



REDIE. Revista Electrónica de
Investigación Educativa

E-ISSN: 1607-4041

redie@uabc.edu.mx

Universidad Autónoma de Baja California
México

Martínez Abad, Fernando; Lizasoain Hernández, Luis; Castro Morera, María; Joaristi
Olariaga, Luis María
Selección de escuelas de alta y baja eficacia en Baja California (México)
REDIE. Revista Electrónica de Investigación Educativa, vol. 19, núm. 2, abril-junio, 2017,
pp. 38-53
Universidad Autónoma de Baja California
Ensenada, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=15550741004>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

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

Vol. 19, Núm. 2, 2017

Selección de escuelas de alta y baja eficacia en Baja California (México)

Selection of High and Low-Effectiveness Schools in Baja California (Mexico)

Fernando Martínez Abad (1) fma@usal.es
Luis Lizasoain Hernández (2) luis.lizasoain@ehu.es
María Castro Morera (3) maria.castro@edu.ucm.es
Luis María Joaristi Olariaga (2) luis.joaristi@ehu.es

(1) Universidad de Salamanca

(2) Universidad del País Vasco

(3) Universidad Complutense de Madrid

(Recibido: 30 de octubre de 2014; Aceptado para su publicación: 5 de noviembre de 2015)

Cómo citar: Martínez, F., Lizasoain, L., Castro, M. y Joaristi, L. M. (2017). Selección de escuelas de alta y baja eficacia en Baja California (México). *Revista Electrónica de Investigación Educativa*, 19(2), 38-53. <https://doi.org/10.24320/redie.2017.19.2.960>

Resumen

El presente estudio planteó la identificación y selección de escuelas de Baja California con alta y baja eficacia a partir de los puntajes de ENLACE y pruebas de contexto. Se obtuvo una muestra representativa de 85 escuelas y alrededor de 7000 estudiantes en los años 2010 y 2011, y se aplicaron modelos jerárquicos lineales para el cálculo de las puntuaciones residuales en el nivel de centro. Tras el ajuste de los 8 modelos finales, que aíslan el efecto de las variables contextuales significativas, se seleccionan, bajo criterios estadísticos, 15 centros con alto y bajo residuo. Los resultados muestran que las distribuciones de las escuelas seleccionadas son similares en ambos casos a la población de referencia. Los resultados afirman la robustez del procedimiento de selección implementado y destacan la importancia del mismo para la detección de buenas prácticas en las escuelas de alta eficacia.

Palabras clave: Calidad de la educación, éxito escolar, evaluación, eficacia del centro de enseñanza.

Abstract

This study proposed to identify and select high and low-effectiveness schools in Baja California based on ENLACE (Quality of Educational Attainment Exams) scores and context tests. A representative sample of 85 schools and around 7,000 students was obtained for 2010 and 2011. Hierarchical linear models were applied to calculate residual scores at a school level. After adjustment of the 8 final models, which isolate the effect of significant contextual variables, statistical criteria were used to select 15 schools with high and low residual scores. The results show that the distributions of schools selected are similar in both cases to the reference population. The results show the robustness of the selection procedure used and highlight its significance in detecting good practices in high-effectiveness schools.

Keywords: Quality of education, school success, evaluation, school effectiveness, achievement control.

I. Introducción

El desarrollo de pruebas de evaluación del logro tanto a nivel internacional como nacional (Educational Resources Information Center and National Science Foundation, 1996; Mullis, Martin, Kennedy, Trong y Sainsbury, 2009; OECD, 2012; Tatto et al., 2008; UNESCO, 2005), entre las que se encuentran las pruebas ENLACE, desarrolladas a lo largo de todo el territorio mexicano, se ha generalizado de cara a asegurar la calidad, el control, la estandarización y la mejora de los resultados educativos de los estudiantes.

Las pruebas ENLACE surgieron con la intención de generar información sobre los conocimientos y capacidades que los estudiantes tienen, a fin de servir de apoyo a la implementación de procesos efectivos y pertinentes de planificación educativa y política pública (Martínez, Chaparro, Lizasoain, Caso y Urias, 2014). Paralelamente a esta escalada de pruebas de rendimiento a gran escala durante los últimos 25 años, que deriva en multitud de estudios comparativos acerca del rendimiento escolar de estudiantes y escuelas de diferentes regiones (Drabowicz, 2014; Kjaernsli, 2004) surge la preocupación por estudiar ciertas variables diferenciales que no son controlables por las instituciones escolares, y que inciden en el rendimiento final de los estudiantes (González, Caso, Díaz López y López, 2012; Joaristi, Lizasoain y Gamboa, 2012; Jornet, González y Perales, 2012).

Fruto de la investigación generada en este ámbito, los modelos jerárquicos lineales se posicionan –dada su naturaleza anidada, multivariante y multinivel– como una alternativa potente y válida para ejercer un control sobre el efecto moderador de estas variables contextuales en el rendimiento académico de los estudiantes y centros educativos y eliminando el efecto moderador que ejercen estas variables contextuales sobre el rendimiento (Bryk y Raudenbush, 1992; Goldstein, 1995; Willms y Raudenbush, 1989).

En este contexto, la noción de eficacia diferencial de una escuela se equipara a menudo con análisis de ganancias residuales basadas en evaluaciones del rendimiento de los alumnos, como por ejemplo el sistema de evaluación de Dallas (Meyer, 1996; Webster y Mendro, 1997) o la estrategia de Fitz-Gibbon (1997). Se puede definir la ganancia residual de una escuela como la diferencia entre el rendimiento medio observado que poseen los estudiantes de la escuela testeados y el rendimiento estimado a partir de las variables contextuales explicativas incluidas en el modelo. Así, se podría hablar de esta ganancia como una medida del valor añadido de la escuela, entendiendo el valor añadido como una puntuación media del centro superior a la que cabría esperar después de controlar los efectos de las variables contextuales.

Aunque parece que no existen grandes diferencias entre los resultados obtenidos por los modelos transversales y los puramente longitudinales (Martínez, 2009), el empleo de estos últimos, que incluyen medidas repetidas del rendimiento de los estudiantes, se ha generalizado en la investigación científica (OECD, 2008), al recoger más información en el modelo.

En contraste con los modelos lineales jerárquicos de carácter transversal, los estudios longitudinales proponen el análisis de modelos de efectos mixtos con medidas repetidas. Valga como ejemplo el sistema de evaluación del valor añadido de Tennessee (Sanders y Horn, 1995) o los análisis de Bryk, Thum, Easton y Luppescu (1998) en Chicago. En ambos estudios se estima la verdadera ganancia basándose en las puntuaciones individuales de los alumnos y las diferencias entre el crecimiento estimado y alcanzado, sin considerar ganancias brutas. De este modo, el modelo estadístico aplicado para el estudio del crecimiento de los estudiantes y de los centros educativos aprovecha de mejor manera la información disponible en las aplicaciones apareadas de las pruebas, ya que se tiene en cuenta la matriz de varianzas-covarianzas completa (Bryk y Weisberg, 1976; Goldstein y Spiegelhalter, 1996).

