



ALCALDÍA DE JAMUNDÍ
VALLE DEL CAUCA



Gobierno de los
Ciudadanos

Documento de Investigación # 2

Secretaría de Planeación y Coordinación
Laboratorio de Datos - Jamundí

Determinantes del desempeño educativo en Jamundí (2020)

Julio Alfredo Escobar Lizano
Juan Camilo Herrera Palacio
Christian André Villegas Solís
Juan David Yela Correa

Andrés Felipe Ramírez Restrepo
Alcalde de Jamundí

...

Diciembre, 2021

Determinantes del desempeño educativo en Jamundí (2020)

Julio Alfredo Escobar Lizano¹; Juan Camilo Herrera Palacio²; Christian André Villegas Solís³; Juan David Yela Correa³

Resumen: Esta investigación busca encontrar los principales determinantes que afectan el desempeño de los estudiantes de Jamundí en educación media, medido a través de los resultados en las pruebas Saber 11. La estimación se realiza utilizando modelos multinivel y MCO; el insumo principal tanto para la selección de variables como para la motivación de este estudio es el análisis descriptivo realizado previamente sobre el sector educativo en el municipio [1]. Se concluye que el modelo más adecuado es el multinivel, y con este se encuentran dependencias de los resultados frente a variables del contexto de los estudiantes como el nivel socioeconómico y la distancia del alumno a la sede educativa, así como también variables propias de la sede educativa, tales como lo son la ubicación de la sede y la proporción de profesores mayores de 40 años.

Abstract: This research aims to find the main determinants that affect the performance of Jamundí secondary school students, measured by the Saber 11 test results. The estimation is done by using multi-level models and OLS; the main input for the variables selection and the motivation of this research is the descriptive analysis previously done about the education sector in the city. It is concluded that the most appropriate model is the multi-level model which shows dependencies of the results on the students context variables such as socioeconomic level and distance between the student and the educational center, there are educational center variables as well such as its location and the proportion of teachers over 40 years of age.

¹ Magíster en Economía y Políticas Públicas - Universidad de los Andes

² Economista - Pontificia Universidad Javeriana

³ Economista - Universidad del Valle

Introducción

A la fecha se ha avanzado significativamente en el reconocimiento de la educación como herramienta de movilidad social, destacándose los aportes hechos por la Comisión Internacional para la Educación en el Siglo XXI, los cuales han reflexionado acerca del papel de esta ante las transformaciones económicas, tecnológicas y políticas, y sobre la misión que tiene el proceso educativo para promover los ideales de paz, democracia y justicia social [5]; en esta misma línea sobre el papel de la educación en la sociedad desde un ámbito integral y multisectorial también se encuentran estudios como [9]; [18]; [19]. Robertson y Reynolds (2010) [18] destacan la relevancia a corto y largo plazo de los resultados obtenidos en media escolar; de esta manera, dada la importancia de la educación en el individuo y la sociedad, entender los determinantes del desempeño educativo en media escolar, es un insumo fundamental para invertir eficiente y eficazmente los recursos públicos.

En aras de generar un mayor impacto mediante la inversión de los recursos públicos de la Alcaldía de Jamundí destinados al sector educativo, esta investigación busca encontrar los principales determinantes del rendimiento de los estudiantes de grado once del municipio en el año 2020, se toma como insumo fundamental para la elección de variables los principales resultados descriptivos encontrados en [1]. Estos resultados muestran las relaciones entre el desempeño educativo y las condiciones socioeconómicas de los estudiantes, así como la relación existente con el entorno escolar, mostrando dos niveles de análisis, el nivel de estudiante y el nivel de sede educativa. Es por esto que se utiliza un análisis mediante modelos jerárquicos lineales (multinivel) sin dejar de lado el comparativo con estimaciones tradicionales como Mínimos Cuadrados Ordinarios (MCO). Los resultados evidencian fuertes dependencias del rendimiento educativo frente a variables contextuales de las sedes educativas, así como también del contexto de los alumnos. A continuación el documento se desarrolla en las siguientes secciones: estado del arte; metodología; estimación y resultados; conclusiones.

Estado del arte

Desde el sector público, son muchos los debates que se han generado en torno a las diferentes formas en que la educación representa un determinante del progreso y desarrollo social. Según Duncan (2007) el papel que juega la educación es fundamental, representa un elemento básico para el progreso de la ciudadanía, debido a que, los procesos de formación académica están estrechamente relacionados a resultados en el largo plazo, mejorando las condiciones de vida de los que se educan.

Esta idea también se sustenta en los planteamientos de Robertson y Reynolds (2010), los cuales argumentan que un buen desempeño en el campo académico en las etapas del sistema educativo, sobre todo en la secundaria, tiene un impacto positivo directo sobre el empleo y los ingresos económicos a lo largo de la vida, ya que, alcanzar buenas calificaciones en este nivel, facilita el acceso a la educación superior. Por otra parte, no acceder o abandonar los procesos de formación educativa, genera no solo consecuencias negativas a nivel individual sino problemáticas colectivas que afectan a toda la sociedad, bajos salarios, criminalidad y violencia, aumento de conductas

psicológicas de riesgo (depresión, ansiedad, suicidios, etc.), entre otras, que en el largo plazo se traducen en limitantes al desarrollo, la prosperidad y bienestar de toda la comunidad [19].

El siguiente aspecto a analizar es el desempeño académico, indicador que representa los avances o retrocesos durante el proceso de formación educativa en los distintos niveles establecidos. El desempeño o rendimiento académico, es ese acumulado de conocimientos y capacidades escolares adquiridas y exhibidas en un área o materia específica por los estudiantes dentro de su proceso de formación, los cuales son medidos por cualquier mecanismo de evaluación establecido, principalmente la aplicación de exámenes estandarizados internos o externos [10]. No obstante, estos mecanismos tradicionales de evaluación suelen limitar el concepto de rendimiento académico, puesto que, cada estudiante cuenta con unas potencialidades y aptitudes distintas, por lo tanto, es fundamental plantear estrategias de evaluación que permitan identificar de manera integral el rendimiento académico de los estudiantes [2]. Esto conduce a dimensionar que existen un conjunto de determinantes que influyen de manera favorable o desfavorable en el rendimiento y desempeño académico de los estudiantes durante su proceso de formación.

