

Clasificación de colegios según las pruebas Saber 11 del ICFES: un análisis usando modelos marginales (MM)¹

Schools classification according ICFES Saber 11 tests: An analysis using marginal models (MM)

Classificação de colégios segundo as provas Saber 11 do ICFES: uma análise com modelos marginais (MM)

Carlos Alberto Gómez Silva²

Universidad Nacional de Colombia
caragomez@unal.edu.co

Recibido: 29/06/2015

Aprobado: 24/11/2015

-
- 1 Este documento hace parte de la investigación realizada en la tesis para optar al título de Maestría en Ciencias Estadística de la Universidad Nacional de Colombia. Una versión de trabajo de este documento se puede encontrar en Archivos de Economía del Departamento Nacional de Planeación. <https://colaboracion.dnp.gov.co/CDT/Estudios%20Economicos/421.pdf>
 - 2 Economista. Especialista en Estadística.

Resumen

Utilizando la clasificación de los colegios según las pruebas Saber 11 del ICFES, se realiza un análisis longitudinal para el período 2001-2011. Esta clasificación es analizada usando los modelos marginales para respuestas ordinales, considerando variables explicativas como la naturaleza jurídica y la jornada de los colegios. Como resultados principales se encontró que los colegios no oficiales tienen, en promedio, las mejores clasificaciones, en contraposición con los colegios oficiales; esto evidencia un aumento en la brecha en la clasificación entre estos dos tipos de establecimiento en el período 2001-2011 para las trece principales ciudades de Colombia. No se encontraron efectos de la jornada completa sobre la clasificación de los establecimientos educativos.

Palabras clave: datos ordinales; modelos marginales; datos categóricos longitudinales; evaluación de la educación.

Códigos JEL: I21; C14.

Abstract

Using the school classification according to tests ICFES Saber 11, a longitudinal analysis was performed for the period 2001 – 2011. This classification was analyzed using marginal models for ordinal responses, and considering explanatory variables as the school legal nature and the timetables. The main results found that the private schools have, on average, better ratings, as opposed to public schools; this shows an increase in the gap in the standings between these two types of establishments in the period 2001-2011 for the thirteen main cities in Colombia. No full-time effects on the classification of schools were found.

Keywords: ordinal data; marginal models; longitudinal categorical data; education evaluation.

Resumo

Com base na classificação dos colégios segundo as provas Saber 11 do ICFES, se realiza uma análise longitudinal para o período 2001-2011. Essa classificação é analisada usando os modelos marginais para respostas ordinais, tendo em vista variáveis explicativas como a natureza jurídica e a jornada dos colégios. Dentro dos resultados principais, se encontrou que os colégios não oficiais têm, em média, as melhores classificações, em contraposição com os colégios oficiais; isto demonstra um aumento na brecha na classificação entre estes dois tipos de estabelecimento no período 2001-2011 para as treze principais cidades da Colombia. Não se encontraram efeitos da jornada completa sobre a classificação dos estabelecimentos educativos.



Este trabajo está bajo la licencia Creative Commons Attribution 3.0

¿Cómo citar este artículo? / How to quote this article?

Gómez, Carlos. «Clasificación de colegios según las Pruebas Saber 11 del ICFES: un análisis usando modelos marginales (MM)». *Sociedad y economía*, No. 30 (enero - junio 2016): 69-89.

Palavras-chave: dados ordinais; modelos marginais; dados categóricos longitudinais; avaliação da educação.

1. Introducción

Evaluar la calidad de la educación es importante, puesto que esta tiene efectos tanto en el ejercicio de la ciudadanía como en la productividad de las personas. Una educación secundaria de mala calidad afecta las posibilidades de ingresar a la educación superior, e incide en la deserción académica y en la inserción exitosa en el mercado laboral. En este sentido, para potenciar la movilidad social es importante que todas las personas puedan acceder a una educación de calidad (Barrera-Osorio y Maldonado 2012; Rangel 2006).

En Colombia, el Instituto Colombiano para la Evaluación de la Educación (ICFES) elabora las pruebas estandarizadas para la evaluación de los estudiantes en diferentes niveles educativos³. Estas pruebas tienen como objetivo identificar el nivel de conocimientos de los estudiantes en diferentes áreas, así como generar elementos de juicio para la evaluación del desempeño de las instituciones educativas. De las pruebas que realiza el ICFES, las pruebas Saber 11 son las que tienen más tradición dado que se vienen realizando desde 1968⁴ y son utilizadas por las instituciones de educación superior como una herramienta para la selección de los estudiantes; de esta manera, los resultados que un estudiante obtiene en estas pruebas son relevantes para la movilidad académica, además de ser un mecanismo de autoevaluación de las instituciones educativas (Iregui, Melo y Ramos 2006).

A partir de estas pruebas, el ICFES clasifica las instituciones educativas a partir de una categorización de las mismas mediante siete categorías de rendimiento, que van desde “muy inferior” hasta “muy superior”⁵, y la cual es determinada a partir de los puntajes obtenidos por los estudiantes en las áreas del núcleo común: lenguaje, matemáticas, biología, física, química, ciencias sociales, filosofía e inglés. Esta clasificación tiene en cuenta los puntajes de los estudiantes y la homogeneidad de estos dentro de una misma institución, y es realizada a través de una normalización de los puntajes de cada área, teniendo como referencia la media y la desviación estándar de los resultados nacionales de la aplicación del examen, en cada una de las áreas evaluadas en el núcleo común.

Esta clasificación realiza una comparación relativa de las instituciones educativas al establecer una categoría para cada establecimiento educativo, por cada

3 Actualmente el ICFES cuenta con las pruebas SABER 5, SABER 9, SABER 11 y SABER PRO, las cuales tienen como propósito la evaluación de los estudiantes de quinto, noveno, once y la educación superior, respectivamente.

4 Consultado en octubre de 2014 en <http://www.icfes.gov.co>. Es necesario tener en cuenta que desde 1968, el examen realizado por el ICFES ha experimentado diferentes cambios en la metodología; actualmente se usa la establecida desde el 2000, aunque con algunas variaciones entre años.

5 En la resolución 489 de octubre de 2009 está descrita la metodología para la clasificación de instituciones y se definen las siguientes categorías de rendimiento: muy inferior, inferior, bajo, medio, alto, superior y muy superior.

uno de los años en que la institución inscribió estudiantes para presentar las pruebas. Se debe resaltar que el ICFES ha realizado ajustes metodológicos en las pruebas, de manera que en el período 2001-2011 se tienen tres diferentes metodologías: 2001-2006, 2007-2010 y 2011.

Usando esta clasificación se han definido diferentes políticas públicas que buscan generar incentivos para mejorar la calidad de las instituciones educativas. Por ejemplo, a partir de 2011, mediante el decreto 1055 de 2011 se estableció un sistema de incentivos para los rectores de los establecimientos educativos oficiales, clasificados de muy inferior a alto, que mejoren su clasificación en relación con el año inmediatamente anterior, y los colegios clasificados en superior y muy superior que mantengan o mejoren aquella categoría; estos incentivos se mantienen hasta 2014, como lo muestra el decreto 171 de 2014. Además, desde el Ministerio de Educación Nacional -MEN- se han propuesto mecanismos de aumentos de matrícula diferenciales, de acuerdo con los resultados de los establecimientos en las pruebas de Estado (Barrera-Osorio y Maldonado 2012).

Esta clasificación ha sido considerada como un instrumento de política pública, por lo que analizar su comportamiento en el tiempo es un elemento relevante para el análisis de la evaluación de la educación, que evidencia tendencias en las brechas educativas, según las características generales de los establecimientos educativos. Al respecto, es preciso considerar las dificultades de comparación temporal de las pruebas Saber dado los múltiples cambios metodológicos señalados en estas pruebas.