Por tanto, se entiende como valor añadido en educación la diferenciación y aislamiento de los factores puramente escolares que influyen en los resultados académicos, de los otros factores individuales y sociales que resulten significativos en este rendimiento. Para identificar aquellos centros escolares que obtienen los mejores y peores resultados, se analizan las puntuaciones obtenidas mediante técnicas estadísticas multivariantes, y más específicamente modelos lineales jerárquicos. Este conjunto de técnicas

estadísticas respetan la estructura anidada de los datos, que es la habitual en educación, y permiten estudiar conjuntamente los efectos de las variables de nivel y de contexto, facilitando el análisis de los efectos de las covariables en cada nivel de anidamiento (Bryk y Raudenbush, 1992; Goldstein, 1995).

Dada la dificultad técnica para realizar estudios puramente longitudinales con tamaños de muestra adecuados a las técnicas multivariantes requeridas y los aceptables resultados obtenidos por el enfoque transversal, existen multitud de trabajos vinculados con esta perspectiva con una orientación similar a la que se presenta. Se puede señalar el estudio de Marchesi y Martín (2002), que lleva a cabo una evaluación en el ciclo de Educación Secundaria contemplando los niveles estudiante y escuela. Por su parte, Gaviria, Martínez-Arias y Castro (2004) implementaron un estudio sobre los factores asociados a la eficacia escolar en el sistema educativo brasileño teniendo en cuenta 3 niveles: estudiante, profesor y estado federado. Cabe destacar también los análisis de Murillo y Hernández-Castilla (2011a, 2011b), que abarcan los niveles de logro socio-afectivos de los estudiantes mediante un modelo multinivel de 4 niveles (estudiante, aula, centro y país); por último, otro modelo multinivel, en este caso de 3 niveles (estudiante, escuela y país), es aplicado por Murillo y Román (2011) para analizar el valor añadido que muestra el centro escolar a partir de la base de datos del Segundo Estudio Regional Explicativo y Comparativo (SERCE) de la UNESCO-OREAL de 2009.

A partir de este contexto, ubicamos el estudio aquí presentado en la primera fase de un proyecto más amplio.¹ La identificación y selección de los centros escolares con mayor y menor residuo tiene como objetivo final buscar elementos comunes en ambos grupos y compararlos, de cara a localizar e informar a la comunidad educativa acerca de buenas prácticas que se están implementando en los centros de alto residuo (en contraste con los de bajo residuo) que están colaborando en la mejora del rendimiento académico. Dicha búsqueda de las buenas prácticas se llevará a cabo a través de un estudio de caso en profundidad de los centros seleccionados. Así pues, la primera fase del estudio, que aquí se describe, trata de discriminar y seleccionar escuelas de alta y baja eficacia eliminando previamente el efecto distorsionador de las variables contextuales significativas, de cara a realizar posteriormente un análisis e identificación más profundos de los factores no contextuales asociados al rendimiento (labor docente, organización y coordinación, liderazgo del equipo directivo) que los centros de alta eficacia implementan en contraste con los centros de baja eficacia.

Cabe esperar, por tanto, que los centros educativos cuyo residuo con respecto al rendimiento académico sea superior al esperado desarrollen ciertas prácticas educativas positivas que ayuden a explicar parte de la variabilidad que el modelo planteado no fue capaz de explicar por sí mismo. Al contrario, es también de esperar que aquellos centros con residuo inferior al esperado desarrollen su labor bajo ciertas condiciones y dificultades que lastren sus niveles de rendimiento, sin excluir la posibilidad de eventuales prácticas educativas que puedan considerarse poco adecuadas.

Al respecto, algunos trabajos ejemplifican el estado de la cuestión y asumen diseños o enfoques metodológicos similares a los aquí planteados, como el de Jackson y Lunenburg (2010), que estudian 24 centros escolares con diversos niveles de rendimiento medio, y los caracterizan como ejemplares, reconocidos, aceptables y no aceptables. Por su parte, Ray, McCormak y Evans (2009) aplican modelos de valor añadido para caracterizar los centros seleccionados en Gran Bretaña. Por otro lado, Wright (2009), en el contexto de la *No Child Left Behind Act* (NCLB), obtiene un conjunto de buenas prácticas centradas en el clima organizacional a partir un estudio de caso de las escuelas que no cumplen el estándar de progreso anual adecuado y que establecieron alguna estrategia para obtener mejoras. Gritter (2006) estudia los factores determinantes del alto rendimiento a partir del análisis comparado de las escuelas de alto rendimiento del programa *Just for the kids*. Por último, Huffman, Pankeke y Muñoz (2006) combinan un estudio inicial multinivel con la aplicación de entrevistas a diversos agentes educativos de los centros de alto valor añadido seleccionados.

El presente estudio, enmarcado en un proyecto más amplio destinado a la detección de buenas prácticas

¹Caracterización de las buenas prácticas en las escuelas secundarias de alto valor añadido en Baja California, proyecto 189833, Conacyt (México). Convocatoria de Investigación para la Evaluación de la Educación 2012. Responsable técnico: Dr. Joaquín Caso Niebla.

que colaboren en la mejora del rendimiento académico de las escuelas, tiene como objetivo principal la identificación y selección, mediante modelos lineales jerárquicos, de las escuelas de Baja California con mayor y menor residuo, o ganancia con respecto al valor esperado, a partir de los datos que proporcionan las pruebas ENLACE de Matemáticas y Español aplicadas en 2010 y 2011 y los Cuestionarios de Contexto aplicados por la Unidad de Evaluación Educativa de la UABC (Caso, Díaz, Chaparro y Urías, 2011; Caso, Chaparro, Díaz y Urías, 2012). Cabe destacar también otros objetivos en el trabajo, como la identificación de los principales factores contextuales asociados al rendimiento académico en el caso de las escuelas de Baja California (México) y la definición de un procedimiento estable y eficaz para la selección de centros educativos de alta y baja eficacia a partir de las ganancias residuales.