En este contexto nacen los estudios de Eficacia Escolar (EE), la cual parte del supuesto que el éxito académico es maleable por condiciones externas, por tanto, buscan relaciones causales entre los insumos financieros o materiales, procesos de enseñanza y otros, que pueden explicar el desempeño académico. El punto de partida de estos trabajos es Coleman (1966), el cual buscó encontrar la combinación de insumos escolares para maximizar el desempeño estudiantil. En un segundo momento, se buscaron variables de procesos, como lo son las prácticas pedagógicas y ambientes de los colegios, las cuales resultaron ser significativamente importantes para explicar los resultados académicos. A estos trabajos le siguieron investigaciones relacionadas con el impacto de las características de los planteles educativos (profesores, salones, etc) y finalmente en los años 90 se iniciaron las investigaciones que relacionan el contexto familiar de los estudiantes y las características de los colegios como los factores claves en el desempeño académico [21].

Noguera y Wells (2011) probaron que los contextos sociodemográficos difíciles pueden afectar desfavorablemente el rendimiento académico de los estudiantes desde tres aspectos, el primero consiste en que estos no obtienen apoyos extraescolares, es decir, ayudas por fuera de los espacios escolares, como por ejemplo, refuerzos académicos en distintas áreas del conocimiento, tampoco cuentan con actividades extracurriculares relacionadas con prácticas deportivas, artísticas y culturales que potencien sus capacidades cognitivas. El segundo aspecto, radica en que continuamente están expuestos a condiciones álgidas de seguridad y bienestar; de igual forma su acceso a la salud es limitado, las condiciones de vivienda desfavorables y cuentan con una mala alimentación. Finalmente, en el tercer aspecto, plantean que la mayoría de padres no tiene acceso a altos niveles de capital social, ya que, también estuvieron expuestos a condiciones contextuales adversas, lo que complica la situación interna y el contexto inmediato de los estudiantes, en la medida que carecen de apoyo familiar que los anime a continuar y obtener buenos resultados en la escuela.

Por su parte, Scheerens y Creemers (1989) establecieron que la institución educativa, los docentes y particularidades de las aulas deben ser parte de dichas dimensiones, por ende, sería preciso estudiar el grado de incidencia y la manera como influyen en el desempeño académico. En ese sentido, los horarios de los cursos, los tamaños de los grupos, las condiciones de la biblioteca y otras características estructurales y funcionales de las instituciones educativas, influyen directamente en el desempeño académico de los estudiantes. En cuanto al tamaño de los grupos, se destaca el estudio experimental STAR project implementado en 2008 en el estado de Tennessee, el cual encontró evidencia suficiente para afirmar que el tamaño de los cursos incide negativamente en el rendimiento académico [12].

Por otro lado, los recursos didácticos y metodológicos representan herramientas significativas de apoyo que potencian cualquier práctica educativa, por ende, es un factor clave que contribuye a mejorar el desempeño de los estudiantes; otro aspecto importante son los factores pedagógicos, principalmente las funciones de los docentes y las propuestas pedagógicas establecidas por cada institución [20]. Además, los profesores representan un papel fundamental dentro de cualquier proceso educativo, ya que, son los mediadores dentro de dichos procesos, acompañando y siendo el principal apoyo de los estudiantes en su proceso de formación.

Diversos autores han estudiado también han estudiado la importancia de facilitar el acceso a servicios por parte de la población más vulnerable, destacándose el impacto de la dificultad de acceso a los servicios de salud [16], y a una garantía alimenticia [11]. Frente a este aspecto se resalta el estudio de Tale (2001) [22] que evaluó el impacto de la distancia del colegio a los hogares de los estudiantes, sobre los resultados académicos en educación primaria. Por otro lado, se encuentra la investigación de Owioye (2011) [17] para escuelas de secundaria en Nigeria, y el artículo de Lin, Chieh y Li (2014) [13] para el área rural de Taiwán. Del primero, se encuentra que existe una diferencia significativa entre los resultados académicos alcanzados por los estudiantes de escuelas urbanas y rurales, favoreciendo a los primeros. Mientras que el segundo estudio valida los problemas que genera contar con escuelas a larga distancia, y también encuentra un efecto diferencial en la existencia de dificultades de acceso a los colegios entre los adolescentes y los niños, siendo mayor en los primeros. En general estas investigaciones son un marco de referencia para los gobernantes en la implementación de políticas de transporte escolar y en la construcción de instituciones en áreas remotas; no obstante, aún no se establece con certeza los mecanismos por los cuales se generan dichas relaciones de causalidad.

Dado el reconocimiento de las diferentes aristas que componen el desempeño académico, en la literatura de la eficacia escolar se han venido implementando los modelos multinivel. Entre los autores destacados en determinar estas relaciones conjuntas se encuentra Coleman (1996) [4], el cual encontró que el 10 % de la varianza del rendimiento se explica por factores de la escuela. Por su parte, Erbring and Young (1979) [6] encontraron que los procesos psicológicos median dentro de las relaciones individuales y del grupo. A su vez que Vélez (1994) [23] realizó un estudio de eficiencia educativa para 18 países de Latinoamérica, mientras Heneveld y Craig (1996) [21] hicieron lo mismo para alumnos de África subsahariana. Entre los estudios más recientes dentro de esta literatura, se encuentra el estudio de Ludtke et al (2009)

[15], donde evaluaron el impacto del ambiente de aprendizaje, realizando una evaluación de cómo debería de introducirse la percepción agregada, encontrando que la estimación es muy sensible a la medida utilizada. Por otro lado, en la publicación de frontera, se destaca el estudio de Liouaeddine y Bijou (2021) [14], quienes evaluaron el desempeño académico de los estudiantes de Marruecos, encontrando que los indicadores de pobreza y género fueron los más explicativos. Aunque este tipo de métodos son utilizados principalmente para la evaluación de modelos educativos, últimamente se han implementado en el campo de negocios, psicología, consumo de energía, etc [14], donde sobresale la investigación de Jeong y Kim (2019) [3], donde comparan la trayectoria laboral asociada a la educación vocacional y la educación en general, realizando una evaluación de los programas Vocational education and Training (VET) implementados por diferentes países.