Teniendo en cuenta estos elementos, el presente trabajo realiza una aproximación longitudinal a la evaluación de la educación en Colombia bajo la consideración que la clasificación de las instituciones educativas se realiza año a año, lo cual constituye una base de comparación en el tiempo.

Así, se presenta una propuesta metodológica para evaluar la influencia de diferentes variables administrativas de los colegios en relación con la clasificación de los establecimientos escolares en las pruebas Saber 11 para el período 2001-2011 en las trece principales ciudades de Colombia⁶; se resalta que esto no se puede equiparar a una comparación sobre el desempeño del establecimiento educativo sino a la probabilidad de ser clasificado por el ICFES en una categoría dada a través del tiempo. Este análisis se realiza teniendo como soporte teórico los modelos marginales, los cuales permiten modelar la media de la población, incluyendo las posibles asociaciones de las diferentes observaciones de un mismo plantel educativo.

Este trabajo consta de seis secciones incluyendo esta introducción. En la segunda sección se realiza una revisión de los estudios de la educación en Colombia y enseguida se presenta una descripción de la información disponible. En la cuarta sección se realiza una descripción de los modelos marginales para respuestas ordinales, los cuales son la base para el análisis de la clasificación de los colegios. La aplicación de esta metodología se presenta en la quinta sección, la cual muestra el análisis para el caso de la independencia de las observaciones

6 Las principales ciudades consideradas en este documento son: Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales, Pasto, Pereira, Cúcuta, Ibagué, Montería, Cartagena y Villavicencio.

y se contrasta cuando se considera la dependencia de las observaciones. En la última sección se concluye.

2. Estudios de evaluación de la educación: una revisión de literatura

El análisis de la calidad de la educación ha sido una temática recurrente dentro de los estudios de capital humano. Dentro de las diferentes aproximaciones de la calidad, se tienen los análisis que evalúan la calidad en términos de los insumos (número de estudiantes por profesor, salarios de los profesores, gasto educativo, entre otras) y los que miden la calidad en términos de los resultados en pruebas estandarizadas (Thomas, Wang y Fan 2001). Dentro del primer enfoque, que considera la calidad educativa según los insumos de la educación, se tienen estudios con perspectiva económica que analizan los efectos de la calidad sobre la rentabilidad de la educación; por ejemplo, Card y Krueger (1992) encuentran una relación positiva entre los recursos de los establecimientos educativos y los retornos en los ingresos salariales para los hombres en Estados Unidos en la primera mitad del siglo XX.

Los estudios que utilizan el concepto de calidad por resultados tienen diferentes análisis que utilizan las pruebas estandarizadas internacionales como las pruebas realizadas por el Programa Internacional para la Evaluación de los Estudiantes (Programme for International Student Assessment -PISA-) y el estudio de Tendencias en el Estudio Internacional de Matemáticas y Ciencias (Trends in International Mathematics and Science Study -TIMSS-). Este enfoque busca establecer relaciones entre variables de los individuos y los establecimientos educativos, para explicar el desempeño en las pruebas estandarizadas de evaluación de educación.

Para el caso colombiano se encuentran trabajos que analizan la calidad de la educación a partir del uso de las pruebas Saber. En estos estudios, con diferentes enfoques y metodologías estadísticas, se evalúa el impacto de las características de los estudiantes y de los establecimientos educativos. En algunos casos se analizan las diferencias en el desempeño entre los establecimientos oficiales y los no oficiales, y en otros se muestra que las diferencias en el desempeño están relacionadas en una proporción importante por el contexto de los estudiantes (ingreso de sus familias, educación de los padres, entre otras variables).

Por ejemplo, Bonilla (2011) evalúa los efectos de la jornada escolar sobre la calidad de la educación en Colombia. A partir del uso de variables instrumentales y de las pruebas Saber 2009, el autor concluye que estudiar en jornada completa tiene impactos positivos sobre los resultados académicos, especialmente cuando se compara con los estudiantes de la jornada de la tarde. En este estudio se encontró que, en promedio, registran mejor desempeño en las pruebas Saber los hombres, los habitantes de las cabeceras municipales, los estudiantes que no pertenecen a minorías étnicas y los que no trabajan. También evidencia que los resultados mejoran con el nivel educativo de la madre, el aumento del ingreso familiar y el valor de la pensión. Igualmente, identifica que tienen mejores

resultados, en promedio, los estudiantes de instituciones no oficiales, luego de controlar por las demás variables mencionadas.

El trabajo de Bonilla y Galvis (2011) usa variables instrumentales para analizar los impactos de la profesionalización docente sobre el desempeño de los estudiantes en las pruebas Saber 11; la variable instrumental que usan es el número de estudiantes de la institución, como una aproximación del tamaño del colegio. Según los autores, todas las regresiones que realizan tienen corrección por efecto clúster en cuanto a sede/jornada y determinan que no hacen uso de modelos multinivel debido a que los resultados de estos modelos no difieren mucho de los modelos de mínimos cuadrados; no obstante, no se comprueba esta afirmación. Dentro de los resultados de interés, concluyen que la profesionalización docente tiene un impacto positivo sobre el desempeño de los estudiantes, en especial en el área de matemáticas. Sin embargo, al diferenciar los resultados por tipo de plantel, encuentran que estos efectos son menores en los colegios oficiales que en los no oficiales.

El trabajo de Castaño (1998) usa modelos jerárquicos para determinar el porcentaje de variabilidad del rendimiento de los estudiantes que es explicado por el “efecto colegio”. A partir del análisis de una estructura de dos niveles, analiza los resultados en matemáticas de las pruebas Saber 3, aplicada a los estudiantes de tercero de primaria. Según estos resultados, aproximadamente 29% de la variabilidad del rendimiento de los estudiantes es explicada por el colegio.

Un análisis más completo que el anterior, y que tiene en cuenta la estructura multinivel de los datos, es el realizado en el trabajo de Duarte, Moreno y Bos (2012). En ese documento se describe la desigualdad de los aprendizajes a partir del análisis de las pruebas Saber 5 y 9 del 2009, asociándolas al nivel socioeconómico de las familias, y al tipo de gestión y ubicación geográfica de las instituciones. Como principales resultados se encuentra que la correlación en el desempeño es más importante entre las escuelas que dentro de las mismas, lo que implica una alta segregación de los alumnos, según el nivel socioeconómico. Se muestra que solo una tercera parte de los estudiantes de quinto y dos de cada cinco en noveno tienen resultados satisfactorios o avanzados de lectura, lo que evidencia un bajo logro educativo en el país.

Según estos autores, en el ámbito nacional 37,5% de la varianza de los resultados de las pruebas Saber es atribuible a las características de las escuelas, el resto (62,5%), a características de los estudiantes; se señala que un aumento de una desviación estándar en el índice socioeconómico calculado por ellos implica una variación de 8 puntos en lectura de grado quinto.

Por su parte, el trabajo de Iregui *et al.* (2006) propone una función de producción del sistema educativo utilizando técnicas de frontera estocástica, verificando los factores que afectan el rendimiento de los estudiantes a partir de las variables de los colegios y las características socioeconómicas de los estudiantes. Esta técnica permite evaluar la eficiencia de cada uno de los colegios en la muestra, midiendo el potencial de mejorar el desempeño de acuerdo con algunas características de cada establecimiento educativo. Haciendo uso de las pruebas Saber de 2002, los autores encuentran una relación positiva entre la educación de los maestros, las variables de infraestructura de los colegios y las variables ambientales de los estudiantes (el ingreso medio de los hogares, la ubicación del

colegio, el bachillerato académico, la jornada). Los autores encuentran que los establecimientos oficiales presentan, en promedio, un menor logro académico que los no oficiales; sin embargo, encuentran evidencia que los colegios no oficiales se benefician de las condiciones del entorno más favorables, teniendo en cuenta que asisten estudiantes de familias con mayores ingresos.