II. Método

Los datos que se tomaron en cuenta para este estudio parten de la población de estudiantes, docentes y directores de escuelas de educación secundaria del estado de Baja California (México). La selección inicial de las escuelas se llevó a cabo en el 2010 y 2011, mediante el método de muestreo por conglomerados en tres etapas, con probabilidades proporcionales al tamaño y con un nivel de confianza del 95%, siendo la última unidad de selección la escuela y la unidad de observación el propio estudiante. La totalidad de los agentes educativos señalados de las escuelas seleccionadas son evaluados en cada centro.

La muestra participante en el estudio consistió en estudiantes de primer y segundo curso (que en 2011 cursaban segundo y tercero, respectivamente); así como sus profesores, de Español y Matemáticas, y los directores de los planteles donde estudiaban durante los años 2010 y 2011.

La muestra final se obtuvo operando con ambas bases de datos y seleccionando aquellos centros educativos evaluados tanto en 2010 como en 2011. Así se obtuvieron 7036 estudiantes para Español y 6762 para Matemáticas tanto en los años 2010 como 2011; 153 docentes de Español y 151 de Matemáticas en 2010 y 128 docentes de Español y 129 de Matemáticas en 2011; y 71 directores en 2010 por 72 en 2011, como se muestra en la tabla I.

Tabla I. Distribución de los estudiantes por grado académico, asignatura y año

	2010		2011		total
	n	%	n	%	
Matemáticas	1o. 4009	59.3	2o. 4013	59.3	6762
	2o. 2753	40.7	3o. 2749	40.7	
Español	1o. 4336	61.6	2o. 4342	61.7	7036
	2o. 2700	38.4	3o. 2694	38.3	

La distribución por sexo y país de nacimiento de los estudiantes fue, tanto en Matemáticas como en Español, prácticamente, la misma. Mientras que en Matemáticas participaron 3246 hombres, un 48.6%, por 3438 mujeres, un 51.4%, en Español lo hicieron 3397 hombres (48.9%) por 3547 mujeres (51.1%). Por otro lado, al contar con datos pareados, las distribuciones por estudiante son iguales en las muestras de los años 2010 y 2011.

En cuanto a las escuelas seleccionadas, se obtuvo un total de 85, por las cuestiones comentadas previamente. Las diferencias entre el número de escuelas y el número de directores en la muestra se deben a que la unidad de análisis estudiada es lo que se ha denominado escuela-turno, es decir, la combinación de una escuela por turno de docencia (matutino, vespertino o discontinuo). Esta es una característica específica del sistema educativo mexicano, pues cuando una escuela presta sus servicios en dos turnos, de facto pasa a operar como dos escuelas diferentes en la medida en que, en la gran mayoría de los casos, únicamente se comparte el edificio y los recursos físicos, pues tanto el plantel de docentes como el equipo directivo son habitualmente distintos. En la tabla II se puede observar la distribución de los centros de la muestra por modalidad y municipio.

Tabla II. Distribución de la muestra de centros por modalidad y municipio

		n	%
Modalidad	General	39	45.9
	Particular	16	18.8
	Técnica	20	23.5
	Telesecundaria	10	11.8
Municipio	Ensenada	18	21.2
	Mexicali	27	31.8
	Playas de Rosarito	4	4.7
	Tecate	5	5.9
	Tijuana	31	36.5

Las distribuciones por modalidad y municipio se ajustan al censo existente en Baja California, por lo que la selección de los conglomerados no parece ser fuente de sesgo muestral.

ENLACE es una prueba objetiva estandarizada, de aplicación censal en todo México, que ofrece un diagnóstico individual de los estudiantes en varias materias del currículo. Se centra en evaluar el conocimiento alcanzado por los estudiantes en su escolarización formal. Las pruebas ENLACE en la educación básica tratan de evaluar conocimientos y habilidades que los estudiantes poseen en Español y Matemáticas (Dirección General de Evaluación de Políticas, 2010, 2011).

La información relativa a las variables de contexto se obtiene como parte de la Estrategia Evaluativa Integral 2010 y 2011 (Caso, Díaz, Chaparro y Urías, 2011; Caso, Chaparro, Díaz y Urías, 2012). Las variables incluidas en el estudio, puestas a prueba de cara a su inclusión en los modelos, se muestran en el anexo I del presente trabajo.

2.1 Procedimiento

El primer paso consistió en la selección de la muestra final a partir de un proceso de emparejamiento de los datos de 2010 y 2011. Debido a este proceso, aproximadamente un tercio de los datos fue eliminado de la muestra. Este procedimiento fue seguido por un segundo, en el que se eliminaron los centros educativos que contaran con menos de 10 estudiantes.

Una vez definida la muestra final fueron necesarias tres fases independientes: La primera, consistente en la elección del modelo concreto bajo el cual se aplicarían los modelos jerárquicos; la segunda, enfocada a la aplicación de los modelos, selección de covariables y cálculo del residuo de cada escuela para su selección final; y la última, dedicada a la descripción y caracterización general de las escuelas seleccionadas.

En cuanto a la elección del modelo más adecuado, dada la propia naturaleza de los datos, en los que se disponía únicamente de dos medidas repetidas, se descartó la aplicación de modelos longitudinales. Así, se decidió entre los siguientes:

1. *Modelo de ganancia residual*: Se emplea como variable criterio la última medida de rendimiento del estudiante y como covariable la medida de rendimiento previo, junto con las variables contextuales significativas.
2. *Modelo de ganancia bruta*:² La variable criterio es la diferencia bruta entre la puntuación del estudiante en la última y la primera medida de rendimiento y las predictoras las covariables significativas.

² A pesar de no haber asegurado la equiparación vertical de las escalas, se considera que ENLACE cuenta con características técnicas que aportan cierta seguridad al respecto y, por consiguiente, prueba el modelo como un acercamiento inicial.

3. *Modelo contextualizado sin ganancia*: Se añade como variable criterio la puntuación del estudiante en la última prueba de ENLACE y predictoras las variables contextuales seleccionadas.

Los modelos de ganancia residual se desestimaron al observar que, a pesar de que se obtenían coeficientes de correlación intraclase superiores al 10%, el rendimiento previo de los estudiantes explicaba más del 70% de la varianza total inter-escuela en cada uno de los 4 modelos (por curso y materia). Se entiende así que incorporar como covariable el rendimiento previo afectaría al nivel de significación del resto de covariables de manera importante.

Los modelos de ganancia bruta se rechazan también, al comprobar que los valores del índice de correlación intraclase (icc) son inferiores al 10% en 3 de los 4 modelos planteados. Teniendo en cuenta que para la selección de las escuelas se analizaron los residuos en la varianza del segundo nivel, no resulta deseable partir de poca varianza en el mismo. Así, se seleccionan los modelos contextualizados sin ganancia, de los que se obtendrán 8 medidas de residuo por centro escolar (por curso, año y materia).

