added models, English educational system, assessment system, accountability system, growth models.

Introducción

La mejora de la escuela es central en la política educativa en Inglaterra. Se anima a las escuelas a desarrollarse y mejorar dentro de un marco de autoevaluación, inspección, consejo y asesoramiento en el nivel local y nacional, y a tomar parte en distintas iniciativas y políticas. La recogida y uso de datos es crucial en este proceso. El análisis de datos tiene lugar en todos los niveles del sistema, desde la evaluación de políticas nacionales, hasta el nivel de alumnos individuales, donde las escuelas usan los datos para seguir el progreso académico y la evaluación de los aprendizajes.

En Inglaterra hay aproximadamente 8,1 millones de estudiantes en 25.000 escuelas de financiación pública e independientes (DfES 2007) -el 7% de los alumnos están en la enseñanza independiente-. Algunos estudiantes con necesidades educativas especiales están escolarizados en escuelas regulares, otros están escolarizados en escuelas específicas de forma separada. Hay aproximadamente 17.400 escuelas primarias, que generalmente acogen alumnos desde los 4 a los 11 años, y alrededor de 3.300

escuelas secundarias, que normalmente cubren las edades de 11 a 16 años (1.750 tienen «seis formas» que cubren también más allá de los 16 años). El tamaño medio de una escuela secundaria es de 980 estudiantes (aproximadamente 140 estudiantes por cohorte anual en términos promedios); las escuelas primarias tienen en promedio 240 alumnos, 40 por año.

El mantenimiento de las escuelas se realiza a través del gobierno local: hay 150 autoridades locales en toda Inglaterra. Las autoridades locales varían considerablemente en tamaño y características. La menor es la de Scilly Isles, con únicamente una escuela, y la mayor es Kent con 103 escuelas secundarias y 466 escuelas primarias. Los datos educativos son usados extensivamente en el nivel de la autoridad local y en el más amplio nivel regional (por ejemplo, para observar el progreso de los alumnos de Londres). Los análisis son realizados por redes o grupos geográficos de escuelas, por las propias escuelas y dentro de las escuelas por áreas temáticas, profesores y por grupos de estudiantes con características específicas. Este artículo estudia algunos de los análisis desarrollados centralmente con datos recogidos de los alumnos individuales y de las escuelas (no hay una recogida de datos nacional que vincule a los profesores con logros de los alumnos). Se centra en los recientes desarrollos en el uso de modelos de valor añadido y en indicadores para la mejora escolar.

El punto de partida para cualquier análisis de mejora de las escuelas es el desarrollo de pruebas fiables y consistentes o datos de evaluación. El actual sistema en Inglaterra para alumnos de hasta 16 años implica evaluaciones al final de las cuatro etapas clave. Estas etapas están definidas como parte del Currículo Nacional que, introducido en 1988, establece las materias y los programas de estudio que las escuelas públicas están obligadas a cubrir. Las etapas claves 1, 2, 3 y 4 cubren los rangos de edad de 5 a 7 años, de 7 a 11 años, de 11 a 14 años y de 14 a 16 años respectivamente. Había pruebas elaboradas para los alumnos de 16 años (GCSEs, y previamente a éstas, «Niveles O» y los exámenes CSE), pero el Currículo Nacional lideró el desarrollo de un sistema nacional de medición y de evaluación del profesorado al final de cada etapa clave. Esto significa que actualmente no hay resultados de pruebas anuales de cada cohorte, en su lugar las pruebas miden el rendimiento en periodos que oscilan entre los 2 y los 4 años.

Los logros académicos en las etapas de 1 a 3 están evaluados con niveles nacionales de currículum referidos a criterio. El sistema de medida depende de calificaciones independientes y del Curriculum Authority (QCA) y de la Agencia Nacional de Evaluación (NAA). Los niveles del test están equiparados entre años utilizando anclajes y pruebas de pre-test. Ellos han diseñado lo que en el nivel 4 de la etapa clave 2 es

equivalente al nivel 4 de la etapa clave 3, sin embargo los modelos de valor añadido no dependen de una equiparación vertical o media del crecimiento en una escala común. En la etapa clave 4 hay un rango de calificaciones que pueden compararse utilizando una escala común, pero esta escala no está relacionada con los niveles de las etapas previas. Más información sobre las pruebas puede encontrarse en Ray (2006) y en la página web de QCA.

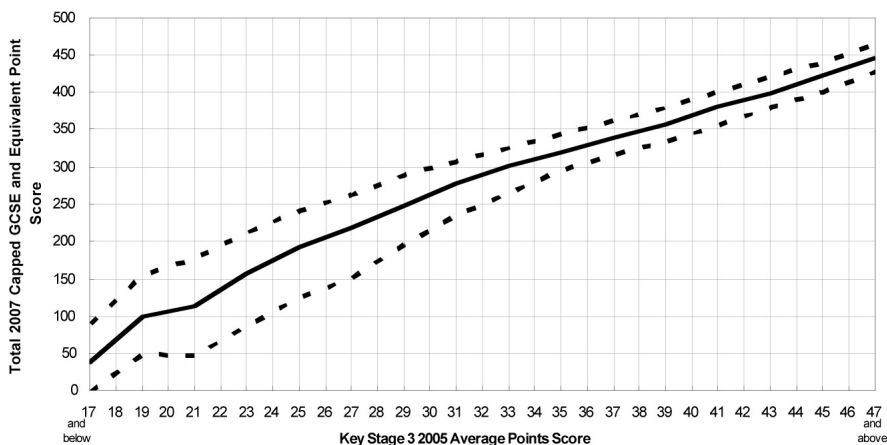
Es posible el uso de datos del nivel de escuelas, por ejemplo, en la proporción de alumnos que alcanzan un determinado estándar, en los que consideran conjuntamente patrones nacionales y para conocer las diferencias locales entre escuelas. Sin embargo, las bases para los análisis en los últimos años han utilizado datos de tests de los alumnos individuales, donde sus resultados en una determinada etapa clave están vinculados al rendimiento previo en el año anterior. Una vez que se establecieron tests comparables al nivel nacional y el sistema estableció cómo recoger los resultados, pudo ser posible crear bases de datos vinculadas, donde el progreso de un alumno individual o grupo de estudiantes pudo ser medido.