Otros tipos de estudios se enfocan en tratar de medir el desempeño futuro en la educación superior, según los resultados de los estudiantes en las pruebas de la educación media. Barón y Bonilla (2011), a partir de la estimación de un modelo probit, encuentran una relación inversa entre los resultados de las pruebas Saber 11 y la probabilidad de graduarse de un programa de pregrado en educación; según los resultados, definen que la probabilidad de graduarse en un pregrado en educación es cinco veces más alta cuando un estudiante obtuvo un resultado entre los más bajos que cuando se obtuvo un resultado entre los más altos, identificando mayores diferencias para las mujeres. En este trabajo usan como medida de análisis el percentil promedio en las pruebas del ICFES de los graduados de educación y lo comparan con otras áreas; esta metodología la usan para comparar a personas que pertenecen a diferentes cohortes. En el documento se coincide en que las características de las familias y del estudiante explican más la variación del rendimiento académico de los estudiantes que las características de los colegios; resultados similares a los de la literatura internacional.

Si bien esta literatura consultada evidencia la importancia de las características personales de cada estudiante, se resalta que una parte importante de la variabilidad de los resultados está relacionada con las características de los establecimientos educativos. No obstante, en esa literatura no se realiza una evaluación longitudinal de las características de los colegios en los resultados de las pruebas Saber 11, lo cual ha restringido la identificación de las brechas entre tipos de establecimientos a través del tiempo. Esta falta de comparación temporal en el desempeño de los planteles educativos se debe en parte a los múltiples cambios metodológicos de las pruebas Saber.

Como una alternativa de análisis, este trabajo parte de la consideración de que la clasificación de las instituciones se realiza año a año, y dado que esta clasificación es de tipo ordinal se puede realizar una comparación relativa de las instituciones educativas. Teniendo en cuenta esto, se realiza una comparación de las distribuciones condicionales de la probabilidad que tiene un colegio de estar clasificado en una categoría dada, para las diferentes variables características de los establecimientos educativos y un período de tiempo dado.

A partir de estos elementos, es posible realizar un seguimiento a la clasificación de las instituciones educativas entre 2001 y 2011, teniendo como marco de referencia las diferentes variables administrativas asociadas a la institución educativa: la naturaleza jurídica (oficial, no oficial) y la jornada (continua, mañana, tarde). Adicionalmente, como una aproximación al enfoque educativo de la institución, se considera la presencia de las palabras “técnico”, “técnica”, “tecnológico” y “tecnológica” en el nombre de la institución, como una aproximación al carácter técnico o académico del establecimiento.

Así, dentro de este marco se tienen las siguientes preguntas que pretenden ser respondidas a la luz de la información disponible.

- ¿La naturaleza del colegio (oficial – no oficial) influye en la clasificación del colegio? Si es así, ¿este efecto ha tenido cambios entre 2001 y 2011?
- ¿Existe una relación entre la jornada del colegio y la clasificación que este obtiene en las pruebas Saber 11? ¿Cómo ha variado esta relación en el tiempo?
- ¿Existe alguna interacción en los efectos de la jornada y la naturaleza del colegio?

Estas preguntas son relevantes dado que están relacionadas con los efectos de características administrativas de los colegios sobre su clasificación en las pruebas Saber 11. Al respecto, se debe reconocer que el análisis de los colegios implica el riesgo de la *falacia ecológica*, al considerar que el contexto de los individuos, en este caso los colegios, define su desempeño, por lo que estos resultados deben tomarse con precaución (Courgeau 2007). Sin embargo, existe evidencia que el efecto de los colegios es muy importante en el desempeño de los estudiantes; además, algunos efectos del contexto familiar, como es el nivel educativo de los padres, se transmite mediante la elección del plantel educativo (Gaviria y Barrientos 2001). De esta manera, hacer una revisión sobre cómo ha sido la transición de la clasificación de los colegios en el tiempo es relevante en el conocimiento de las dinámicas educativas.

3. Clasificación de los establecimientos educativos en las pruebas Saber 11

En esta sección se describe la información de la clasificación de los establecimientos educativos de las jornadas completa, mañana y tarde para las principales trece ciudades de Colombia en el período 2001-2011. Siguiendo a Bonilla (2011) y Bonilla y Galvis (2011), no se consideran las instituciones educativas con jornada nocturna o sabatina; dado que la mayoría de los estudiantes en estas jornadas son adultos, se tienen amplias diferencias con la población de los establecimientos de las demás jornadas.

En la tabla 1 se presenta la clasificación de los colegios en las pruebas Saber 11 para el período de análisis. En esta tabla se verifica un crecimiento en el número de establecimientos educativos que registraron estudiantes en las pruebas, al pasar de 2.683 en 2001 a 3.403 en 2010. Para este análisis se unificó la categoría *inferior* con la *muy inferior* debido a las bajas frecuencias que registró esta última categoría para algunos años (menos de cinco establecimientos).

De acuerdo con la distribución porcentual de la tabla 1, la participación de los colegios clasificados en la categoría *muy superior* se incrementó durante el período: en 2001 la participación de estos colegios fue 2,91% y en 2011, de 17,36%, lo que indica un incremento de 14,45 puntos porcentuales (p.p.). De igual manera, la participación de los colegios clasificados en la categoría *superior* tuvo un incremento considerable: en 2001 su participación fue de 9,65%, con un aumento de 12,31 p.p., para tener una participación de 21,61% en 2011.

Como un comportamiento derivado, los colegios clasificados en la categoría *inferior* tuvieron una reducción en su participación en el período: en 2001 esta fue 5,9%, mientras que en 2011 fue 0,9%. De igual manera, en 2001 el porcentaje de colegios clasificados en desempeño *bajo* fue 30,5%, mientras que en 2011 fue

Tabla 1. Número de colegios, según la clasificación en las pruebas Saber 11, 2001-2011

Frecuencia											
Clasificación del colegio	Año										
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Inferior	154	131	215	49	123	26	137	185	131	153	24
Bajo	840	783	742	382	629	322	796	752	659	631	206
Medio	932	929	958	1053	1027	907	1045	1064	1000	917	773
Alto	420	524	541	857	596	829	517	572	690	697	762
Superior	259	321	330	463	397	594	423	417	462	544	625
Muy superior	78	96	106	176	225	421	290	302	347	461	502
Total general	2.683	2.784	2.892	2.980	2.997	3.099	3.208	3.292	3.289	3.403	2.892
Distribución porcentual											
Clasificación del colegio	Año										
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Inferior	5,74	4,71	7,43	1,64	4,10	0,84	4,27	5,62	3,98	4,50	0,83
Bajo	31,31	28,13	25,66	12,82	20,99	10,39	24,81	22,84	20,04	18,54	7,12
Medio	34,74	33,37	33,13	35,34	34,27	29,27	32,57	32,32	30,40	26,95	26,73
Alto	15,65	18,82	18,71	28,76	19,89	26,75	16,12	17,38	20,98	20,48	26,35
Superior	9,65	11,53	11,41	15,54	13,25	19,17	13,19	12,67	14,05	15,99	21,61
Muy superior	2,91	3,45	3,67	5,91	7,51	13,59	9,04	9,17	10,55	13,55	17,36
Total general	100	100	100	100	100	100	100	100	100	100	100

Fuente: ICFES. Cálculos propios a partir de las bases de datos del ICFES

7,3%. Por su parte, los colegios clasificados en desempeño *medio* también tuvieron un comportamiento de reducción de su participación, con 35,3% y 25,2%, para 2001 y 2011, respectivamente.