Dado que se disponía de 8 modelos independientes, se plantearon 3 criterios para la selección final de las escuelas de alto y bajo residuo:

1. Que la escuela estuviera presente al menos en 6 de las 8 bases de datos manejadas. Tras comprobar este supuesto, la muestra de centros se redujo a 65 unidades.
2. Que la escuela, para ser seleccionada como de residuo alto, se situara al menos en 5 de los 8 modelos planteados en puntuaciones superiores al P66. Además, en el caso de pertenecer en algún residuo al tercio de puntuaciones inferior, debía obtener una puntuación diferencial entre las veces que situó en el tercio superior y en el tercio inferior superior a 4. En el caso de la selección de las escuelas de residuo bajo, el procedimiento es inverso al descrito.
3. El tercer criterio se relacionó con el percentil medio que ocupaban las escuelas en lo que respecta a la distribución de los residuos. Las escuelas seleccionadas de residuo positivo y negativo estaban situadas en el tercio superior e inferior de la distribución de percentiles medios, respectivamente. Para el cálculo de este índice se aplicó la ecuación,

$$\overline{\text{Percentil}}_i = \frac{\sum_{j=1}^m \frac{\text{Rango}_{ij}}{n_j}}{m}$$

m = Número de modelos en los que está presente la escuela i
 Rango_{ij} = Posición directa que ocupa la escuela i en el modelo j
 en función del residuo obtenido
 N_j = Número de escuelas que están presentes en el modelo j

2.2 Análisis de los datos

Se aplicaron modelos jerárquicos lineales de dos niveles (estudiante y escuela) y se evaluó su idoneidad a partir del cálculo del icc y el coeficiente de determinación obtenidos. Se introdujeron tanto las variables de primer nivel, como las de segundo nivel, que son agregadas a partir de las otras, como covariables de efectos fijos, sin contemplar los efectos aleatorios.

El residuo del centro en cada modelo se calcula a partir de la diferencia entre la puntuación obtenida de la predicción teniendo en cuenta efectos fijos y aleatorios en el segundo nivel y la predicción contemplando sólo los efectos fijos de las covariables. A partir de estos residuos se obtienen rangos y rangos medios.

Por último, se procedió al cálculo de medias y frecuencias de diferentes variables contextuales para la caracterización de los centros seleccionados.

III. Resultados

En primer lugar, una vez seleccionado el modelo contextualizado sin ganancia para llevar a cabo los análisis, se probaron los 8 modelos nulos. La tabla III muestra en todos los casos buenos valores del ICC.

Tabla III. Ajuste de los modelos nulos de los modelos contextualizados sin ganancia

		Modelo Nulo		
		icc	Var N1	Var N2
Matemáticas 2010	1o.	28.4%	9909.4	3922.7
	2o.	33.2%	8693.2	4328.4
Matemáticas 2011	2o.	31.9%	10952.3	5135.3
	3o.	32.6%	11130.7	5385.8
Español 2010	1o.	29.9%	10935.8	4669.5
	2o.	26.7%	9458.2	3437.6
Español 2011	2o.	38.0%	10625.3	6513.6
	3o.	29.2%	7671.3	3158.6

Una vez observado el buen comportamiento de los modelos nulos, se procedió al cálculo de cada uno de los 8 modelos incluidos en el estudio. La tabla IV muestra las covariables significativas obtenidas en cada uno de los modelos.

Tabla IV. Covariables significativas para cada modelo*

		Covariables Nivel 1	Covariables Nivel 2
Matemáticas 2010	1o.	n1_sexo n1_mexican n1_edadacc n1_repite n1_expectativas n1_falta.alumno	n2_mod_partic n2_turno n2_sexo n2_isoc n2_recursos.esc
	2o.	n1_mexican n1_edadacc n1_repite n1_expectativas n1_falta.alumno	n2_muni_mex n2_turno n2_expectat n2_recursos.esc
Matemáticas 2011	2o.	n1_sexo n1_repite n1_expectativas n1_falta.alumno	n2_mod_partic n2_sexo n2_falta.alumno n2_isoc
	3o.	n1_repite n1_expectativas n1_falta.alumno	n2_mod_partic n2_falta.alumno n2_recursos.esc
Español 2010	1o.	n1_sexo n1_mexican n1_guarderia n1_expectativas n1_isoc	n2_mod_partic n2_sexo n2_falta.alumno n2_recursos.esc
	2o.	n1_sexo n1_mexican n1_repite n1_expectativas n1_falta.alumno	n2_muni_mex n2_mod_partic n2_turno n2_mexican n2_isoc n2_experesc.prof
Español 2011	2o.	n1_sexo n1_mexican n1_repite n1_expectativas n1_falta.alumno n1_isoc	n2_mod_partic n2_expectativas
	3o.	n1_sexo n1_mexican n1_guarderia n1_expectativas n1_falta.alumno n1_isoc	n2_turno n2_isoc n2_recursos.esc

* En el Anexo I se presentan los nombres completos y una definición de las variables.

A nivel general se observa cómo los modelos obtenidos tienen un número y tipo de covariables similares, incluso entre los que incluyen como variable criterio la puntuación ENLACE en Matemáticas y los que incluyen la puntuación en Español. Las variables más destacadas en el primer nivel son las expectativas académicas del estudiante, que forma parte de los 8 modelos, la puntualidad y asistencia del propio alumno al aula, que se incluye en 7, y el sexo, la condición de repetidor y la condición de nacido en México, que permanecen en 6 modelos. En lo que respecta a las covariables del segundo nivel, se observa una tendencia menos clara, aunque es cierto que el hecho de que la escuela sea privada resulta significativo en 6 casos, los recursos escolares se incluyen en 5 modelos y el índice socioeconómico medio de la escuela en 4.

Si se hace una comparación de los modelos de Español y de Matemáticas, llama la atención que en el caso de la inclusión de las Matemáticas como variable criterio, gana importancia la condición de repetidor del

estudiante en el primer nivel y la existencia de buenos recursos en la escuela en el segundo. Por su parte, en el caso de los modelos que incluyen la puntuación en Español en el criterio, gana importancia, como era de esperar, en el primer nivel que el estudiante sea nativo y el índice socio-económico, tanto en el primer nivel como en el segundo.

El ajuste final de los 8 modelos, que indica mejores valores para los modelos que incluyen el Español que para los que incluyen las Matemáticas, se muestra en la tabla V.