Metodología

Los modelos multinivel suponen que el individuo y los grupos sociales forman parte de un sistema jerárquico, en donde cada uno aporta al desempeño académico de forma independiente (efecto fijo) y a la vez diferencial (efecto aleatorio) [7]. Este modelo supone que los grupos de los colegios toman las características de los alumnos que los componen, creando características particulares en cada grupo que a su vez influyen en los resultados académicos de los estudiantes. Por tanto, tratan de explicar el desempeño académico en relación con las características propias del estudiante (i) y las características del colegio (j). Una diferencia frente a los modelos estadísticos estándar, es que no asumen independencia entre las variables, pues, ante una estructura jerárquica se violaría este supuesto, dando una significancia espuria de los resultados si se estimaran por métodos tradicionales [8].

Para su implementación, la metodología inicia estimando la ecuación (1) sin variables explicativas, donde $Y_{i,j}$ es el puntaje global de la prueba Saber 11 en el año 2020 del estudiante i , perteneciente a la sede educativa j . En esta etapa, el $\beta_{0,j}$ está indexado en j , puesto que equivale al resultado promedio del resultado educativo por colegio, a lo que se le suma el error individual $e_{i,j}$. (β)

$$Y_{i,j} = \beta_{0,j} + e_{i,j} \quad (1)$$

Para conocer los efectos comunes de cada colegio j sobre sus estudiantes, se genera una nueva ecuación de ($\beta_{0,j}$), la cual está en función del promedio del resultado de la prueba para toda la población $\gamma_{0,0}$ y un término de error a nivel de colegio $\xi_{0,j}$, este último siendo el efecto aleatorio a nivel de colegio.

$$\beta_{0,j} = \gamma_{0,0} + \xi_{0,j} \quad (2)$$

Reemplazando (1) en (2), se obtiene la ecuación (3).

$$Y_{i,j} = \gamma_{0,0} + \xi_{0,j} + e_{i,j} \quad (3)$$

En este modelo la varianza de ($\xi_{0,j}$) es (σ_0^2) y la varianza de ($e_{i,j}$) es (σ_1^2), con las que se construye el coeficiente de correlación mediante la ecuación (4), capturando

cuánto representa la varianza del colegio respecto a la varianza total.

$$\rho = \frac{(\sigma_0^2)}{(\sigma_0^2 + \sigma_1^2)} \quad (4)$$

Una vez se encuentre que el aporte a la varianza de (σ_0^2) es significativo, se ha verificado que la aplicación del modelo multinivel es válida y que existe una estructura jerárquica en los datos.

Para el caso donde se espera que la diferencia de los resultados académicos promedio de los colegios varíe por características especiales de este, se debe modificar la ecuación de $\beta_{0,j}$, por tanto se añade a la ecuación (2) los coeficientes $\gamma_{0,1}, \gamma_{0,2}, \dots, \gamma_{0,n}$, que representan el impacto promedio de n características del colegio (C). Dando origen así a la ecuación (5).

$$\beta_{0,j} = \gamma_{0,0} + \gamma_{0,1} * C1_j + \dots + \gamma_{0,n} * Cn_j + \xi_{0,j} \quad (5)$$

Para el nivel de estudiante, el mismo comportamiento esperado se representa mediante la expresión (6), donde se han añadido las m variables (E) propias del estudiante i .

$$Y_{i,j} = \beta_{0,j} + \beta_{1,j} * E1_{i,j} + \dots + \beta_{m,j} * Em_{i,j} + e_{i,j} \quad (6)$$

Puede ocurrir que algunas variables tengan efectos aleatorios, es decir, que su efecto sobre el puntaje del estudiante dependa del tipo de colegio en el que se encuentre. Para este caso, si consideramos que el impacto de la variable $E1$ presenta dicho comportamiento, la ecuación para el $(\beta_{1,j})$ estaría dada por (7). Este tipo de casos aplican en variables como el nivel de pobreza (INSE), en donde la literatura ha mostrado su impacto sobre el resultado del estudiante $(\beta_{1,j})$ en dos sentidos: un efecto promedio de esta variable $\phi_{1,0}$ y un efecto aleatorio $\phi_{1,j}$ que representan el impacto de cada una de las k variables a nivel de colegio (D) que impactan en este efecto.

$$\beta_{1,j} = \phi_{1,0} + \phi_{1,1} * D1 + \dots + \phi_{1,k} * Dk + \mu_{1,j} \quad (7)$$

Se sustituyen en la ecuación (6) las ecuaciones (5) y (7), dando como resultado la expresión (8). Para ejemplificar la modelación, se considera que solo 1 variable cuenta con efectos aleatorios (E1) y las restantes variables solo cuentan con efectos fijos.

$$Y_{i,j} = \gamma_{0,0} + \gamma_{0,1} * C1_j + \gamma_{0,2} * C2_j + \dots + \gamma_{0,n} * Cn_j + \\ + [\phi_{1,0} + \sum_{p=1}^k \phi_{1,p} * Dp_j + \mu_{1,j}] * E1_{i,j} + \sum_{q=2}^m \beta_m E_{q,i,j} + e_{i,j} \quad (8)$$

Que descomponiendo en términos de los efectos resulta en (9).