En la tabla 2 se muestra la distribución de los colegios según diferentes características administrativas. Como se observa, hay un incremento en la proporción de colegios oficiales en el período de análisis: en 2001, eran oficiales 35,2% de los colegios, mientras que en 2011 representaron 46,4%, lo que evidenció un

Tabla 2. Distribución porcentual de colegios de jornadas mañana, tarde o completa, según naturaleza, enfoque y jornada, 2001-2011

Características		Año										
		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Naturaleza	No oficial	64,8	63,6	62,5	60,4	57,7	56,7	56,3	54,5	53,1	53,1	53,6
	Oficial	35,2	36,4	37,5	39,6	42,3	43,3	43,7	45,5	46,9	46,9	46,4
Enfoque	Académico	94,2	93,7	93,9	94,3	94,1	94,2	94,5	94,7	94,7	94,6	95,2
	Técnico	5,8	6,3	6,1	5,7	5,9	5,8	5,5	5,3	5,3	5,4	4,8
Jornada	Mañana	48,9	49,0	49,1	49,1	48,5	48,2	47,2	44,4	44,1	44,1	43,6
	Tarde	26,4	25,8	25,3	24,3	23,8	23,3	22,4	21,5	21,8	21,5	19,9
	Completa	24,7	25,2	25,6	26,5	27,7	28,5	30,5	34,1	34,1	34,3	36,5

Fuente: ICFES. Cálculos propios a partir de las bases de datos del ICFES

incremento de 11,2 p.p. en su participación. Por su parte, la proporción de los colegios de carácter “técnico”, según la definición adoptada en este trabajo, tuvo una reducción en su participación al pasar de 5,8% en 2001 a 4,8% en 2011.

La jornada que tuvo la mayor participación fue la de la mañana con 49,1% en 2003 para finalizar el período con una participación de 43,6% en 2011. La segunda jornada con más participación fue la completa, con un comportamiento creciente en su importancia relativa (en 2001 su participación fue 24,7% y en 2011 de 36,5%).

Un elemento importante para el análisis del desempeño de los establecimientos educativos en los resultados de las pruebas Saber 11 es la relación entre la naturaleza del colegio y la jornada. En la tabla 3 se observa que la proporción de colegios oficiales de jornada completa se incrementó en el período analizado; en 2001, en las trece principales ciudades 11,1% de los planteles oficiales era de jornada completa, mientras que en 2011 lo era 15,9%; no obstante, en los planteles oficiales la jornada mañana fue la que tuvo la mayor participación. Para los planteles educativos no oficiales se presentó la misma tendencia, pero con un

Tabla 3. Distribución porcentual de planteles educativos según naturaleza y jornada, 2001-2011

Naturaleza /Jornada	Año										
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
No oficiales											
Mañana	52,7	53,0	52,8	51,7	50,0	49,1	47,5	42,1	41,2	41,2	39,3
Tarde	15,2	14,0	13,1	11,9	10,6	10,4	9,6	8,1	8,0	7,6	6,4
Completa	32,1	33,0	34,1	36,4	39,4	40,6	42,9	49,8	50,7	51,2	54,3
Oficiales											
Mañana	41,9	42,0	43,0	45,2	46,5	47,1	46,8	47,3	47,4	47,5	48,5
Tarde	46,9	46,5	45,6	43,3	41,8	40,3	38,8	37,5	37,5	37,3	35,5
Completa	11,1	11,5	11,4	11,5	11,7	12,6	14,5	15,2	15,1	15,2	15,9

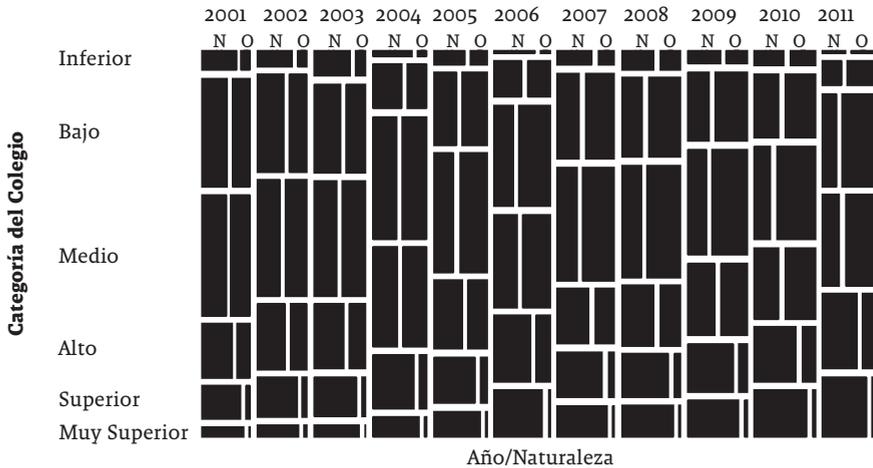
Fuente: ICFES. Cálculos propios a partir de las bases de datos del ICFES.

incremento más marcado: en 2001 la proporción de planteles no oficiales de jornada completa fue 32,1% mientras que en 2011 fue 54,3%, siendo esta la jornada más frecuente en este tipo de colegios.

En relación con el número de estudiantes que presentaron las pruebas Saber 11 por cada plantel, en el período de interés, el 25% de los colegios con menos estudiantes tuvieron 22 o menos estudiantes que presentaron las Pruebas Saber 11 para todos los años. Por otro lado, el 25% de los colegios con más estudiantes tuvieron 78 estudiantes o más inscritos que presentaron las pruebas Saber 11. Por la naturaleza del colegio, a través de una prueba de diferencia de medias, se concluye que el número medio de estudiantes de los colegios oficiales es mayor que el de los colegios no oficiales, con un índice de confianza de 95% (para los primeros, el número medio de estudiantes es 82,5 estudiantes, en tanto que los colegios oficiales es 45,4 estudiantes), lo cual evidencia que los colegios oficiales inscriben, en promedio, más estudiantes que los establecimientos no oficiales.

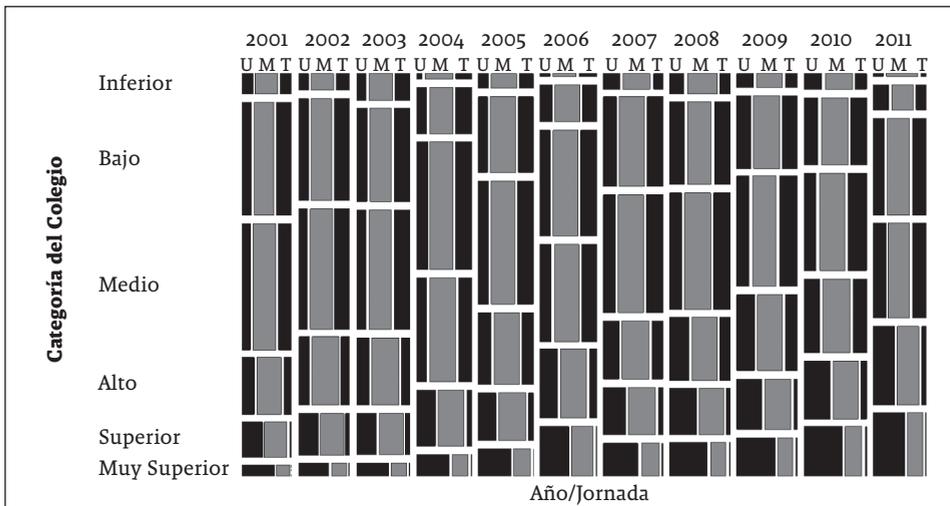
En cuanto a las relaciones existentes entre las variables administrativas de los establecimientos educativos y su clasificación, en la gráfica 1 de mosaicos se puede observar un crecimiento en la participación de los colegios en las categorías superior y muy superior durante el período 2001–2006, y luego una reducción de la participación en el 2007. A partir de ese año se tiene nuevamente un

Gráfica 1. Mosaico de la categoría del colegio, el año y la naturaleza del colegio, 2001-2011



Fuente: ICFES. Cálculos propios a partir de las bases de dato del ICFES.