El desarrollo de los análisis para la mejora de la escuela

Previo a la recolección nacional de datos de tests vinculados, para grupos de escuelas se emplearon análisis de valor añadido para la mejora de las escuelas si sus alumnos realizaban tests específicos, tales como en el programa desarrollado por el centro CEM de la Universidad de Durham (Tymms y Coe, 2003) y en estudios como el de Mortimore et al. (1988) y el de Goldstein et al. (1993). Se puede encontrar un resumen de esos estudios previos en SCAA (School Curriculum and Assessment Authority) (1994). Sin embargo, ninguna de las escuelas tuvo acceso o hizo uso de este tipo de datos. Para solventar este problema, el primer Autumn Package fue realizado en 1998, con patrones nacionales de valor añadido y con patrones estadísticos para grupos de escuelas. Esto permitió a las escuelas comparar sus rendimientos y establecer objetivos¹. Se desarrolló una versión interactiva del programa, donde las escuelas podían introducir sus datos y ver el abanico de análisis útiles de los alumnos o de los grupos de alumnos.

¹ Véase DEPARTMENT FOR CHILDREN SCHOOLS AND FAMILIES, DfEE (1998). La información limitada de valor añadido estuvo disponible un año antes (DfEE/QCA, 1998);] antes no era posible vincular los niveles de rendimiento previo de los alumnos con los resultados en las etapas clave de las escuelas.

GRÁFICO I. Ejemplo de gráfico de valor añadido realizada por el Autumn Package



Las tablas y gráficos nacionales del Autumn Package, de los que el Gráfico I es un ejemplo, no requerían una cobertura completa y una precisión en los vínculos. Sin embargo, para aportar datos de valor añadido para las escuelas individuales, fue importante usar correctamente los datos vinculados para todos los alumnos. En 1997, el gobierno estableció el desarrollo de mejores datos del nivel de alumnos y buscó la introducción de un identificador único para cada alumno que podría ayudar a que los datos fueran seguidos a través del sistema educativo. Estos identificadores únicos del alumno fueron introducidos en 1999, siguiendo con las consideraciones prácticas sobre protección de datos. Sin embargo, los datos vinculados fueron completándose de forma creciente, este desarrollo facilitó la centralización del cálculo y aportó indicadores simples de valor añadido para todas las escuelas primarias y secundarias públicas. Estos indicadores fueron desarrollados consultando e introducidos en las tablas anuales de rendimiento comprendidas entre 2002 y 2004. También fueron usadas por los servicios de inspección escolar, Ofsted, y por organizaciones externas encargadas de crear estándares para las escuelas a través de las Estrategias Nacionales de Primaria y Secundaria. Para una exposición más detallada de esta fase inicial de los indicadores de valor añadido se puede consultar más información en Ray (2006).

La nueva relación con las escuelas

Varios factores llevaron a la siguiente fase de desarrollo de información de valor añadido centralizada. En 2002 se introdujo el Pupil Level Annual School Census (PLASC), ampliando el rango de los datos disponibles sobre las características de los estudiantes de forma centralizada desde considerar sólo género a incluir información útil sobre raza, condiciones sociales, privación, necesidades educativas especiales y primera lengua. Estos nuevos datos eran útiles por sí mismos, pero también ofrecieron la posibilidad de desarrollar modelos de valor añadido que tenían en cuenta esos factores «contextuales». Se estableció un programa de consulta y desarrollo para concebir un nuevo conjunto de «modelos de valor añadido contextuales» que pudieran ser utilizados para aportar información adicional sobre la eficacia de las escuelas. Éstos fueron parte central del New Relationship with Schools lanzado en 2004.

New Relationship with Schools también introdujo School Improvement Partners (SIP's), expertos (a menudo ex directores de centros) quienes podrían trabajar con las escuelas en su autoevaluación. La aproximación tomada por la inspección escolar de Ofsted también varió y se pidió a los inspectores que consideraran un mayor rango de datos de rendimiento, incluyendo el valor añadido. Tuvieron inicialmente que depender de informes escritos (PANDAs) provenientes de un conjunto de estadísticas claves de cada escuela. Ahora, la finalidad era desarrollar PANDAs y el Autumn Package en un nuevo software que sería pilotado con los datos y gráficos de cada escuela y que mostraría un abanico de estadísticas, incluyendo el nuevo valor añadido contextualizado. Este producto, RAISEonline² fue introducido en 2006 y se presenta a continuación.