$$\begin{aligned}
 Y_{i,j} = & \underbrace{\gamma_{0,0} + \gamma_{0,1} * C1_j + \gamma_{0,2} * C2_j + \cdots + \gamma_{0,n} * Cn_j}_{\text{Efecto Fijo}} + \cdots \\
 & \cdots + \underbrace{\gamma_{1,0} * E1_{i,j} + \beta_2 * E2_{i,j} + \cdots + \beta_n * En_{i,j}}_{\text{Efecto Fijo}} + \cdots \\
 & \cdots + \underbrace{\sum_{k=1}^n \gamma_{1,k} * Ck_j * E1_{i,j}}_{\text{Efecto fijo}} + \underbrace{\xi_{1,j} * E1_{i,j} + e_{i,j}}_{\text{Efecto aleatorio}}
 \end{aligned} \tag{9}$$

De la ecuación (9), se destaca que el error asociado a cada β ($\xi_{1,j}$) se asume de media 0 e independiente al error estructural del modelo ($E[\xi_{1,j} * E1_{i,j}] = 0$). En este tipo de modelos, la variación de los coeficientes de cada variable está explicada por ξ_j , donde el subíndice j hace referencia a la clase a la que pertenece. Al estimar este modelo por Mínimos Cuadrados Ordinarios, esta variación del error y los parámetros generan problemas de heterocedasticidad, haciendo que se pierda eficiencia en la estimación. Por tal motivo, siguiendo a Thomsons (1962) se estiman los parámetros mediante el método de Máxima Verosimilitud Restringida (REML por sus siglas en inglés), donde la idea básica de este proceso es que durante la estimación de máxima verosimilitud (ML por sus siglas en inglés) se forman un conjunto de restricciones lineales que no dependen del efecto fijo β , pero sí de la varianza a estimar. Luego, se aplica el método de ML usando la distribución de verosimilitud formada por las restricciones lineales previamente definidas ¹.

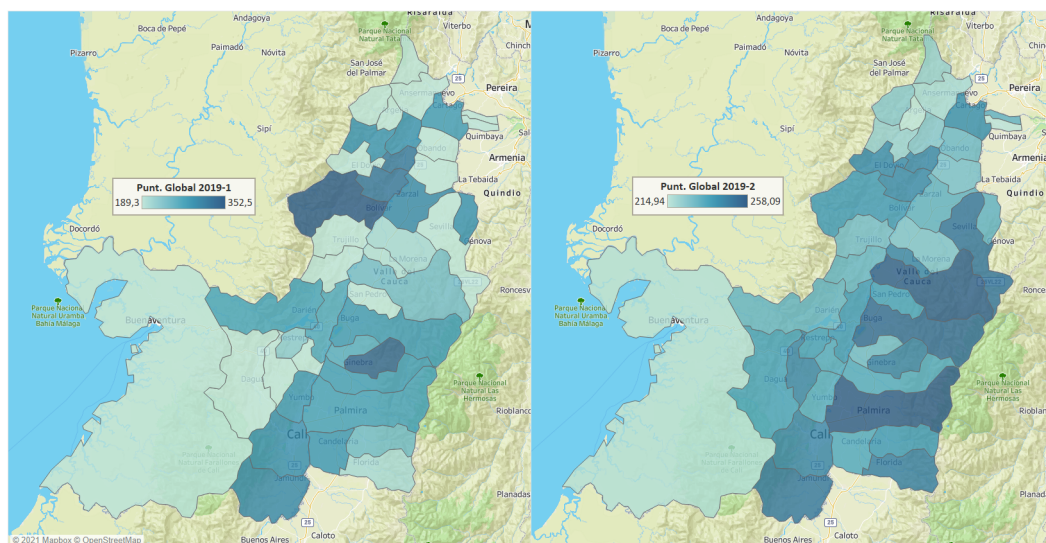
Datos

Los datos que se tuvieron a disposición para el presente estudio fueron: base del Sistema Integrado de Matrícula (SIMAT) y la base ICFES, ambas disponibles para el municipio de Jamundí en el año 2019. La primera base de datos contiene información a nivel de características de estudiantes y a nivel de Colegio, incluyendo información de profesores y aulas de clase. Por otro lado, la base del ICFES contiene los resultados de los estudiantes en la prueba Saber 11 así como otras variables de contexto familiar. En los resultados del ICFES de los colegios públicos se observó un puntaje medio de 235, sin presentar niveles elevados de dispersión respecto a la media (desviación estándar de 11 puntos). En este grupo de colegios sobresale el resultado obtenido por los municipios de Palmira (258), Tuluá (256), Buga (253), Sevilla (250) y Jamundí (249), los cuales obtuvieron puntajes superiores al promedio departamental y Cali (246) (Ver Figura 1).

Siguiendo la información presentada por el DANE (EDUC 2019), en Jamundí se registran 119 sedes educativas, de las cuales 5 se encuentran inactivas, de las 114 sedes restantes el 56 % están ubicadas en los 535 km^2 del área rural del municipio. Por su parte, en la zona urbana se encuentran un total de 50 sedes educativas en 42

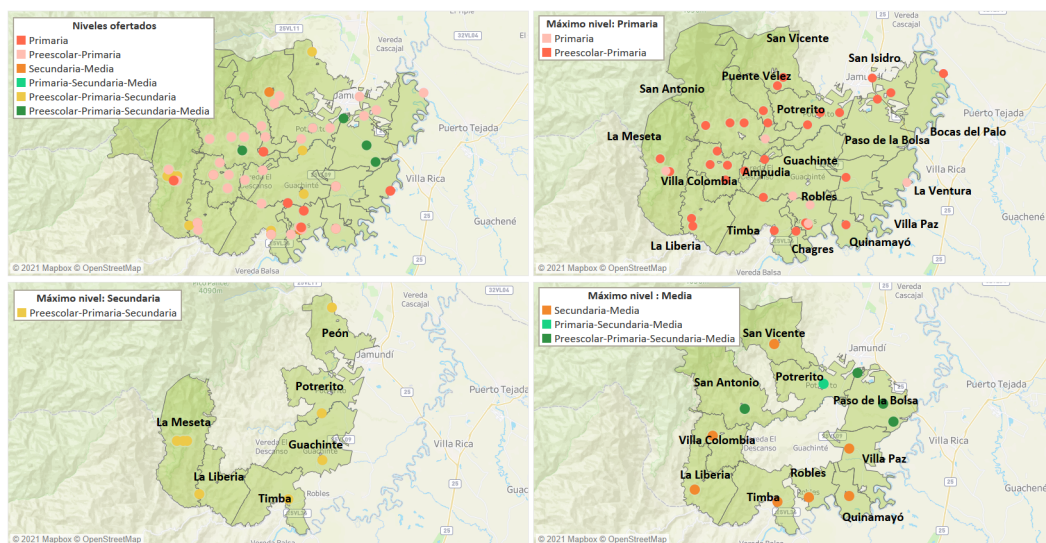
¹Para más detalles revisar el libro "The problem of negative estimates of variance components" de Thompson (1962) o el manual de uso del Paquete xtmixed de stata