Gráfica 2. Mosaico de la categoría del colegio, el año y la jornada del colegio, 2001-2011



Fuente: ICFES. Cálculos propios a partir de las bases de dato del ICFES.

incremento constante de la participación hasta el 2011. También se pueden ver que las participaciones de los colegios no oficiales y de jornada completa son mayores a medida que aumenta la categoría del colegio (gráfica 2).

De acuerdo con este panorama, a continuación se hace una breve descripción de los modelos marginales para datos ordinales, metodología que permite hacer un análisis de la clasificación de los colegios a lo largo del tiempo.

4. Metodología de análisis

Dado que la variable clasificación de los colegios en las pruebas Saber 11 es de tipo ordinal, es preciso utilizar métodos que permitan modelar esta estructura, teniendo en cuenta que para un mismo establecimiento existe más de una observación. Esta estructura tiene la característica de no tener independencia entre las observaciones, por lo que el uso de modelos que usan este supuesto, como los modelos lineales generalizados o los modelos log-lineales, no es adecuado debido a la existencia de patrones de dependencia o asociaciones en los datos (Ziegler, Kastner y Blettner 1998). Cuando todas las covariables son de tipo categórico, este análisis se puede abordar como múltiples tablas de contingencia, de manera que se estiman las probabilidades asociadas a las frecuencias de las tablas marginalizadas.

Para el análisis de datos categóricos resulta útil el método de estimación a través de las ecuaciones de estimación generalizadas (EEG), propuestas por Liang y Zeger (1986) y que son una extensión multivariada de los enfoques de cuasi verosimilitud, dado que incluyen la posibilidad de asumir estructuras de dependencia entre las diferentes observaciones de una unidad muestral. Las EEG permiten la estimación de modelos marginales (MM), los cuales tienen como punto central la estimación de la estructura de la media marginal, tomando la dependencia entre las observaciones como un parámetro molesto (*nuisance parameter*). Con estos modelos se estiman parámetros que determinan las relaciones promedio de la población, usando supuestos sobre las asociaciones entre las observaciones.

Estas ecuaciones realizan una estimación separada de los parámetros relacionados con la media poblacional y los parámetros de asociación de las observaciones, sin embargo, las EEG permiten obtener parámetros consistentes a pesar de especificaciones erróneas de la estructura de correlación (Touloumis, Agresti y Kateri 2013). Además, el uso de EEG permite analizar observaciones con diferentes tamaños de muestra o con valores perdidos, lo cual no es posible hacer con metodologías como mínimos cuadrados ponderados (Stokes, Davis y Koch 2012.)

Para ilustrar la estimación mediante EEG, considere Y_{it} la respuesta multinomial del sujeto $i=1, \dots, N$ en el momento $t=1, \dots, T_i$. Es posible definir una variable $Y_{itk}=I(Y_{it}=k)$ para $k=1, \dots, J$, siendo $I(A)$ una función indicador del evento A , de manera que \mathbf{Y}_{it} es equivalente al vector de tamaño $(I-1)$, $\mathbf{Y}_i=(\mathbf{Y}'_{i1}, \dots, \mathbf{Y}'_{i(T_i-1)})'$. Junto con esta variable de respuesta, considere $\mathbf{x}_i=(x'_{i1}, \dots, x'_{iT_i})'$ la matriz de covariables de tamaño $T_i(I-1) \times p$, que incluye el intercepto.

Las probabilidades marginales de obtener la respuesta $Y_{it}=k$ dadas las covariables \mathbf{x}_i es $\pi_{itk}=E(Y_{itk} | \mathbf{x}_i) = Pr(Y_{itk}=1 | \mathbf{x}_i)$, siendo $\pi_{it}=(\pi_{it1}, \dots, \pi_{it(I-1)})$. Este enfoque, al igual que los modelos lineales generalizados, relaciona la esperanza $E(Y_{itk})$ con las covariables a través de una función de enlace, es decir,

$$g(\pi_{it}) = x_{it} \beta.$$

Como punto fundamental para desarrollar este método de análisis de datos ordinales, la función de enlace $g(\pi_{it})$ establece una relación lineal entre la media de la variable dependiente y las variables explicativas. La principal aproximación a los datos ordinales se hace a través de modelos de *odds* proporcionales, las cuales definen una relación multiplicativa de las covariables sobre la relación de *odds* de las probabilidades acumuladas de una variable. En esta estructura, se asume que la relación de *odds* marginal está definido por

Ecuación 1 $\text{logit}[E(Y_{ijc} \leq c)] = \alpha_c - x_{ij}^t \beta.$

Este modelo puede ser visto como una regresión logística de la dicotomización de la variable, que tiene como objetivo modelar las asociaciones entre las observaciones (Stokes, Davis y Koch 2012; Agresti y Natarajan 2001; Christensen 2011), por lo que el modelo se puede representar como

Ecuación 2
$$\text{logit}[Y_{ijc} \leq c] = \log \frac{\pi(Y_{ijk} \leq k)}{\pi(Y_{ijk} > k)} = \alpha_c - \beta x$$

Siendo $\pi(\cdot)$ la probabilidad de que la variable respuesta esté en alguna de las categorías determinadas. Esta aproximación determina un efecto constante para cada una de las variables en cualquier categoría de la variable dependiente, es decir, asume el mismo efecto (β) para cualquier punto de corte (definido por las constantes de modelo β_c). Para el caso analizado en este documento, se tiene el mismo efecto para una variable explicativa a través de todas sus categorías.

Con el procedimiento de ecuaciones de estimación generalizadas (EEG) la estimación de los parámetros β está dada por la solución de las ecuaciones de estimación dadas por:

Ecuación 3
$$U(\beta, \alpha) = \frac{1}{N} \sum_{i=1}^K D_i V_i^{-1} (Y_i - \pi_i) = 0$$

Donde $D_i = \delta \pi_i / \delta \beta$, y la matriz de pesos determinada por $V_i = V_i(\beta, \alpha)$ de tamaño $T_i (I-1) \times T_i (I-1)$ es la matriz de covarianzas de trabajo, la cual depende de un valor α que depende a su vez de los valores de β (Touloumis, Agresti y Kateri 2013). Estas ecuaciones permiten obtener estimadores consistentes iteradamente, de manera que los estimadores de β se usan para la estimación de $\hat{\alpha}$, y viceversa.

Un estimador de la matriz de covarianzas del estimador $\hat{\beta}$ es el llamado estimador “sándwich”, el cual está determinado por

Ecuación 4
$$\Sigma_g = \lim_{N \rightarrow \infty} \left[\sum_{i=1}^N D_i' V_i^{-1} D_i \right]^{-1} \left[\sum_{i=1}^N D_i' V_i^{-1} \text{cov}(Y_i) V_i^{-1} D_i \right] \left[\sum_{i=1}^N D_i' V_i^{-1} D_i \right]^{-1}$$

Siendo la $\text{cov}(Y_i) = \text{cov}(Y_i | x_i)$ la matriz de covarianzas de tamaño $T_i (I-1) \times T_i (I-1)$ la verdadera función de covarianza para el sujeto i .