Los modelos de valor añadido contextual

Una vez que los datos de PLASC se vincularon con la información del progreso del alumno, se pusieron a disposición de la sociedad, hacia finales del año 2002. Estadísticos en el Ministerio de Educación comenzaron a analizarlos para entender las relaciones entre las variables y qué podrían decir sobre el rendimiento nacional. Los planteamientos

² Puede verse en <https://www.raiseonline.org/>

de una selección de académicos fueron consultados para establecer las futuras direcciones de los trabajos sobre valor añadido y, aunque no se encontró consenso entre sus opiniones, hubo un fuerte apoyo para el desarrollo de modelos más complejos que los utilizados, con los nuevos datos. Fuera del Ministerio de Educación, los estadísticos habían construido modelos de valor añadido que consideraban factores contextuales, por ejemplo el Fischer Family Trust. Asesorados por la Oficina Nacional de Auditorías (2003) se desarrolló información de rendimiento para escuelas que tenían en cuenta no sólo el rendimiento previo, sino también «otras influencias externas del rendimiento», basadas en los datos de PLASC.

En octubre de 2004 se presentó un prototipo contextualizado de modelo de valor añadido (CVA) a las escuelas. Al año siguiente un sistema de puntuaciones CVA fue pilotado para usarlo en las tablas de rendimiento y utilizado por primera vez para la mejora de las escuelas en el año siguiente. El modelo CVA 2005 para escuelas secundarias cubría el progreso de los estudiantes desde el final de la etapa clave 2 hasta la etapa clave 4, y se encuentra desarrollado con detalle en Ray (2006). El modelo equivalente para 2007 se presenta en la Tabla I.

El modelo de regresión de valor añadido más simple (1), sin factores contextuales, incluiría $y_{ij(t_2)}$ que son los resultados para la etapa clave 4, del alumno i en la escuela j , y como variable explicativa $y_{ij(t_1)}$, que es el rendimiento previo medio en la etapa clave 2 calculado sobre los cinco años anteriores. Los coeficientes de regresión son a y b ; e_{ij} y u_j son los valores aleatorios, que son independientes, residuos normalmente distribuidos con media cero [0] y varianza constante. El modelo de coeficientes puede estimarse utilizando los estimadores Mínimos Cuadrados Ordinarios (OLS) y la media de los residuos de los alumnos para cada escuela se utilizaría como medida del valor añadido.

$$y_{ij(t_2)} = a + by_{ij(t_1)} + u_j + e_{ij} \quad (1)$$

En el modelo de valor añadido contextualizado para escuelas sostenidas con fondos públicos se emplean actualmente cuatro variables para crear modelos del modelo de rendimiento previo (que se describe a continuación), un total de diez variables contextuales del nivel del alumno y dos del nivel de las escuelas. Como se ve en (2), por ejemplo, x_{5ij} representa el primer factor contextual del alumno y y_{4j} es el primero de los dos predictores del nivel de escuela. (Nótese que las variables explicativas están medidas en 2007 y pueden también ser asignadas al subíndice de la etapa clave 4, aunque la mayoría serían igualmente aplicables en momentos previos, como por ejemplo: género, grupo racial).

$$y_{ij(t_2)} = \beta_0 + \beta_1 y_{1ij(t_1)} + \dots + \beta_4 y_{4ij(t_1)} + \beta_5 x_{5ij} + \dots + \beta_p x_{pij} + \beta_q y_{qj(t_1)} + \beta_s y_{sj(t_1)} + \varepsilon_{ij} \quad (2)$$

Más que usar la aproximación de Mínimos Cuadrados Ordinarios, los indicadores CVA se derivan de un modelo multinivel (Goldstein, 2003), obtenidos con MLwiN, un paquete estadístico que estima los efectos fijos –rendimiento previo y factores contextuales– y calcula los residuos de nivel 1 y 2³. Los residuos de nivel 1 muestran la variación de los resultados de los alumnos en relación con sus escuelas. Los residuos de nivel 2 muestran los resultados de las escuelas en relación con los resultados nacionales esperados, dados los factores medidos por los efectos fijos. Estos residuos del nivel 2 son las puntuaciones de valor añadido.

Factores controlados por el modelo

La variable dependiente «modelizada» es la puntuación total obtenida en la etapa clave 4, basada en las puntuaciones de cada alumno en sus mejores ocho calificaciones. Esto aporta una medida continua del logro desde el alumno más capaz, calificado con ocho A* o más hasta los alumnos que rinden menos, calificados con las menores puntuaciones (G) e incluye calificaciones de bajo nivel (tales como en alfabetización básica y numeración) tomadas principalmente de alumnos con necesidades educativas especiales.

Las medidas de los resultados fueron elegidas más allá de un simple límite (como el que se establecería si un alumno ha obtenido cinco o más A* -C), así el modelo resultante refleja bajo –o sobre– rendimiento de los alumnos en todas las partes de la escala más que centrarse sólo en los estudiantes situados en el límite marcado por las puntuaciones C/D. También aporta un útil diagnóstico en el nivel de alumnos, mostrando lo que cada estudiante ha logrado en comparación con alumnos «similares» a nivel nacional. Además de establecer la puntuación hasta las ocho mejores ocho busca capturar tanto la cantidad como la calidad de las puntuaciones en las asignaturas, sin aportar incentivos a las escuelas para admitir alumnos con un excesivo número de calificaciones.