Figura 1: Puntaje global ICFES de municipios del Valle del Cauca (2019)



km². En cuanto a la naturaleza de estas sedes, en la zona rural el 95 % son públicas, mientras que en la zona urbana la participación de los establecimientos oficiales cae a un 28 %. En cuanto a la ubicación de las sedes educativas, se observa que solo 3 sedes de las 64 activas ofrecen todos los niveles, desde preescolar a media. Mientras, que en total solo 18 % de los establecimientos cuentan con educación media, la cual es finalmente la que permite completar el bachillerato. En la zona rural, 8 corregimientos de los 19 que existen en Jamundí (42 %) cuentan con todos los niveles de educación, destacándose Paso de la Bolsa y San Antonio que cuentan con todos los niveles en una sola sede educativa.

Figura 2: Niveles educativos ofertados por las sedes de la zona rural



Al realizar el cruce de información entre las bases de datos SIMAT e ICFES 2019, inicialmente se encuentra un total de 857 alumnos y un total de 15 sedes educativas, donde el 65 % están en la zona urbana. Dado que algunos estudiantes tenían infor-

mación faltante, fue necesario aproximar mediante información de sus vecinos. No obstante, hubo estudiantes donde no fue posible corregir la información, dejando un total de 831 estudiantes con información completa para realizar la estimación, de los cuales el 57 % son mujeres y el 1 % presentan alguna discapacidad. En el Cuadro 1 se aprecian algunos datos descriptivos, se observa que en promedio la planta profesoral de los colegios públicos de Jamundí cuenta con un 41 % de profesores con posgrado, no obstante, se evidencia una alta dispersión entre los colegios (de 36 %), donde hay colegios sin este tipo de profesores. En relación a la distancia, se encontró que en promedio los estudiantes están a 2.6 kilómetros de distancia de sus colegios, no obstante, se registró una alta dispersión en donde el quintil más alto de la distribución debe recorrer más de 5 kilómetros para llegar al colegio, mientras que el primer quintil debe recorrer menos de 1 kilómetro.

Cuadro 1: Resumen base de datos

	N	Media	DE	Min	Max
Puntaje Global	831.00	244.97	41.16	148.00	371.00
INSE	831.00	49.07	7.20	23.45	71.75
Edad	831.00	17.30	0.92	15.00	20.00
Distancia	831.00	2667	3313	56.22	18822
% Estudiantes Discapacitados	831.00	0.01	0.08	0.00	1.00
Género Femenino	831.00	0.57	0.50	0.00	1.00
% profesores con posgrado	831.00	0.41	0.36	0.00	0.98
% profesores con posgrado > 40	831.00	0.77	0.16	0.00	1.00
% Estudiantes sede urbana	831.00	0.65	0.48	0.00	1.00
% Estudiantes jornada Mañana	831.00	0.46	0.50	0.00	1.00
% Estudiantes jornada Tarde	831.00	0.10	0.31	0.00	1.00
% Estudiantes jornada Única	831.00	0.44	0.50	0.00	1.00
% Estudiantes educación tradicional	831.00	0.87	0.34	0.00	1.00

En cuanto a las características de los estudiantes, registraron una media de 49 puntos en el INSE y desviación estándar (DE) de 7,2. Al desagregar por zona, se encuentran diferencias sistemáticas en los estudiantes que asisten a sedes en la zona urbana y rural, siendo superior el puntaje de los estudiantes que asisten a la zona urbana. Sin embargo, la relación positiva entre INSE y el puntaje global se mantiene en ambos sectores y son similares las pendientes estimadas, 0.75 y 0.78 respectivamente (Ver Figura 3). Por otro lado, una pregunta natural al observar tantos corregimientos sin el ciclo completo de educación, es si la distancia que deben recorrer los estudiantes para llegar a los colegios ha sido un limitante para su formación académica. Una primera hipótesis se resalta en la Figura 4, donde se observa que los estudiantes que viven cerca a sus colegios, tienen un mejor resultado en su desempeño académico, por los beneficios anuados de contar con mejores condiciones económicas (INSE), esto lo muestra la pendiente más pronunciada.

Figura 3: Correlación entre INSE y logro académico desagregado por Zona de la institución educativa