Tabla V. Ajuste final de los modelos aplicados

		Modelo Nulo (%)		
		icc	R2 N1	R2 N2
Matemáticas 2010	1o.	13.60	4.8	62.18
	2o.	13.90	3.2	68.5
Matemáticas 2011	2o.	15.70	4.8	71.1
	3o.	18.50	3.8	54.8
Español 2010	1o.	12.54	4.5	67.9
	2o.	8.31	5.9	76.5
Español 2011	2o.	14.77	6.3	65.4
	3o.	10.75	6.4	72.6

La varianza explicada de los modelos planteados para el rendimiento en Español es superior a los modelos para el rendimiento en Matemáticas en ambos niveles. Asimismo, el icc en el caso de los modelos de Español alcanzan valores más bajos, que indican un porcentaje menor de variabilidad en el segundo nivel tras el cálculo de los mismos.

A pesar de que se obtienen valores del icc por encima del 10% y de que se obtengan varianzas explicadas bajas en el primer nivel, dado que el objetivo del estudio no es tanto obtener modelos con niveles de explicación y ajuste elevados, sino identificar y seleccionar las escuelas de residuo más alto y más bajo, se considera que los modelos son adecuados y se puede proceder con la siguiente fase del mismo.

Tras el proceso de aplicación de los criterios de selección de los centros con residuo positivo y negativo (descritos en el apartado anterior) 30 centros son elegidos, 15 con residuos positivos (RP) y 15 con residuos negativos (RN). En primer lugar, la tabla VI muestra las puntuaciones obtenidas por los centros con alta eficacia tanto en el residuo y percentil medios, como en las veces que pertenecen al tercio superior e inferior de puntuaciones y la diferencia obtenida.

Tabla VI. Centros seleccionados con residuos positivos

Centro	Residuo*	Percentil	P ₁ -P ₃₃	P ₆₆ -P ₁₀₀	Diferencia
RP_01	1.74	0.07	8	0	8
RP_02	1.41	0.14	7	0	7
RP_03	1.21	0.19	7	1	6
RP_04	1.19	0.19	7	1	6
RP_05	1.17	0.23	6	1	5
RP_06	0.89	0.25	6	0	6
RP_07	0.88	0.24	6	0	6
RP_08	0.88	0.22	8	0	8
RP_09	0.72	0.31	6	1	5
RP_10	0.59	0.29	5	0	5
RP_11	0.56	0.30	5	0	5
RP_12	0.56	0.32	6	0	6
RP_13	0.56	0.31	6	0	6
RP_14	0.50	0.31	5	0	5
RP_15	0.50	0.32	5	0	5

* Puntuaciones residuales estandarizadas N (0.1) en cada modelo obtenido.

Por su parte, 15 son los centros seleccionados como de baja eficacia o residuo negativo. Los valores obtenidos en este caso se muestran en la tabla VII.

Tabla VII. Centros seleccionados con residuos negativos.

Centro	Residuo*	Percentil	P ₁ -P ₃₃	P ₆₆ -P ₁₀₀	Diferencia
RN_01	-1.79	0.93	0	8	-8
RN_02	-1.65	0.91	0	7	-7
RN_03	-1.47	0.83	0	6	-6
RN_04	-1.21	0.88	1	6	-5
RN_05	-1.11	0.78	0	6	-6
RN_06	-1.09	0.86	0	7	-7
RN_07	-1.06	0.79	0	6	-6
RN_08	-1.05	0.86	0	7	-7
RN_09	-0.85	0.81	0	7	-7
RN_10	-0.81	0.74	0	5	-5
RN_11	-0.73	0.78	0	6	-6
RN_12	-0.70	0.72	1	6	-5
RN_13	-0.65	0.70	0	5	-5
RN_14	-0.54	0.71	0	5	-5
RN_15	-0.53	0.71	0	6	-6

* Puntuaciones residuales estandarizadas N (0.1) en cada modelo obtenido.

En comparación, mientras que la mayor parte de los centros de alta eficacia se sitúa en puntuaciones medias-bajas del percentil medio, los centros de eficacia inferior se sitúan en percentiles medios más extremos, en el lado superior. El análisis de las puntuaciones diferenciales y de las veces que los centros están en los tercios superior e inferior confirma que los centros de residuo negativo seleccionados se comportaron como tal de manera más sistemática que los de residuo positivo a lo largo de los 8 modelos.

En lo que respecta a la caracterización general de los centros seleccionados, se presenta, en primer lugar, la tabla VIII, que muestra, además del percentil medio de los centros de alto residuo, las calificaciones obtenidas en las pruebas ENLACE en Español y Matemáticas, el índice socioeconómico medio, los recursos medios, la modalidad y el municipio en el que están ubicados.

Tabla VIII. Características generales de los centros seleccionados de alta eficacia

Centro	Percentil	Cal. Mat.*	Cal. Esp.*	ISOC*	Recs.*	Modalidad	Municipio
RP_01	0.07	2.52	2.83	0.97	0.77	Particular	Mexicali
RP_02	0.14	0.53	1.00	-0.69	0.86	Telesec.	Ensenada
RP_03	0.19	1.34	0.53	0.49	1.13	General	Tijuana
RP_04	0.19	1.51	1.82	1.24	0.67	Particular	Mexicali
RP_05	0.22	-0.17	-0.06	-1.25	-0.45	General	Ensenada
RP_06	0.23	2.56	1.56	1.41	1.83	Particular	Ensenada
RP_07	0.24	1.01	0.55	0.54	0.07	Técnica	Ensenada
RP_08	0.25	-0.19	0.79	0.15	1.20	General	Tijuana
RP_09	0.29	1.01	0.61	0.35	1.33	General	Tijuana
RP_10	0.30	1.66	1.16	0.36	1.33	Particular	Tijuana
RP_11	0.31	0.47	0.45	0.51	0.71	Técnica	Mexicali
RP_12	0.31	1.29	1.36	1.09	1.39	Particular	Tecate
RP_13	0.31	-0.03	0.88	-0.67	-0.10	Telesec.	Ensenada
RP_14	0.32	1.29	2.13	1.09	0.09	Particular	Tijuana
RP_15	0.32	0.31	0.00	-0.14	-1.08	General	Tijuana

* Puntuaciones estandarizadas N (0.1) para cada modelo obtenido

Llama la atención la selección de algún centro como RA_05 que, a pesar de tener puntuaciones inferiores a la media tanto en Matemáticas como en Español, con un índice socioeconómico muy bajo y una calidad de recursos del centro baja, se mantienen como de alta eficacia. Dado que el objetivo final del proyecto es localizar prácticas educativas que contribuyan en la mejora del rendimiento académico de los

estudiantes, el análisis de este tipo de centros, como son también el RA_13 o RA_15, tiene un interés especial para las siguientes fases del proyecto en el que se integra este estudio.