Para poder solucionar las ecuaciones de estimación, es preciso definir el parámetro α , que está relacionado con la matriz de covarianzas de “trabajo” V_i . Para

definir la asociación entre las observaciones se considera el uso de relaciones de *odds* locales, los cuales son usados para el análisis de asociaciones en tablas de contingencia mediante la estimación de modelos marginales a partir de EEG, según la metodología presentada en Touloumis *et al.* (2013).

Para esto, se considera $L = \frac{T(T-1)}{2}$ el número de pares de tiempo, siendo $T = \max\{T_1, \dots, T_N\}$. Por cada pareja de tiempos se forman L conjuntos de tablas de contingencia marginalizadas de tamaño $I \times I$. A partir de estas tablas, se define

$$\alpha = (\theta_{1121}, \dots, \theta_{112(I-1)}, \dots, \theta_{(T-1)1T1}, \dots, \theta_{(T-1)(I-1)(T-1)})$$

Donde $\theta_{jj't'}$ es la relación de *odds* local en el punto (j, j') en la tabla de contingencia marginalizada correspondiente a la pareja de momentos (t, t') para $t < t'$. La relación de *odds* cuando se consideran todas las posibles asociaciones entre los tiempos y las categorías está definido de la siguiente manera:

Ecuación 5
$$\log \theta_{jj't'} = \phi_{tt'} (\mu_{jt} - \mu_{t(j+1)}) (\mu_{j't'} - \mu_{t'(j'+1)})$$

En este se puede observar que el logaritmo de la relación de *odds* depende de cada momento t, t' y cada categoría j, j' , de manera que bajo este se construye la matriz de covarianzas con el mayor número de parámetros. No obstante, dado que las tablas de contingencia son cuadradas, es posible tener modelos más parsimoniosos a través del uso de parámetros homogéneos para las categorías de respuesta en cada par de momentos (Touloumis, Agresti y Kateri 2013). Dentro de estas estructuras que simplifican $\log \theta_{jj't'}$, se tienen las siguientes:

1. Uniforme: $\log \theta_{jj't'} = \phi$. Asume una estructura de asociación homogénea entre todos los pares de tiempos y categorías adyacentes.
2. Intercambiabilidad de categorías: $\log \theta_{jj't'} = \phi_{tt'}$. Asume que las relaciones de *odds* cambian entre períodos, pero no depende de la categoría de la variable, es decir, es homogénea por categoría.
3. Intercambiabilidad temporal: $\log \theta_{jj't'} = \phi (\mu_j - \mu_{j+1}) (\mu_{j'} - \mu_{j'+1})$. Considera una asociación homogénea en el tiempo, por lo que no asume ningún tipo de dependencia temporal, implicando igualdad en las relaciones de *odds* locales en los puntos de corte de las categorías (j, j') y (j', j) cuando $j \neq j'$.
4. RC (row column): $\log \theta_{jj't'} = \phi_{tt'} (\mu_{jt} - \mu_{t(j+1)}) (\mu_{j't'} - \mu_{t'(j'+1)})$. Tiene más parámetros que las anteriores estructuras dado que determina efectos heterogéneos tanto para los pares de momentos como para las categorías. No obstante, es más parsimoniosa que la determinada en la ecuación 5. De acuerdo con Touloumis *et al.* (2013), esta estructura debe ser utilizada con estructuras de asociación extremadamente complicadas.

La inclusión de estas estructuras de asociación, en conjuntos de información que tienen variaciones en el tiempo y respuestas multinomiales fuertemente correlacionadas, produce resultados más eficientes que cuando se consideran modelos cuya correlación de “trabajo” es de independencia. Así, para los datos ordinales, es recomendable utilizar la estructura uniforme para patrones de correlación intercambiables en el tiempo y la estructura de intercambiabilidad en categorías para estructuras más complejas (Touloumis, Agresti y Kateri 2013). Estas dos estructuras son consideradas en este trabajo.

5. Resultados

En esta sección se presentan los resultados del análisis de la categoría del colegio tomando como variables explicativas la naturaleza jurídica del colegio (oficial, no oficial), la jornada (completa, mañana, tarde), el enfoque educativo según la aproximación del nombre del establecimiento (técnico y académico) y el año en el que se realiza la clasificación.

De esta manera, el modelo definido tiene la forma $\text{logit}[\text{Pr}(Y_{it} \leq c)] = X'_{it} \beta$, donde $c=1, \dots, 6$ corresponde a cada una de las categorías de clasificación del establecimiento educativo, siendo 1 la categoría “inferior” y 6 la categoría “muy superior”. Cada uno de los años están identificados por $t=2001, \dots, 2011$, e $i=1, \dots, 4221$ identifica a cada uno de los establecimientos educativos considerados. X'_{it} es la matriz que contiene cada una de las variables independientes sobre las cuales se analiza la clasificación del establecimiento educativo i en el momento t . A partir

de esta estructura se tienen $L = \frac{11(11-1)}{2} = 55$ pares de comparación temporal sobre las cuales se construye la matriz de correlación de trabajo, de tamaño 55×55 .

En la tabla 4 se pueden ver las estimaciones del modelo marginal bajo independencia y cuatro estructuras de asociación. Como se puede observar, existen diferencias cuando se realiza el análisis bajo independencia y bajo estructuras de asociación. En particular, se destaca la pérdida de significancia de las variables asociadas a la jornada de los establecimientos y el tamaño de las instituciones asociado con el quintil del número de estudiantes inscritos (se tienen cambios de significancia al 99% a no significancia al 90%). Así mismo, se observan reducciones en los efectos de las diferentes categorías (en el anexo 1 se puede ver una comparación gráfica de los cambios de los parámetros estimados para cada una de las estructuras de asociación).

Estos cambios en los modelos implican ajustes en términos de la importancia de unos factores en relación con sus efectos sobre la clasificación de los colegios cuando se considera la independencia de las observaciones y cuando no. Así, la jornada de los colegios pierde poder explicativo cuando se incluyen las asociaciones de los datos.

En las diferentes estructuras de asociación se tienen parámetros y errores estándar similares, por lo que se puede suponer como válida la estructura de correlación de intercambiabilidad de categoría, la cual asume asociaciones variables entre períodos, pero constantes para cada una de las $C-1$ categorías.

De acuerdo con la metodología de análisis, la interpretación del modelo está relacionada con el logaritmo de la relación de *odds*, determinado por la ecuación 2, de manera que los parámetros estimados indican una variación proporcional a la relación de *odds*, bajo

$$\frac{\pi(Y_{ijk} \leq k)}{\pi(Y_{ijk} > k)} = \exp[\alpha_c] \exp[\beta x]$$

Para mejorar la interpretación, es útil multiplicar los parámetros β por -1 , con el fin de tener una interpretación directamente proporcional a la relación de

odds del desempeño de los colegios, es decir, que valores positivos impliquen una mayor probabilidad de estar en categorías más altas⁷.