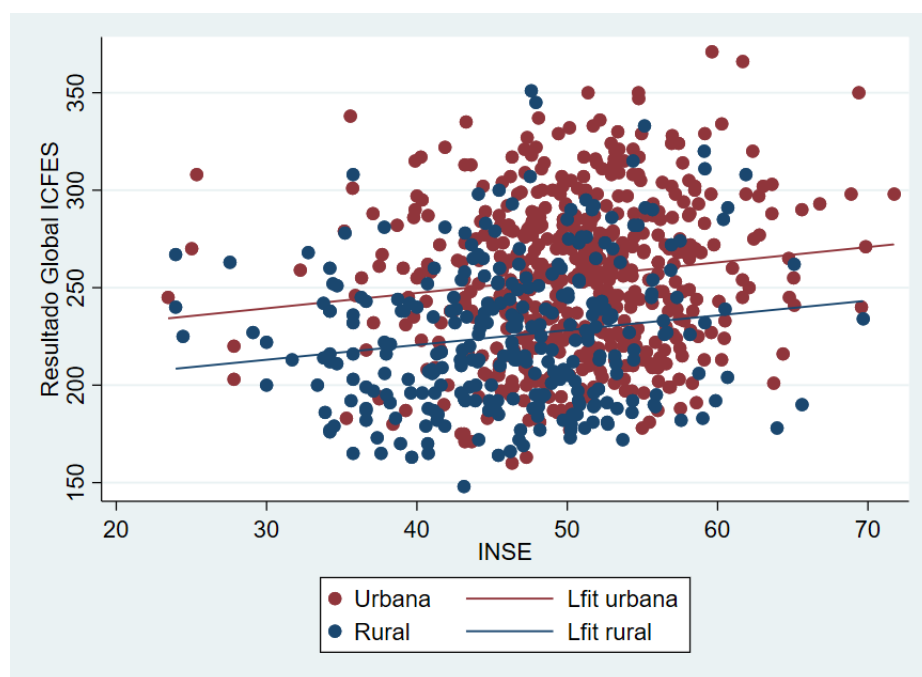
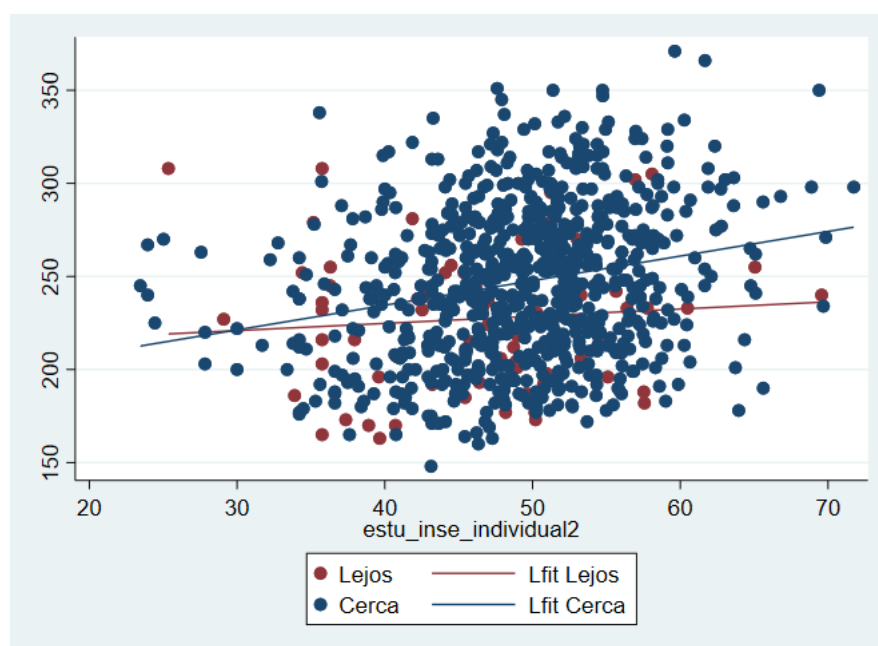


Figura 4: Correlación entre INSE y logro académico desagregado por distancia



Estimación y Resultados

El modelo vacío estimado sugiere que el porcentaje de la varianza que está explicado por las características del estudiante, según el Cuadro 2 fue el 70 % $(1 - \frac{458,18}{458,18+1365,78})$. Al comparar este resultado con el 90 % propuesto por Coleman (1996), se concluye que sí se requiere implementar un análisis multinivel.

Según el análisis descriptivo hecho previamente, se determinaron los niveles de efec-

Cuadro 2: Regresión modelo vacío

	(1) Puntaje Global
Constante	231.6*** (39.31)
Efecto Aleatorio	
Varianza de β_0 (σ_0^2)	458.18 (190.5)
Varianza residual (σ_1^2)	1365.78 (67.6)
N	831
adj. R^2	
t Estadístico en paréntesis	
* p<.10, ** p<.05, *** p<.01	

tos aleatorios existentes ². En relación al intercepto (β_0), la ecuación (10) recoge las variables a nivel de colegio que parecen generar un resultado diferencial del estudiante i: promedio INSE, % profesores con posgrados, % profesores > a 40 años y zona de ubicación de la sede.

$$\beta_{0,j} = \text{Promedio INSE}_j + \gamma_{0,1} \text{Zona Urbana}_j + \dots \quad (10)$$

$$\dots + \gamma_{0,2} \% \text{ Profesores Posgrados}_j + \gamma_{0,3} \% \text{ Profesores} > 40 \text{ años}_j + \xi_j$$

En cuanto al nivel de estudiantes, se determinó que las variables INSE y distancia, tienen posiblemente un efecto diferencial (β_i) en la forma en que impactan a la variable Resultado global (desempeño), de acuerdo a las condiciones del colegio, tales como: el porcentaje de profesores mayores de 40 años y el porcentaje de profesores con posgrado que atienden a estos estudiantes en su proceso académico. Por tanto, surgen la ecuación β_1 para el INSE y β_4 para la distancia, las cuales se incorporan en la ecuación (13) a nivel de estudiante.

$$\beta_{1,j} = \gamma_{1,0} + \gamma_{1,1} \% \text{ Profesores mayores a 40 años}_j + \xi_{1,j} \quad (11)$$

$$\beta_{4,j} = \gamma_{4,0} + \gamma_{4,1} \% \text{ Profesores con posgrados}_j + \xi_{4,j} \quad (12)$$

$$\begin{aligned} \text{P.Global}_{i,j} = & \beta_{0,j} + \beta_{1,j} \text{INSE}_{i,j} + \beta_2 \text{Edad}_{i,j} + \beta_3 \text{Femenino}_{i,j} + \dots \\ & \dots + \beta_{4,j} \text{Lejos}_{i,j} + \beta_5 \text{Lejos} * \text{INSE}_{i,j} + \beta_6 \text{Edad} * \text{Lejos}_{i,j} + \dots \\ & \dots + \beta_7 \text{Discapacidad}_{i,j} + \beta_8 \text{Estrato}_{i,j} + \beta_9 \text{Jornada única} + e_{i,j} \end{aligned} \quad (13)$$