Se espera que aquellos centros con condiciones ambientales y estructurales de partida desfavorecedoras tengan más dificultades para desempeñar su labor educativa, y que hagan importantes esfuerzos en la implementación de prácticas que mitiguen esta desventaja inicial. Se entiende que en el caso de estos centros que han sido seleccionados como de alto residuo, y con índices socioeconómicos y recursos disponibles menores a la norma, las prácticas implementadas han sido exitosas, y cobra vital relevancia estudiar sus claves y aspectos diferenciales para la puesta en común a la comunidad científica y educativa.

Por otro lado, cabe destacar que, si bien en la modalidad particular el porcentaje alcanza cuotas algo superiores (40% frente al 18.8% inicial) y en el general algo inferiores (33.3% frente al 45.9% inicial), la selección de centros por municipio y modalidad está repartida de manera similar a la distribución general de centros obtenida inicialmente. La tabla IX muestra los centros seleccionados con residuo negativo, o baja eficacia.

Tabla IX. Características generales de los centros seleccionados de baja eficacia

Centro	Percentil	Cal. Mat.*	Cal. Esp.*	Isoc*	Recs.*	Modalidad	Municipio
RN_01	0.93	-0.15	-0.44	0.46	0.53	Particular	Tijuana
RN_02	0.91	-0.64	-0.51	0.76	1.67	Particular	Mexicali
RN_03	0.88	-0.98	-1.23	-0.16	0.13	General	Mexicali
RN_04	0.86	-1.86	-1.17	-0.08	-0.05	General	Mexicali
RN_05	0.86	-0.66	-0.61	-0.18	-0.64	Técnica	Ensenada
RN_06	0.83	-0.03	0.27	1.11	1.35	Particular	Tijuana
RN_07	0.81	-0.66	-0.81	-0.28	-0.53	Técnica	Tijuana
RN_08	0.79	-1.18	-1.10	-0.43	-1.06	General	Mexicali
RN_09	0.78	0.74	0.48	0.81	1.02	Particular	Ensenada
RN_10	0.78	-1.07	-1.08	-0.49	-0.56	General	Mexicali
RN_11	0.74	-0.79	-0.98	-0.39	-0.59	Técnica	Ensenada
RN_12	0.72	0.64	0.09	-0.42	0.84	Particular	Tijuana
RN_13	0.71	-0.14	0.00	-0.29	0.02	General	Ensenada
RN_14	0.71	-0.82	-0.77	-0.71	-1.42	General	Tecate
RN_15	0.70	0.47	0.78	1.02	-0.40	Particular	Ensenada

* Puntuaciones estandarizadas N (0.1) para cada modelo obtenido

En este caso, destaca la existencia de centros como el RB_02 o RB_06, de carácter particular, con un potencial extraordinario dado el índice socioeconómico y los recursos disponibles, pero que se mantienen como de baja eficacia. Otro centro como el RB_09 se muestra como de gran interés, ya que obtiene puntuaciones altas en ambas pruebas de ENLACE y posee tanto un índice socioeconómico como recursos altos. Al igual que en el análisis de las escuelas de residuo alto cobraban especial relevancia aquellas con mayores dificultades, en este caso resalta la importancia de llevar a cabo un estudio pormenorizado en aquellas escuelas que disponen de muchas facilidades para alcanzar altas cuotas en el rendimiento académico de sus estudiantes, pero que no las obtienen. En suma, ¿qué tipo de prácticas implementadas de manera habitual por este tipo de centros no están siendo eficaces en la mejora del rendimiento de los estudiantes?

En este caso, el reparto de centros por modalidad y municipio sigue estando bien distribuido, a excepción de la inexistencia de centros de telesecundaria incluidos.

IV. Conclusiones

El empleo de modelos lineales jerárquicos para la identificación de centros con valor añadido en educación se ha generalizado partiendo de los esfuerzos que las administraciones públicas nacionales llevan a cabo para la evaluación general del logro de sus estudiantes y con la ayuda de los cuestionarios de contexto. En este caso, debido a la naturaleza de los datos disponibles, no fue posible aplicar modelos

longitudinales, a los que en estricto sentido les correspondería la denominación de valor añadido siguiendo la recomendación de la OCDE (2008). Sin embargo, la aplicación del modelo transversal contextualizado sin ganancia permitió aplicar 8 modelos independientes, localizando los centros que se distribuyen de manera sistemática en residuos bajos y altos.

El procedimiento, que aísla el efecto de las variables contextuales significativas, junto con los 3 criterios de selección empleados, aporta robustez a la selección de los centros más y menos eficaces, dando credibilidad a todo el proceso. Todos los modelos aplicados resultaron con un número y tipo de covariables semejante, lo que también aporta seguridad al trabajo realizado en el estudio. Por otra parte, el criterio empleado de incluir como covariables en los modelos únicamente aquellas variables estrictamente contextuales, es decir, aquellas que se puede considerar que operan independientemente de las prácticas y estrategias adoptadas por los docentes y las escuelas, resulta ser una garantía de equidad para la selección e identificación de las escuelas de uno y otro tipo, en la medida en que así se detrae sólo el efecto de lo meramente contextual.

No se debe olvidar que el rendimiento académico es un fenómeno multicausal en el que intervienen infinidad de variables íntimamente relacionadas entre sí (Lee y Shute, 2010) que, de acuerdo con González et al. (2012), pueden agruparse en personales, familiares, escolares y sociales. Los resultados aquí mostrados identifican y aíslan el influjo de los factores contextuales significativos, tanto personales como familiares, escolares y sociales, asociados al rendimiento académico. A pesar de que la mera identificación de estos factores ya tiene asociado un interés evidente en la investigación educativa, la principal aportación perseguida tiene que ver con la identificación de los factores no contextuales que contribuyen al valor añadido localizado en las escuelas de alto rendimiento. Como ya se ha indicado, esta cuestión se aborda en las fases posteriores del proyecto.

En cuanto a los modelos, cabe destacar la importancia general de las expectativas de los sujetos con respecto a su futuro académico-laboral y la asistencia y puntualidad a las clases. Por otro lado, adquieren importancia el índice socioeconómico y la nacionalidad del estudiante en los modelos que incluían como variable criterio el rendimiento en Español, en contraste con el rendimiento en Matemáticas. Mientras que un estudiante motivado y que acude a todas las clases parece sinónimo de éxito académico general, un bajo índice socioeconómico y no tener el Español como lengua materna tiene más influencia negativa en el rendimiento en esta materia.