La variable explicativa más importante es el rendimiento previo de cada estudiante y este ha sido «modelizado» de forma tan precisa como ha sido posible. La puntuación

³ Para más información sobre MLwin, véase: <http://www.mlwin.com/features/index.html>.

media (APS), basada en las notas obtenidas en Inglés, Matemáticas y Ciencias, es usada, junto con otros dos términos adicionales para medir la «diferencia» entre los resultados de Inglés y Matemáticas. Esto es matemáticamente equivalente a incluir cada una de las asignaturas de forma separada, pero tiene la ventaja de que es menos compleja en el manejo de los datos perdidos y también en la simplificación de la presentación. Se incluye también el término cuadrático de APS, reflejando el hecho de que la relación entre los resultados de la etapa clave 4 y la etapa clave 2, anterior, sea no lineal. También realizamos un ajuste posterior para los efectos «techo» y «suelo» para tratar con grupos muy pequeños de estudiantes en los extremos de la escala donde las predicciones del modelo están fuera del rango del rendimiento recogido por la escala de medida. Hacer ajustes en el techo de la escala es importante, aunque sólo el 2% de los alumnos están en ese rango, su tratamiento tiene un efecto desproporcionado en el pequeño número de selectas «grammar schools» que existen en ciertas áreas. Más información está disponible en la página web de Tablas de Logro y de Rendimiento DCSF: <http://www.dcsf.gov.uk/performancetables/>

Las decisiones sobre qué variables contextuales había que añadir en el modelo tras controlar el rendimiento previo estuvieron basadas en un conjunto mixto de consideraciones estadísticas, educativas y prácticas. Dada la necesidad de aportar información sobre el valor añadido para cada escuela es necesario restringir la elección de la información en los datos nacionales. Así, la finalidad fue generar estimaciones de los efectos de la escuela en el aprendizaje desde los residuos del modelo, e incluir las variables explicativas necesarias para recoger los factores que están fuera del control de la escuela. La elección de las variables contextuales del nivel del alumno tomadas en cuenta son las que han mostrado investigaciones internas y externas sobre los factores que explican la varianza de los resultados de una prueba. La Tabla I aporta la relación completa de variables explicativas y sus coeficientes (para ver la Tabla I, váyase al final del artículo).

No se recogen datos de clase social, ingresos familiares o niveles educativos de los padres a nivel nacional de todos los alumnos. Sin embargo, se dispone de alguna información relacionada con niveles de deprivación. Los niños cuyos padres reciben asistencia social y económica, y algunos beneficios relacionados, pueden reclamar ayuda para el comedor (Free School Meals, FSM). Además del FSM, los modelos utilizan algunas otras medidas de deprivación social que proceden del hecho de que los censos escolares recopilan los códigos postales de los domicilios de los alumnos que pueden vincularse a datos del área local. Se han intentado probar varios indicadores locales pero actualmente los modelos utilizan una medida llamada IDACI -Índice de Deprivación Económica que afecta a los Niños (Income Deprivation Affecting Children

Index)-. Es el porcentaje de niños con menos de 16 años en el área local que viven en familias que reciben ciertos beneficios o con bajos ingresos.

Las necesidades educativas especiales cubren un amplio rango de necesidades que a menudo están interrelacionadas también con necesidades específicas que habitualmente se relacionan con particulares tipos de prejuicios. Los niños con necesidades educativas especiales tendrán necesidades y requerimientos que pueden caer al menos en una de estas cuatro áreas: comunicación e interacción; cognición y aprendizaje; desarrollo conductual, emocional y social; y necesidades sensoriales o físicas. El modelo distingue entre dos niveles de necesidades educativas especiales: «Escuela Acción» (donde el profesor de la clase o el coordinador de necesidades educativas especiales aportan ayuda especial) y «Escuela Acción Plus», donde la intervención de la Escuela Acción no ha producido mejoras y se precisa asesoramiento externo.

El censo escolar recoge datos de 18 grupos étnicos, con un código 19 disponible para «no clasificados» dado que la información sobre este dato es voluntaria. Todos estos grupos fueron incluidos en el modelo como indicadores individuales. Una categoría étnica no significativa fue incluida porque faltaba por razones prácticas y de presentación, sería mejor incluir todas las categorías más que combinar dos de ellas con otros grupos sobre la base de los datos de un año (incluir estas variables produce muy pocas diferencias en el modelo general). En definitiva, hay términos de interacción para los grupos étnicos y el estatus de ayuda al comedor.

Las otras variables del nivel de alumnos abarcan un abanico de factores que están asociados con distintas tasas de progreso. Además, en Inglaterra, dado que la mayoría de los alumnos del mismo grupo anual han nacido dentro del mismo año, su mes de nacimiento es significativo: los alumnos que han nacido más tarde tienden a tener resultados más bajos pero realizan progresos más rápidos. Otra variable asociada con bajos resultados pero con rápido progreso debida en parte al dominio de la primera lengua de los alumnos -Inglés y «otra lengua distinta al inglés»-. Ahora incluimos una interacción de términos para el Inglés y un lenguaje adicional (EAL) y los niveles de rendimiento previo. La movilidad de los estudiantes tiende a reducir el progreso y el modelo CVA toma este hecho en cuenta al utilizar el dato del censo escolar sobre la fecha de entrada en la escuela de cada alumno. Finalmente, este es un indicador para los alumnos que están al cuidado de sus autoridades locales, viviendo con padres de acogida localizados en hogares o en otras formas de cuidado residencial o situándose en casas con sus padres.