²Para revisar las hipótesis de efectos diferenciales en la media β_0 de la estimación y el efecto diferencia en la pendiente β_i , remitirse a los Anexos

En el Cuadro 3 se registran las estimaciones realizadas. Las estimaciones 1 y 2 se realizaron mediante MCO, por tanto, cuentan con problemas de heterocedasticidad descrita en la sección metodología, lo cual se puede confirmar con las pruebas realizadas en los anexos, donde ambas tienen este problema sumado a que los errores no son normales. Las estimaciones 3 y 4, cuentan con corrección de varianza y se realizaron mediante REML. En cuanto a los resultados, a nivel de colegio, las 4 estimaciones muestran que contar con una mayor proporción de profesores con posgrados aumenta el puntaje global, siendo no significativo para los modelos 3 y 4. Por su parte, el contar con profesores de más de 40 años reduce los resultados académicos entre 34 y 85 puntos, según las 4 estimaciones propuestas. En las estimaciones multinivel, el efecto del INSE promedio arrojó que aumentos en 1 punto del INSE del colegio, se traducen en promedio en 3 puntos del puntaje global del estudiante, mientras que la ubicación de los colegios de la zona urbana obtienen un puntaje mayor en 14 puntos respecto a los rurales, siendo ambos coeficientes significativos en los 3 modelos.

A nivel de estudiante, se encuentra que un aumento en una unidad del INSE individual en los modelos 1 y 3 reportan un incremento en el resultado global en 0.38 y 0.39 puntos, respectivamente. No obstante, al incluir la interacción % profesores mayores a 40 años e INSE, este coeficiente se vuelve negativo y pierde significancia estadística (2 y 4). Este resultado es consistente con lo encontrado a nivel de estrato, en donde la diferencia entre el estrato 0 y 1 no es estadísticamente significativa, mientras que, sí se encuentran pruebas de que los estratos 2, 3 y 4, cuentan con mayores resultados respecto al estrato 0. En cuanto a la edad, se encuentra que en estudiantes con un año más al promedio, el resultado global del ICFES disminuye 7.2 puntos en promedio; este resultado podría explicarse debido a que en promedio la edad es de 17 años, por tanto, un año adicional puede indicar estudiantes que se encuentran repitiendo año electivo, por bajo desempeño académico. En relación al sexo, en Jamundí las mujeres obtienen 12 puntos menos que los hombres. Finalmente, en ninguno de los 4 modelos estimados, la discapacidad arrojó un resultado estadísticamente significativo.

Como se observó en la sección de datos, varios corregimientos no cuentan con el ciclo completo de bachillerato, lo que implica que deben desplazarse de su hogar al colegio. Para capturar este efecto, se clasificó a los estudiantes que se encuentran a más de 5.5 kilómetros del colegio (4 quintil de la distribución). De esta variable, se encontró un signo positivo en los 4 modelos estimados; sin embargo, es posible la presencia de problemas de endogenidad producto de la ubicación espacial del municipio, en donde los barrios de mayores ingresos no cuentan con la presencia de colegios. Al interactuar esta variable con el INSE se corrige este problema, encontrándose que estudiantes con el mismo nivel socioeconómico, al estar lejos tienen un impacto negativo de 1,03 puntos. Esto se explica, porque los costos incurridos en tiempo y tipo de transporte, pueden ser compensados por transporte particular, mejor alimentación, mejores condiciones de internet, etc., propias de estudiantes pertenecientes a familias de altos ingresos. Finalmente, siguiendo a Lin, Chieh y Li (2014) [13] se cruzó la edad de los estudiantes y la distancia, encontrando que al tener la misma edad, la distancia afecta negativamente el desempeño académico, explicado en que, a mayor edad los estudiantes deben asumir su transporte a los colegios de manera

independiente, donde muchos lo hacen en transportes no motorizados o servicios públicos.

Para evaluar los 4 modelos estimados, como se explicó al inicio de esta sección, se descarta la utilización de la estimación por MCO, dada la estructura de los datos jerárquicos con los que se cuenta para este problema. Adicional a esto, a las estimaciones por REML se les realizó una prueba de bondad, la cual comprueba si la varianza del componente aleatorio es 0, lo que mostraría la necesidad de descartar la estimación multinivel; en esta prueba se rechazó la hipótesis nula para los modelos 3 y 4, aceptando la utilización de la modelación multinivel. Una vez se tuvo este resultado, la elección del modelo se redujo en escoger entre los modelos 3 y 4. Para esta labor, se realizó la prueba de Ruppert, Wand, y Carroll (1973), la cual compara modelos con pendiente aleatoria y un modelo de intercepto aleatorio; acorde al resultado se decidió utilizar el modelo sin pendiente aleatoria, descartando el efecto supuesto en las variables distancia e INSE. No obstante, se resalta que el modelo 4 parece indicar que contar con profesores más educados puede corregir los efectos adversos de la distancia. Finalmente, se realizó una comparación de los 4 modelos mediante el criterio Akaike's Information Criterion (AIC) (Akaike, 1987), bajo el cual se escoge el tercer modelo.³