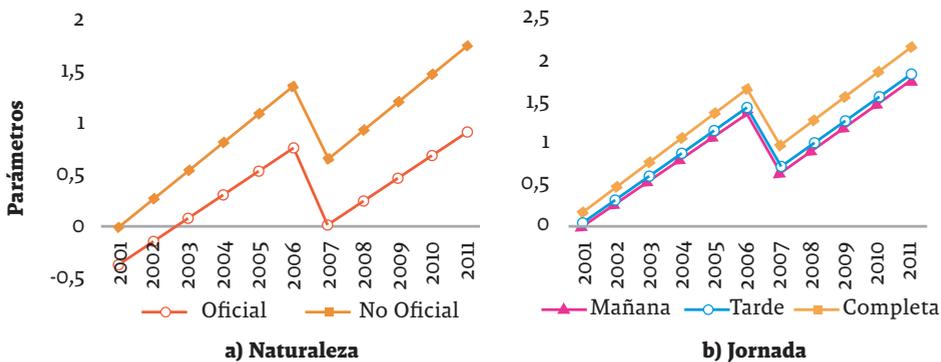
Esto implica que la relación de odds, sin tener en cuenta las constantes, quedan definidas como

$$\frac{\pi(Y_{ijk} > k)}{\pi(Y_{ijk} \leq k)} = \exp[-\beta x]$$

De acuerdo con los parámetros estimados del modelo “category exchange” de la tabla 4, se definen los perfiles de las gráficas 3 y 4.

Como se observa, los colegios oficiales considerados tienen una relación de odds 30% menor de estar en categorías más altas que los colegios no oficiales (gráfica 3.a). Dado que la interacción de la naturaleza con el año es significativa, se observa un efecto de “divergencia” entre los dos tipos de colegios analizados, de manera que en 2011 esta relación es menor en 150%. Así, si bien la clasificación que realiza el ICFES tiende a aumentar durante todo el período de análisis para todos los establecimientos considerados, aunque con un cambio estructural en 2007 que resultó ser significativo, se identifica un aumento en la brecha para todo el período, dadas las demás variables de análisis.

Gráfica 3. Perfiles estimados de la naturaleza y la jornada sobre la clasificación del establecimiento educativo, 2001-2011



En el caso de la jornada (gráfica 3), tomando la jornada de la mañana como referencia, no se observan diferencias en la relación de odds con la jornada de la tarde, en tanto que para la jornada completa la relación de odds de estar en categorías más altas es superior en 20% que la de la jornada de la mañana.

7 Una alternativa de interpretación de los parámetros se da mediante la transformación $\exp[-\beta]$. Con esta especificación se observan directamente los efectos multiplicativos de las covariables sobre la relación de odds. Se puede observar que para $-\beta > 0$ el efecto multiplicativo de una variable sobre $\frac{\pi(Y_{ijk} > k)}{\pi(Y_{ijk} \leq k)}$ es menor que uno, lo que indica una mayor probabilidad de $\pi(Y_{ijk} \leq k)$ y viceversa.

Tabla 4. Parámetros y errores estándar estimados bajo diferentes estructuras de asociación entre las observaciones

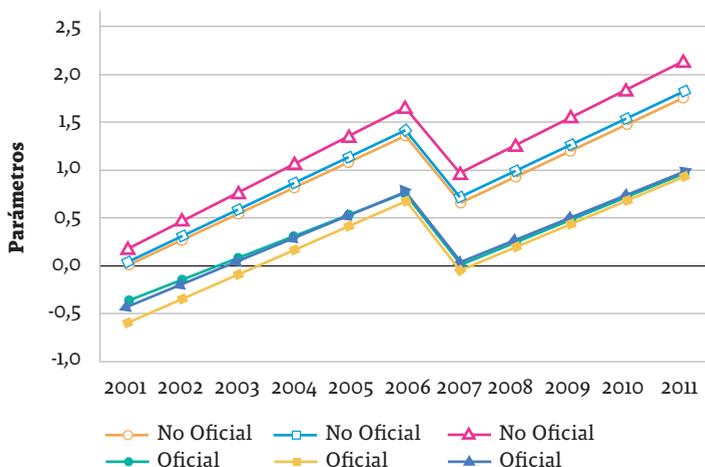
Parámetros	Independencia		Category Exchange		Uniforme		Time Exchange		Row - Column	
	Beta	s.d.	Beta	s.d.	Beta	s.d.	Beta	s.d.	Beta	s.d.
Alpha 1	-2,598	0,084	-2,068	0,058	-2,068	0,058	-2,197	0,063	-2,197	0,063
Alpha 2	-0,441	0,066	-0,090	0,047	-0,090	0,047	-0,091	0,049	-0,091	0,049
Alpha 3	1,139	0,066	1,310	0,049	1,310	0,049	1,362	0,051	1,362	0,051
Alpha 4	2,283	0,071	2,382	0,056	2,382	0,056	2,412	0,057	2,412	0,057
Alpha 5	3,599	0,081	3,710	0,071	3,710	0,071	3,669	0,070	3,669	0,070
Año	-0,297	0,009	-0,272	0,006	-0,272	0,006	-0,278	0,006	-0,278	0,006
Después de 2007: sí	1,243	0,068	0,973	0,049	0,973	0,049	0,633	0,050	0,633	0,050
Jornada: Tarde	-0,440	0,075	-0,041	0,037	-0,041	0,037	-0,006	0,040	-0,006	0,040
Jornada: Completa	-0,831	0,103	-0,184	0,076	-0,184	0,076	-0,107	0,087	-0,107	0,087
Naturaleza: Oficial	0,444	0,084	0,365	0,069	0,365	0,069	0,273	0,070	0,273	0,070
Enfoque: Técnico	0,223	0,118	0,210	0,093	0,210	0,093	0,206	0,100	0,206	0,100
Calendario: B	-0,108	0,083	0,011	0,037	0,011	0,037	0,020	0,036	0,020	0,036
Calendario: F	1,678	0,125	1,000	0,129	1,000	0,129	1,033	0,131	1,033	0,131
Quintil estudiantes: 2	0,307	0,050	-0,002	0,018	-0,002	0,018	0,011	0,018	0,011	0,018
Quintil estudiantes: 3	0,173	0,057	-0,008	0,020	-0,008	0,020	-0,003	0,021	-0,003	0,021
Quintil estudiantes: 4	-0,154	0,057	-0,030	0,020	-0,030	0,020	-0,025	0,021	-0,025	0,021
Quintil estudiantes: 5	-0,286	0,049	-0,009	0,018	-0,009	0,018	-0,006	0,018	-0,006	0,018
Año: después de 2007	-0,015	0,009	0,001	0,006	0,001	0,006	0,039	0,006	0,039	0,006
Tarde: Oficial	0,706	0,149	0,267	0,094	0,267	0,094	0,229	0,089	0,229	0,089
Completa: Oficial	0,592	0,138	0,240	0,101	0,240	0,101	0,133	0,108	0,133	0,108
Año: Oficial	0,088	0,009	0,046	0,006	0,046	0,006	0,050	0,006	0,050	0,006
Año: Tarde	-0,001	0,008	-0,004	0,004	-0,004	0,004	-0,003	0,004	-0,003	0,004
Año: Completa	-0,037	0,015	-0,021	0,009	-0,021	0,009	-0,023	0,009	-0,023	0,009
Año: Tarde: Oficial	-0,018	0,015	-0,024	0,009	-0,024	0,009	-0,025	0,009	-0,025	0,009
Año: Completa: Oficial	0,057	0,017	0,008	0,011	0,008	0,011	0,009	0,011	0,009	0,011

Fuente: cálculos propios. Niveles de significancia: 0 **** 0,001 *** 0,01 ** 0,05 * 0,1 ' ' 1

Con el supuesto de independencia, este efecto es de 130%, de manera que al utilizar las asociaciones existe una corrección importante hacia un efecto más conservador de la jornada.

Cuando se analizan las interacciones entre jornada y naturaleza jurídica del establecimiento, es importante observar que los dos efectos resultan significativos al 95%; sin embargo, estos actúan contrarrestando el efecto de la jornada sobre los colegios no oficiales, de manera que no se observan diferencias significativas para los colegios de jornada completa, mañana y tarde cuando estos son oficiales (gráfica 4).