Hay ventajas y desventajas en la inclusión de variables contextuales del nivel de escuelas. Por ejemplo, se pueden controlar cuestiones adicionales que afectan a todos

los alumnos en la escuela, pero pueden ser parcialmente endógenas, al medir factores que pueden estar influenciados por la propia eficacia de la escuela. Ellos sitúan a las escuelas en un nivel más igualitario de tal manera que las estimaciones de valor añadido pueden aportar un indicador más justo de eficacia escolar, pero al hacer este ajuste, la información puede no ser útil para los padres u otros sectores que están menos interesados en realizar comparaciones sobre esta base. Los modelos CVA se usan para una variedad de finalidades y la decisión fue tomada para incluir dos factores del nivel de escuelas en los modelos para las escuelas secundarias, describiendo la media general del rendimiento previo en el acceso a la secundaria, pero no se incluyen factores adicionales como el nivel de mestizaje étnico en la escuela, el nivel de deprivación social y otros varios más. Estas variables fueron los niveles generales medios y la dispersión (medida a través de la desviación típica) del rendimiento previo APS.

La modelización multinivel (MLM), teniendo en cuenta la estructura de los datos educativos, ofrece un conjunto de complejos modelos más allá de los Mínimos Cuadrados Ordinarios (OLS). La complejidad añadida es vista en ocasiones como una desventaja, particularmente cuando los modelos que resultan son bastante similares, en ambos términos de la estimación de los efectos fijos y de los residuos de valor añadido. En Inglaterra ambas versiones de los modelos OLS y MLM han sido calculadas y comparadas (Ray, 2006). La principal diferencia práctica es que las estimaciones de los modelos de valor añadido multinivel incorporan la concentración (shrinkage). El grado de concentración depende del tamaño de la escuela: las escuelas más pequeñas se concentran hacia la media nacional. No hay una solución fácil para el problema de interpretar el valor añadido en las escuelas pequeñas. La restricción del valor añadido de una escuela a partir de un determinado tamaño significa que la mejora de las escuelas y los sistemas de rendición de cuentas dejan a algunas escuelas fuera. Hacer promedios por más de un año puede aportar figuras más robustas pero impide las comparaciones anuales.

El modelo para escuelas primarias

El modelo empleado para las escuelas primarias es bastante parecido al modelo de secundaria descrito previamente. La principal diferencia es que los términos del nivel de escuelas se omiten: esto es debido a que el patrón de ingresos es menos variable

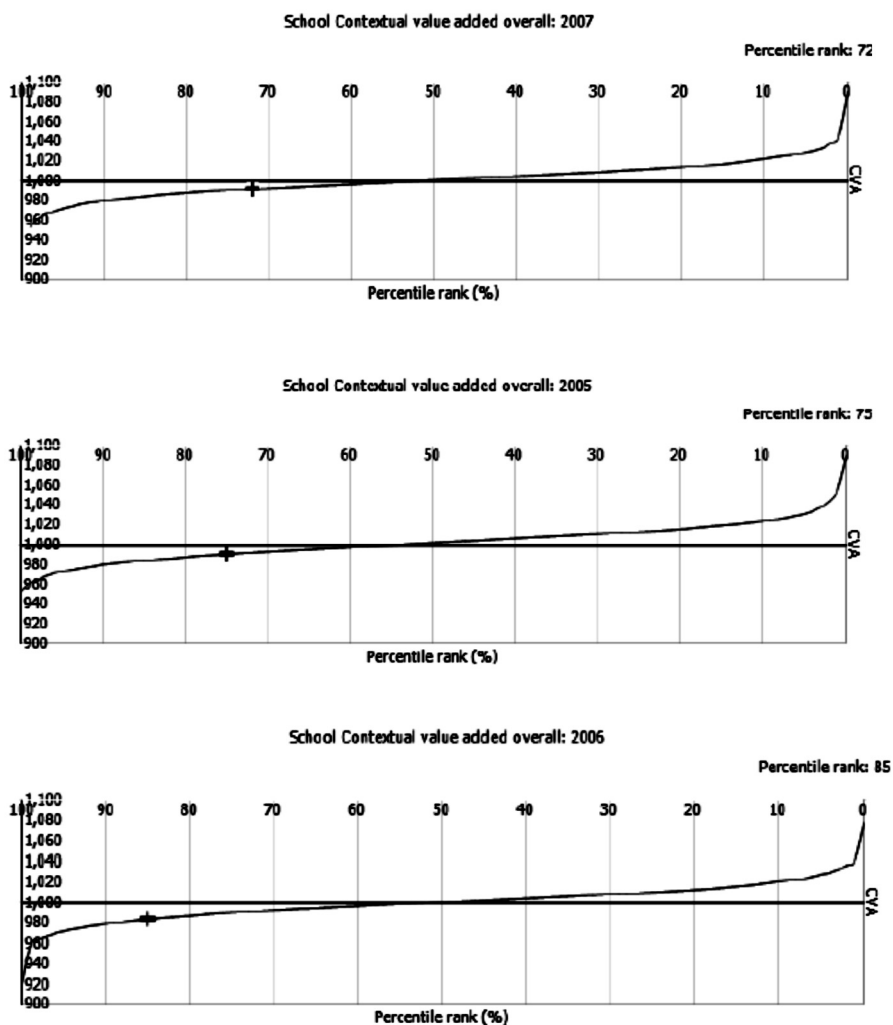
que en las escuelas secundarias, y también porque, con menores números de estudiantes en cada cohorte en primaria, el uso de la misma variable de rendimiento previo en dos niveles puede llevar a modelos más inestables.

El uso de valor añadido contextual con RAISEonline

RAISEonline fue introducido en 2006. Aporta un rango más extenso de datos que las tablas de rendimiento, incluyendo valor añadido para una amplia gama de medidas de resultados y para subgrupos de alumnos dentro de cada escuela. Las escuelas usan RAISEonline como una parte de su autoevaluación y en el proceso de establecer metas que adoptan con la ayuda de los School Improvement Partners. Los datos también están a disposición de los inspectores de Ofsted para su empleo en la evaluación de la extensión de las mejoras escolares o en la capacidad de las mismas para mejorar. Sin embargo, las estadísticas detalladas en RAISEonline no están disponibles para el público en general.