Si se analizan las características de los centros seleccionados con respecto a la modalidad y el municipio, cabe valorar las distribuciones de ambos grupos similares a las de la muestra general estudiada. Con la excepción en las escuelas de residuo negativo de la modalidad telesecundaria y del municipio de Playas de Rosarito en ambas selecciones, la distribución de municipios, modalidades, rendimiento, nivel socioeconómico y recursos es similar al localizado en la muestra original. Así, cabe afirmar que se están aislando realmente los efectos causados por las variables contextuales sobre el rendimiento académico, ya que se seleccionaron escuelas con condiciones similares a las de la muestra inicial.

A pesar de estas cuestiones generales, que aportan al estudio de validez y fiabilidad, es necesario destacar algunas particularidades que restan fortaleza al mismo. En primer lugar, destaca el hecho de la no inclusión del rendimiento previo como covariable en los modelos planteados. Este factor implica que se supone una distribución aleatoria de los estudiantes por centro, cuestión discutible, razón que aporta cierto sesgo a los resultados obtenidos. Por otro lado, no se debe olvidar que se han obtenido valores del ICC superiores al 10% en 7 de los 8 modelos finales probados, con varianzas explicadas pequeñas en el primer nivel. Así, a pesar de que la intención principal del estudio no ha consistido en obtener modelos impecables a nivel estadístico, cabe la posibilidad de que se hayan obviado algunas variables contextuales valiosas o que las variables introducidas cuenten con un error de medida importante.

Teniendo en cuenta estas fortalezas y debilidades, cabe señalar, en suma, que el presente estudio ha alcanzado los objetivos planteados inicialmente, y que se sitúa como punto de partida para futuros trabajos y estudios de caso más específicos para la detección de buenas prácticas educativas integrales implementadas por los centros educativos del estado de Baja California, en México, de cara a la mejora del rendimiento académico de sus estudiantes. La detección de estas prácticas educativas es una

información valiosa tanto para las autoridades educativas estatales como para los propios agentes educativos de las diferentes escuelas.

Referencias

Bryk, A. S. y Raudenbush, S. W. (1992). *Hierarchical linear models: applications and data analysis methods*. Estados Unidos: Sage.

Bryk, A. S., Thum, Y. M., Easton, J. Q. y Luppescu, S. (1998). Assessing school academic productivity: The case of Chicago School Reform. *Social Psychology of Education*, 2(1), 103-142. doi: 10.1023/A:1009657609112

Bryk, A. S. y Weisberg, H. I. (1976). Value-added Analysis: A dynamic approach to the estimation of treatment effects. *Journal of Educational Statistics*, 1(2), 127-155. doi:10.3102/10769986001002127

Caso, J., Chaparro, A., Díaz, C. y Urías, E. (2012). Propiedades psicométricas de las escalas, cuestionarios e inventarios de la Estrategia Evaluativa Integral 2011: Factores Asociados al Aprendizaje. UEE RT 12-002. México: Universidad Autónoma de Baja California. Recuperado de <http://uee.uabc.mx/uee/nivel1.php?p=estudiosYproyectos/reportesTecnicos>

Caso, J., Díaz, C., Chaparro, A. y Urías, E. (2011). Propiedades psicométricas de las escalas, cuestionarios e inventarios de la Estrategia Evaluativa Integral 2010: Factores Asociados al Aprendizaje. UEE RT 11-002. México: Universidad Autónoma de Baja California. Recuperado de <http://uee.uabc.mx/uee/nivel1.php?p=estudiosYproyectos/reportesTecnicos>

Dirección General de Evaluación de Políticas. (2010). ENLACE BÁSICA. Manual técnico 2010. México: Secretaría de Educación Pública. Recuperado de http://www.enlace.sep.gob.mx/ba/manuales_tecnicos/

Dirección General de Evaluación de Políticas. (2011). ENLACE BÁSICA. Manual técnico 2011. México: Secretaría de Educación Pública. Recuperado de http://www.enlace.sep.gob.mx/ba/manuales_tecnicos/

Drabowicz, T. (2014). Gender and digital usage inequality among adolescents: A comparative study of 39 countries. *Computers & Education*, 74, 98-111.

Educational Resources Information Center and National Science Foundation (1996). *Third International Mathematics and Science Study (TIMSS)*. Washington, DC: US Dept. of Education. Recuperado de <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED401127>

Fitz-Gibbon, C. T. (1997). The value-added national project: final report. Feasibility studies for a national system of value-added indicators. Londres: SCAA.

Gaviria, J. L., Martínez-Arias, R. y Castro, M. (2004). Un Estudio multinivel sobre los factores de eficacia escolar en países en desarrollo: el caso de los recursos en Brasil. *Education Policy Analysis Archives*, 12(20). Recuperado a partir de <http://epaa.asu.edu/ojs/index.php/epaa/article/download/175/301>

Goldstein, H. (1995). *Multilevel statistical models*. Londres: Oxford University Press.

Goldstein, H. y Spiegelhalter, D. J. (1996). League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society*, 159(3), 384-443.

González, C., Caso, J., Díaz-López, K. y López, M. (2012). Rendimiento académico y factores asociados: aportaciones de algunas evaluaciones a gran escala. *Bordón. Revista de pedagogía*, 64(2), 51-68. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3960787>