Gráfica 4. Perfiles estimados de la naturaleza y la jornada e interacciones sobre la clasificación del establecimiento educativo, 2001-2011



Finalmente, con respecto a los colegios catalogados como “técnicos” se observa una reducción en la relación de *odds* en 18% con respecto a los colegios “académicos”, dadas las otras características como jornada y naturaleza del establecimiento. De acuerdo con los parámetros estimados, no se observaron efectos significativos para el tamaño de los colegios, medidos a través del quintil del número de estudiantes, cuando se consideran asociaciones en las observaciones.

6. Discusión y conclusiones

En la literatura consultada sobre la evaluación de la educación para Colombia, no se encuentran análisis en el tiempo que evidencien los efectos sobre la educación de las diferentes características de los establecimientos educativos. Esta situación se debe a la dificultad de comparar los resultados de las pruebas Saber del ICFES entre un año y otro, dados los múltiples cambios metodológicos que han presentado dichas pruebas, de esta manera, este valioso instrumento no ha sido utilizado en todo su potencial, principalmente por estas dificultades metodológicas.

Dada la ausencia de este tipo de estudios, en este trabajo se realizó una comparación de la clasificación de los colegios realizada por el ICFES, como una alternativa de aproximación al análisis de los resultados educativos a través del tiempo. Al utilizar el análisis categórico se solventan algunos problemas de comparabilidad al poder realizar una comparación de las distribuciones de variables ordinales entre los diferentes años, considerando las asociaciones de esta clasificación con las diferentes variables administrativas de los establecimientos. Esto es posible dado que la clasificación de los colegios se realiza dentro de cada año, de manera que se hace una comparación año a año de la posición relativa de los establecimientos.

Como lo muestran los resultados, se resalta la baja importancia de la jornada como predictor de la clasificación de los establecimientos educativos. Según el análisis, cuando se incluye la dependencia de las observaciones a través de la estructura de asociación de los *odds* locales, dejan de ser significativas las variables asociadas a la jornada -especialmente la jornada completa- cuando un establecimiento es oficial. Por su parte, los resultados asociados a la naturaleza del establecimiento son significativos dentro de todas las aproximaciones utilizadas, lo que indica la mayor importancia de otros elementos administrativos de los planteles oficiales, diferentes de la jornada, en el desempeño de las pruebas Saber.

Estos elementos indican que cuando se consideran las variables naturaleza jurídica y jornada de manera separada, se tienen resultados significativos para las dos variables; sin embargo, cuando se consideran las interacciones entre estas variables, así como las estructuras de dependencia de las observaciones, los resultados cambian sustancialmente. Resultados similares han sido documentados para países de la OCDE en el documento de Baines (2007), donde se recopilan antecedentes sobre la inexistencia de una relación directa entre las horas gastadas en la escuela, las tareas dejadas para realizar en la casa, el gasto en tecnología y los mejores resultados en las pruebas PISA.

Por último, en este documento se observa una mejora en la clasificación de todos los establecimientos educativos, sin embargo, esta mejora ha sido más sustancial para los establecimientos no oficiales que para los oficiales, de manera que esto indica un aumento en la brecha de los resultados entre los dos tipos de establecimiento.

Como una deuda importante de este trabajo está la necesidad de incluir variables de la infraestructura educativa para poder determinar cuáles son los efectos de variables de insumo educativo en la clasificación de los colegios a través del tiempo. Una posibilidad de realizar esto es mediante el uso de técnicas multivariantes que permitan la clasificación de los establecimientos educativos según sus características relacionadas con diferentes variables de infraestructura.

Referencias bibliográficas

Agresti, A. y R. Natarajan. «Modeling clustered ordered categorical data: A survey». *International Statistical Review*, n° 69(3), 2001: 345-371.

- Baines, L. «Learning from the World: Achieving More by Doing Less.» *Phi Delta Kappan*, n° 89(2), 2007: 98-100.
- Barón, Juan, Leonardo Bonilla y Catherine Rodríguez. «La calidad de los maestros en Colombia: Desempeño en el examen de Estado del ICFES y la probabilidad de graduarse en el área de educación». *Documentos de Trabajo sobre economía Regional*, 2011.
- Barrera-Osorio, Felipe y Dario Maldonado. «Calidad de la Educación Básica y Media en Colombia: Diagnóstico y Propuestas». *Documento CEDE*, 2012.
- Bojesen, Rune. «Analysis of ordinal data with cumulative link models—estimation with the ordinal package». *R-package*, version 13, 2011.
- Bonilla, Leonardo. «Doble jornada escolar y calidad de la educación en Colombia». *Banco de la República-Economía Regional*, 2011.
- Bonilla, Leonardo. «Movilidad inter-generacional en educación en las ciudades y regiones de Colombia». *Banco de la República-Economía Regional*, 2010.
- Bonilla, Leonardo y Luis Galvis. «Profesionalización docente y la calidad de la educación escolar en Colombia». *Documentos de Trabajo sobre economía Regional*, 2011.
- Card, David y Alan Krueger. «Does school quality matter? Returns to education and the characteristics of public schools in the United States». *National Bureau of Economic Research*, 1990.
- Castaño, Elkin. «El efecto colegio sobre la variabilidad del rendimiento en matemáticas». *Lecturas de Economía*, n° 49, 1998: 47-57.
- Courgeau, Daniel. *Multilevel synthesis: from the group to the individual*. Dordrecht: Springer Science & Business Media, 2007.
- Duarte, Jesús, María Soledad Bos y José Moreno. «Calidad, Igualdad y Equidad en la Educación Colombiana (Análisis de la Prueba Saber 2009)». *Notas Técnicas. Banco Interamericano de Desarrollo – BID*, 2012.
- Gaviria, Alejandro y Jorge H. Barrientos. *Determinantes de la calidad de la educación en Colombia*. Documento de trabajo. Bogotá: Fedesarrollo, 2001.
- Heagerty, Patrick J. y Scott L. Zeger. «Marginal regression models for clustered ordinal measurements». *Journal of the American Statistical Association*, n° 91435, 1996: 1024-1036.
- Iregui, Ana María, Ligia Melo y Jorge Ramos. «Evolución y análisis de eficiencia de la Educación en Colombia». *Borradores de Economía. Banco de la República*, 2006.
- Liang, Kung-Yee y Scott L. Zeger. «Longitudinal data analysis using generalized linear models». *Biometrika*, 1986: 13-22.
- Rangel, Claudia Liliana. 2006. How do schools and neighborhoods matter for student achievement?: a study of educational inequality in Cartagena, Colombia. Tesis Master of Sciences in Human and Community Development, University of Illinois at Urbana-Champaign.
- Raudenbush, Stephen W. «Educational applications of hierarchical linear models: A review». *Journal of Educational and Behavioral Statistics*, n° 13.2, 1988: 85-116.
- Stokes, Maura E., Charles S. Davis y Gary G. Koch. *Categorical data analysis using SAS*. SAS institute, 2012.

Thomas, Vinod, Yan Wang y Xibo Fan. «Measuring education inequality: Gini coefficients of education». *World Bank Publications*, Vol. 2525, 2001.

Touloumis, Anestis, Alan Agresti y Maria Kateri. «GEE for multinomial responses using a local odds ratios parameterization». *Biometrics*, n° 69.3, 2013: 633-640.

Ziegler, Andreas, Christian Kastner y María Blettner. «Generalised Estimating Equations: An Annotated Bibliography». *Biometrical Journal*, Vol. 40, n° 2, 1998: 115-139.

Anexo 1. Comparación gráfica de los efectos bajo diferentes supuestos de asociación en las observaciones