RAISEonline es interactivo y ofrece la posibilidad de investigar cualquier aspecto del rendimiento escolar que sea de particular interés. Además, también provee informes estandarizados que pueden imprimirse y que establecen los indicadores claves en un formato común. Este informe completo para una escuela secundaria puede tener, por ejemplo, aproximadamente 80 páginas. Informes excepcionales, más cortos, también están disponibles para las áreas de mayor interés que aparecen con los puntos fuertes o débiles, basadas en las pruebas estadísticas de significación que comparan las figuras de la escuela con las medias nacionales. El informe completo comienza con alguna información sobre la composición de la escuela, por ejemplo, el porcentaje de alumnos en distintos grupos étnicos y una comparación del rendimiento previo de los alumnos con el patrón nacional (las escuelas con un rendimiento previo relativamente superior se espera que obtengan también mejores resultados). El informe aporta las puntuaciones de valor añadido contextual, primero en una tabla y después en varios gráficos.

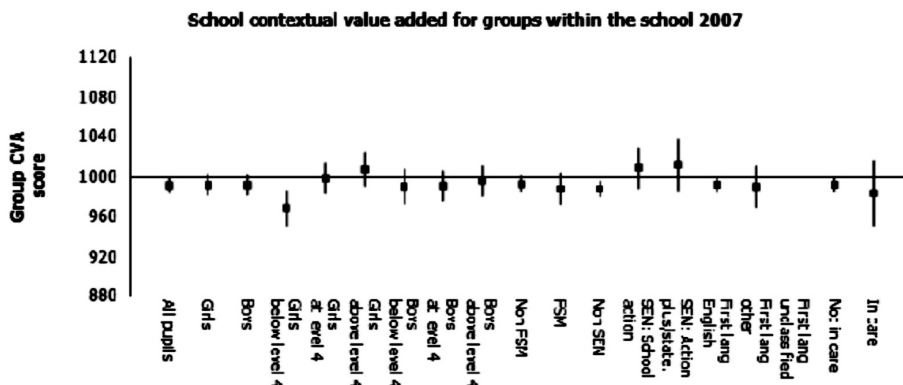
GRÁFICO II. Valor añadido contextual en las etapas clave 2 a 4. General



El Gráfico II reproduce un ejemplo de tres gráficos que se muestran en una página de un informe completo. Los gráficos muestran el valor añadido contextual en diferentes años –gráficos similares están disponibles para distintas asignaturas-. El texto que acompaña indica que: «esta sección aporta la información general de la medida del valor añadido contextual para la escuela respecto a una media nacional de 1.000. La escuela está situada dentro de la distribución nacional para ilustrar el rango de

puntuaciones CVA obtenidas por otras escuelas públicas del conjunto principal. Si el intervalo de confianza para el 95% no se cruza con la línea de la media general, la escuela difiere significativamente de la media general». La CVA para esta escuela aparece por debajo de la media en términos generales, sin embargo, no claramente por debajo en 2006. Una de las cuestiones esenciales para las medidas que están basadas no sólo en una cohorte es si los cambios año a año pueden realmente significar cambios en la eficacia de la escuela. Aquí, por ejemplo, el usuario puede buscar explicación al aparente descenso en 2006 al considerar la naturaleza de la cohorte o de la enseñanza en ese año, o concluir que esta diferencia tan pequeña probablemente no es significativa desde el punto de vista educativo y así centrarse en una variedad de cuestiones sobre por qué la escuela no está por encima de la media en ningún año.

GRÁFICO III. Grupos de alumnos-CVA general

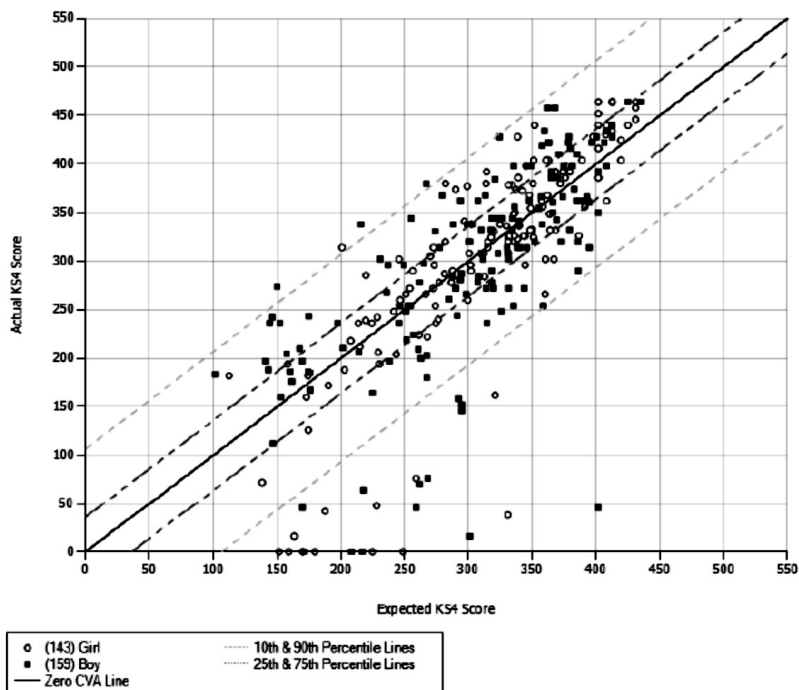


RAISEonline también ofrece valor añadido contextual para subgrupos de alumnos dentro de una escuela. El Gráfico III reproduce un gráfico de ejemplo de un informe completo. Con subgrupos de alumnos es importante reconocer que las pequeñas desviaciones con respecto a la media nacional son difícilmente significativas. Este gráfico muestra que las niñas con bajo rendimiento previo realizan un progreso significativamente menor que lo que se podría esperar (hay 38 niñas en esta categoría).