- Gritter, A. (2006). Great Schools: Identifying Higher-Performing Schools. Center for Education Policy, Applied Research, and Evaluation. Recuperado de http://www.usm.maine.edu/sites/default/files/Center%20for%20Education%20Policy,%20Applied%20Research,%20and%20Evaluation/Great_Schools.pdf
- Huffman, J. B., Pankake, A. y Munoz, A. (2006). The tri-level model in action: site, district, and state plans for school accountability in increasing student success. *Journal of School Leadership*, 16(5), 569-582.
- Jackson, S. A. y Lunenburg, F. C. (2010). School performance indicators, accountability ratings, and student achievement. *American Secondary Education*, 39(1), 27-44.
- Joaristi, L. M., Lizasoain, L. y Gamboa, E. (2012). Construcción y validación de un instrumento de medida del Nivel Socioeconómico y Cultural (NSE) de estudiantes de educación primaria y secundaria. *Bordón. Revista de pedagogía*, 64(2), 151-171. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3960926>
- Jornet, J. M., González, J. y Perales, M. J. (2012). Diseño de cuestionarios de contexto para la evaluación de sistemas educativos: optimización de la medida de constructos complejos. *Bordón. Revista de pedagogía*, 64(2), 89-110. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3960804>
- Kjaernsli, M. (2004). PISA and Scientific Literacy: Similarities and Differences between the Nordic Countries. *Scandinavian Journal of Educational Research*, 48(3), 271-286.
- Lee, J. y Shute, V.J. (2010). Personal and social-contextual factors in K-12 academic performance: An integrative perspective on student learning. *Educational Psychologist*, 45(3), 185-202.
doi: 0.1080/00461520.2010.493471
- Marchesi, A. y Martín, E. (2002) *La evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: SM.
- Martínez, R. (2009). Usos, aplicaciones y problemas de los modelos de valor añadido en educación. *Revista de Educación*, 348, 217-250. Recuperado de http://www.mecd.gob.es/revista-de-educacion/numeros-revista-educacion/numeros-anteriores/2009/re348/re348_10.html
- Martínez, F., Chaparro, A. A., Lizasoain, L., Caso, J. y Urías, E. (2014). Identificación de escuelas de alto de alto y bajo residuo, con base en puntajes de enlace y variables de contexto. UEE RT 14-003. México: Unidad de Evaluación Educativa. Recuperado de <http://uee.uabc.mx/uee/nivel1.php?p=estudiosYproyectos/reportesTecnicos>
- Meyer, R. H. (1996). Value-added indicators of school performance. En E. A. Hanushek y D. W. Jorgenson (Eds.), *Improving America's schools: the role of incentives* (pp. 197-223). Washington: National Academic Press.
- Mullis, I. V., Martin, M. O., Kennedy, A. M., Trong, K. L. y Sainsbury, M. (2009). *PIRLS 2011 Assessment Framework*. Amsterdam: International Association for the Evaluation of Educational Achievement. Recuperado de <http://timssandpirls.bc.edu/pirls2011/framework.html>
- Murillo, F. J. y Hernández-Castilla, R. (2011a). Efectos escolares de factores socio-afectivos. Un estudio Multinivel para Iberoamérica. *Revista de Investigación Educativa*, 29(2), 407-428.
- Murillo, F. J. y Hernández-Castilla, R. (2011b). Factores escolares asociados al desarrollo socio-afectivo en Iberoamérica. *RELIEVE*, 17(2), 1-23.

Murillo, F. J. y Román, M. (2011). ¿La escuela o la cuna? Evidencias sobre su aportación al rendimiento de los estudiantes de América Latina. Estudio multinivel sobre la estimación de los efectos escolares. *Revista de Curriculum y Formación de Profesorado*, 15(3), 27-53. Recuperado de <http://digiбуq.ugr.es/handle/10481/19169#.Vv6wWPnhC70>

OECD (2008). *Measuring improvements in learning outcomes: best practices to assess the value-added of schools*. París: Autor.

OECD (2012). *PISA 2009 technical report*. París: Autor. Recuperado de <http://www.oecd.org/pisa/pisaproducts/50036771.pdf>

Ray, A., McCormack, T. y Evans, H. (2009). Value added in english schools. *Education Finance and Policy*, 4(4), 415-438.

Sanders, W. L. y Horn, S. P. (1995). The Tennessee Value-Added Assessment System (TVAAS): mixed model methodology in educational assessment. En A. J. Shrinkfield y D. Stufflebeam (Eds.), *Teacher Evaluation: Guide to Effective Practice* (pp. 337-350). Boston, MA: Kluwer.

Tatto, M. T., Ingvarson, L., Schulle, J., Peck, R., Senk, S. L. y Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): policy, practice, and readiness to teach primary and secondary mathematics: conceptual framework*. Amsterdam: Michigan State University.

UNESCO (2005). *Segundo Estudio Regional Comparativo y Explicativo (SERCE): análisis curricular*. Santiago de Chile: Autor.

Webster, W. J. y Mendro, R. L. (1997). The Dallas value-added accountability system. En J. Millman (Ed.), *Grading teachers, grading schools: is student achievement a valid evaluation measure?* (pp. 81-99) Thousand Oaks, CA: Corwin.

Willms, J. D. y Raudenbush, S. W. (1989). A Longitudinal Hierarchical Linear Model for Estimating School Effects and Their Stability. *Journal of Educational Measurement*, 26(3), 209-232.
doi: 10.1111/j.1745-3984.1989.tb00329.x

Wright, R. J. (2009). Methods for improving test scores: the good, the bad, and the ugly. *Kappa Delta Pi Record*, 45(3), 116-121.

ANEXO I.

Variables contextuales incluidas en el estudio

N1_ISOC: Índice socioeconómico de cada estudiante.

N1_SEXO: Sexo del estudiante (0=femenino; 1=masculino).

N1_MEXICAN: Si el estudiante ha nacido en México o no (0=no mexicano; 1=mexicano).

N1_GUARDERÍA: ¿El estudiante ha asistido al jardín de niños? (0=no asiste; 1=sí, por un año o menos; 2=sí, por más de un año).

N1_REPITE: ¿Es repetidor el estudiante? (0=no asistió; 1=sí, una vez; 2=sí, dos veces; 3-4-5=sí, más de dos veces en varios cursos).

N1_EDADACC: Edad de acceso a la educación primaria (1=5 años o menos; 2=6 años; 3=7 años; 4=8 años o más).

N1_EXPECTATIVAS: Nivel educativo hasta el que le gustaría estudiar al estudiante (1=secundaria; 2=preparatoria o carrera técnica; 3=licenciatura; 4=maestría; 5=doctorado).

N1_FALTA.ALUMNO: Frecuencia en la no asistencia e impuntualidad del alumno.

N2_TURNO: Variable dicotómica, con una puntuación de 1 si el centro es matutino y de 0 si es vespertino, discontinuo o nocturno.

N2_MUNI_ 'municipio': Variables dicotómicas que reflejan si el centro pertenece a Ensenada, Mexicali, Playas de Rosarito, Tecate o Tijuana.

N2_MOD_ 'modalidad': Variables dicotómicas que reflejan si el centro es de modalidad particular o no.

N2_ISOC: Media de índice socioeconómico por escuela.

N2_SEXO: Promedio de estudiantes masculinos por escuela.

N2_MEXICAN: Promedio de estudiantes mexicanos por escuela.

N2_GUARDERIA: Media de tiempo asistiendo al jardín de infancia por escuela.

N2_REPITE: Media de veces que repiten los alumnos de una escuela.

N2_EDADACC: Edad de acceso media a la escuela.

N2_EXPECT: Expectativas medias de los estudiantes de la escuela.

N2_FALTA.ALUMNO: Media de falta de asistencia e impuntualidad del alumno por escuela.

N2_EXPERESC.PROF: Experiencia escolar media del profesorado de la escuela.

N2_ESTUD.PROF: Nivel de estudios medios por escuela del profesorado.

N2_EXPER.DIRE: Años de experiencia como director del director de la escuela.

N2_RECURSOS.ESC: Recursos educativos generales de los que dispone la escuela y estado de los mismos.

N2_DIASSUSP: Días que se suspendieron las clases en la escuela sin planificación previa.