El Gráfico IV aporta un tercer ejemplo del uso del valor añadido contextual en RAISEonline. Se comparan los actuales resultados de un alumno individual con los resultados que se le predecirían en un modelo de regresión. Las líneas muestran la media nacional y la escuela puede identificar a los alumnos cuyos resultados fueron mayores o menores de lo esperado. Dentro de RAISEonline es posible hacer clic en uno de los

puntos de este gráfico y el sistema ofrecerá detalles de ese alumno en particular. El gráfico también puede filtrar: es posible seleccionar un grupo específico, como los alumnos con ayuda de comedor, y RAISEonline destacará a esos alumnos específicos en el gráfico.

GRÁFICO IV. CVA en etapas clave 2-4. General, predichos frente a los actuales para alumnos (2007)



Coverage 93%

Otros usos del valor añadido en la mejora escolar

Además del uso en las escuelas, Ofsted y School Improvement Partners hacen RAISEonline. La disponibilidad de las estadísticas de valor añadido aporta una potente fuente de información para que las autoridades locales y el gobierno central apoyen un abanico de actividades e iniciativas. Por ejemplo, ha sido posible identificar aparente bajo rendimiento en las escuelas, utilizando el valor añadido contextual y aportar apoyo orientado

al programa RATL (Raising Achievement Transforming Learning, Transformación de la Enseñanza para el Incremento del Rendimiento). La disponibilidad de los resultados de valor añadido también ofrece información para monitorizar el progreso de grupos de escuelas con respecto a políticas específicas y a acuerdos administrativos.

La publicación anual del valor añadido en las tablas de logro y rendimiento de las escuelas e institutos aporta datos nacionales consistentes sobre el rendimiento de las escuelas, para informar a los padres y al público en general, y asegura que las escuelas sean las responsables de sus resultados. Las tablas son un recurso intensivo producido de forma precisa y anual y mantienen deliberadamente un abanico limitado de indicadores clave. Por tanto, no aportan, por ejemplo, un rango completo de resultados vistos en RAISEonline –los usuarios se dirigen a la inspección Ofsted para tener una imagen completa de una determinada escuela–.

Sin embargo, el valor añadido no está disponible a nivel nacional para estudiar la contribución realizada por los profesores individuales, se puede buscar dentro de las escuelas donde la responsabilidad sobre la enseñanza de materias concretas es clara. El valor añadido contextual puede también contribuir a la mejora de la escuela a través de una vinculación indirecta con el salario, por lo que muchos profesores individuales han progresado hacia una mayor escala de salario a través de la acreditación del progreso. Por este motivo, se les pide que aporten evidencia de que es un resultado de su enseñanza: «sus alumnos rinden bien con respecto a su rendimiento previo, realizando progresos tan buenos o mejores respecto de alumnos similares a nivel nacional».

Los modelos de valor añadido tienen por sí mismos una política de uso, demostrando la relativa importancia de factores que impactan en el progreso de los alumnos, y mostrando que los efectos de la escuela son, de hecho, una parte relativamente pequeña de la variación total. Por consiguiente, abordar los déficits del rendimiento no debe centrarse simplemente en las escuelas con peor rendimiento sino que también necesita considerar el bajo rendimiento «dentro» de las escuelas.

Conclusión

La disponibilidad de datos nacionales y, desde 2002, de información sobre las características de los alumnos individuales de todas las escuelas públicas, ha permitido que el sistema educativo inglés desarrolle varios indicadores de valor añadido para las escuelas

y grupos de alumnos dentro de las escuelas. Éstos tienen ahora una importancia capital en las políticas escolares de mejora y aportan información importante para los inspectores escolares, para los gobiernos locales y nacional, para los padres y para el público en general. RAISEonline ofrece un medio para analizar estas estadísticas y este artículo ha aportado ejemplos del tipo de análisis que este paquete estadístico puede ahora contribuir a realizar.

La experiencia en Inglaterra ha demostrado que las medidas de valor añadido pueden desarrollarse y aplicarse y que van de alguna manera más allá de las medidas brutas de rendimiento al tratar de identificar la contribución que hacen las escuelas al progreso de sus alumnos. Sin embargo, el sistema actual es nuevo y todavía necesita cierta discusión sobre las mejores combinaciones de indicadores que se deben utilizar en distintas situaciones y para distintos propósitos. Los aspectos de la aplicación en Inglaterra pueden recomendarse a otros países que buscan desarrollar medidas de valor añadido, así como contribuye a valorar la importancia de pilotar y evaluar, de la consulta con las escuelas y con los expertos en educación y de aportar documentación clara y transparente, así como de las herramientas para ayudar a las escuelas, y a otros usuarios a entender las medidas. El modelo óptimo de uso para un país determinado dependerá de las circunstancias locales y de los datos para analizar.

Referencias bibliográficas

- DEPARTMENT FOR CHILDREN, SCHOOLS AND FAMILIES/ QUALIFICATIONS AND CURRICULUM AUTHORITY, DFEE/QCA (1998). *The Autumn Package.*, London: Department for Education and Employment.
- (2007). Schools and Pupils in England, January 2007 (Final). *Statistical First Release: SFR 30/2007.*
- GOLDSTEIN, H. (2003). *Multilevel Statistical Models* (3rd ed.) London: Arnold.
- GOLDSTEIN, H., RASBASH, J., YANG, M., WOODHOUSE, G., PAN, H., NUTTALL, D. & THOMAS, S. (1993). A Multilevel Analysis of School Examination Results. *Oxford Review of Education*, 19, 425-433.
- MORTIMORE, P., SAMMONS, P., STOLL, L., LEWIS, D. & ECOB, R. (1988). *School Matters: The Junior Years.* Wells: Open Books.
- NATIONAL AUDIT OFFICE (2003). *Making a Difference: Performance of maintained secondary schools in England.* London: The Stationery Office.

RAY, A. (2006). *School Value Added Measures in England: A Paper for the OECD Project on the Development of Value-Added Models in Education Systems*. London: Department for Education and Skills.

SCHOOL CURRICULUM AND ASSESSMENT AUTHORITY (SCAA) (1994). *Value Added Performance Indicators for Schools*. London: School Curriculum and Assessment Authority.

TYMMS, P. & COE, R. (2003). Celebration of the Success of Distributed Research with Schools: The CEM Centre, Durham. *British Educational Research Journal*, 29, 639-653.

Fuentes electrónicas

DCSF ACHIEVEMENT AND ATTAINMENT TABLES WEBSITE. Consultado de:

<http://www.dcsf.gov.uk/performance/tables/>

MLWIN GUIDANCE. Consultado de: <http://www.mlwin.com/features/index.html>

RAISE ONLINE WEBSITE. Consultado de: <https://www.raiseonline.org/>

Dirección de contacto: Andrew Ray. Department for Children, Schools and Families. School Analysis and Research Division. Sanctuary Buildings. Great Smith St, London, SW1P 3BT, United Kingdom. E-mail: Andrew.RAY@dcsf.gsi.gov.uk

TABLA I. Modelo de regresión en 2007 para las etapas claves 2-4

Dependant variable = capped KS4 point store

-2logL = 6304400

Number of pupils = 563,908

Explanatory factor	Variable	Estimate	Std. Error
Prior attainment	Intercept	162.1	13.1
	KS2 student APS	-5.94	0.24
	KS2 APS (using fine grades) – squared	0.38	0.01
	KS2 English PS deviation	1.4	0.07
	KS2 Maths PS deviation	-0.11	0.07
Deprivation	Does student have FSM?	-22.9	0.33
Deprivation of pupil's local area	Deprivation indicator – IDACI score	-59.51	0.67
Special Educational Needs	Does student have SEN Statement/ Action Plus?	-65.76	0.38
	Does student have SEN - school action?	-34.37	0.31
In care	Has the student ever been in care at this school?	-27.1	1.17
Mobility	Student joined other than Jul/Aug/Sep?	-23.43	0.42
	Student joined within last 2 yrs?	-73.55	0.63
Gender	Is student female?	14.52	0.19
Age	Age within year	-12.94	0.3
Language	Is English not the student's first language?	-8.32	8.25
English as Additional Language * Prior	EAL interacted with prior attainment	4.93	0.68
Attainment	EAL interacted with prior attainment squared	-0.14	0.01
Ethnic group	White Irish	-3.61	1.56
	White Irish Traveller	-64.92	7.66
	White Gypsy/Roma	-54.15	5.25
	White Other	10.52	0.79
	Mixed White/Black Caribbean	-3.08	1.03
	Mixed White/Black African	8.03	2.19
	Mixed White/Asian	10.81	1.38
	Any other Mixed ethnic group	6.03	1.08
	Indian	24.08	0.83
	Pakistani	17.53	0.92
	Bangladeshi	22.91	1.47
	Any other Asian ethnic group	24.76	1.35
	Black Caribbean	13.31	0.9
	Black African	29.19	1.04
	Any other Black ethnic group	11.16	1.62
	Chinese	32.78	1.77
	Any other ethnic group	20.68	1.39
Unclassified ethnic group	-7.42	0.74	
Ethnic group * FSM	White Irish and FSM	2.56	4.04
	White Irish traveller and FSM	11.36	11.98
	White Gypsy/Roma and FSM	25.62	7.79
	White other and FSM	26.53	1.79
	Mixed White/Black Caribbean and FSM	8.46	2.08
	Mixed White/Black African and FSM	3.87	4.46
	Mixed White/Asian and FSM	8.27	3.54

Explanatory factor	Variable	Estimate	Std. Error
	Any other Mixed ethnic group and FSM	11.24	2.48
	Indian and FSM	16.04	1.94
	Pakistani and FSM	17.3	1.24
	Bangladeshi and FSM	19.48	1.89
	Any other Asian ethnic group and FSM	24.71	2.87
	Black Caribbean and FSM	18.2	1.79
	Black African and FSM	20.46	1.65
	Any other Black ethnic group and FSM	8.69	3.19
	Chinese and FSM	34.11	5.12
	Any other ethnic group and FSM	30.83	2.32
	Unclassified ethnic group and FSM	7.14	1.95
Level of school prior attainment	School KS3 APS (using fine grades) for CVA	1.96	0.36
Spread of school prior attainment	School std dev of KS3 APS for CVA	-4.82	0.92
	Random components:	Estimate	Std. Error
	Between school variance	319.42	8.77
	Within school variance	4135.6	7.81
	Variance partition coefficient	0.08	