Cuadro 3: Regresión modelo completo

	(1) P.Global	(2) P.Global	(3) P.Global	(4) P.Global
Promedio INSE	3.045*** (3.59)	3.319*** (3.76)	2.759** (2.21)	3.078** (2.42)
Zona urbana	14.50*** (3.04)	14.43*** (2.86)	16.64** (2.12)	16.44** (2.04)
% Profesores Posgrado	11.26* (1.96)	9.770* (1.68)	9.897 (1.05)	7.575 (0.79)
% Profesores 40 años	-40.02*** (-4.28)	-80.29** (-2.03)	-34.61*** (-2.82)	-85.06** (-2.14)
INSE	0.398** (1.98)	-0.306 (-0.43)	0.385** (1.96)	-0.506 (-0.73)
Edad	-7.714*** (-5.34)	-7.683*** (-5.32)	-7.532*** (-5.35)	-7.496*** (-5.33)
Femenino	-12.34*** (-4.81)	-12.20*** (-4.76)	-12.41*** (-4.96)	-12.24*** (-4.90)
Lejos	225.2** (2.38)	212.5** (2.24)	227.9** (2.47)	220.6** (2.39)
Lejos*INSE	-1.213** (-2.25)	-1.042* (-1.91)	-1.036* (-1.95)	-0.923* (-1.73)
Edad*lejos	-10.09** (-2.01)	-10.62** (-2.12)	-10.70** (-2.19)	-11.26** (-2.30)
Discapacitado	-85.70 (-0.45)	-79.96 (-0.42)	-141.1 (-0.75)	-136.0 (-0.73)

³Revisar anexos

Cuadro 4: Continuación Cuadro 3

	(1) P.Global	(2) P.Global	(3) P.Global	(4) P.Global
Discapacidad*INSE	0.706 (0.17)	0.570 (0.14)	1.830 (0.47)	1.717 (0.44)
Estrato 0	-12.81 (-1.53)	-12.32 (-1.48)	-13.08 (-1.60)	-12.63 (-1.55)
Estrato 2	8.654*** (2.86)	8.756*** (2.90)	7.858*** (2.67)	7.944*** (2.70)
Estrato 3	8.534 (1.64)	8.993* (1.73)	8.517* (1.68)	8.939* (1.77)
Estrato 4	-23.59 (-1.13)	-23.46 (-1.13)	-25.86 (-1.27)	-25.85 (-1.28)
Estrato 5	19.43 (0.76)	19.40 (0.76)	14.05 (0.56)	13.99 (0.56)
Jornada única	5.108 (1.44)	5.863 (1.64)	6.025 (0.96)	6.941 (1.11)
% Profesores 40 años * INSE		0.889 (1.00)		1.147 (1.31)
Distancia * % Profesores posgrados		27.42** (2.03)		25.83* (1.84)
Constante	227.6*** (4.86)	245.6*** (4.91)	233.5*** (3.75)	256.7*** (3.98)
Random				
Intercepto β_0			42.504 (28.11)	42.67 (32.57)
Residuales σ^2				
			1213.70 (60.07)	(59.73)
Random Distancia				
$\gamma_{1,1}$				3.12e-11 (5.09e-10)
Random INSE				
$\gamma_{2,1}$				4.62e-07 (0.00014)
LR test vs.linear model (chi2):			13.33***	13.15***
AIC y BIC				
N	831	831	831	831
ll(null)	-4267.9	-4267.9	.	.
ll(model)	-4143.5	-4140.9	-4136.8	-4134.4
df	19	21	21	25
AIC	8324.9	8323.9	8315.6	8318.8
BIC	8414.6	8423.1	8414.7	8436.8
N	831	831	831	831
adj. R^2	0.258	0.263		

t statistics in parentheses

* p<.10, ** p<.05, *** p<.01

Cuadro 5: Prueba de pendiente aleatoria vs intercepto aleatorio

Likelihood-ratio test LR $\chi^2(4) = 4.79$
(Assumption: est3 nested in est4) Prob $>\chi^2 = 0.3099$
Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

Conclusiones

Este trabajo es una aproximación a los determinantes del desempeño académico, el cual nos permitió validar algunas hipótesis que explican la dispersión observada en los resultados de los estudiantes de instituciones oficiales en Jamundí. Puntualmente, se encontró que el ambiente escolar influye en los resultados académicos, donde un colegio con mejores resultados en promedio, influye en el resultado del estudiante; por otra parte, los estudiantes reciben una prima positiva por estudiar en colegios de la zona urbana y los profesores de más de 40 años podrían generar retrocesos en los resultados académicos. Esta última hipótesis podría ser un tema nuevo de investigación, pues la literatura valida que la mayor experiencia de los profesores se traduce en mejores técnicas de aprendizaje, por tanto se debe revisar qué impide que esta experiencia no se traduzca en mejores resultados.

Por otro lado, a nivel de estudiante los hallazgos muestran la incidencia de la distancia en los resultados académicos de acuerdo a niveles de INSE y edad de los estudiantes. Este resultado es común en la literatura, pero ya que se encontró para Jamundí, vale la pena preguntarse por soluciones que puedan reducir estos impactos por medio de sedes cercanas a los estudiantes o plantear sistemas de transporte eficientes y con mayor cobertura, debido a que, actualmente el PAE no beneficia a alumnos de grados superiores. A pesar de los resultados obtenidos, aún es necesario entender los mecanismos por los cuales la distancia incide negativamente con base en la edad y el INSE, por tanto se propone para próximas investigaciones utilizar como variable alternativa el tiempo que recorren los estudiantes a los colegios, o el método de transporte que utilizan. Finalmente, se requiere contar con una medida más amplia del desempeño académico, puesto que, el resultado de una prueba estandarizada no es la única medida para captar esta variable, y según la literatura, deja por fuera otros aspectos a tener en cuenta para una mejor aproximación.

Bibliografía

Reportes técnicos

- [23] Eduardo Velez, Ernesto Schiefelbein y Jorge Valenzuela. *Factores que Afectan el Rendimiento Académico en la Educación Primaria Revisión de la Literatura de América Latina y El Caribe**. 1994.

Reportes sin publicar

- [1] Escobar J. A. y col. «Una aproximación al sector educativo municipal, caracterización de los alumnos y las sedes educativas en Jamundí». 2021.

Artículos

- [2] James D. Allen. «Grades as Valid Measures of Academic Achievement of Classroom Learning». En: *The Clearing House: A Journal of Educational Strategies, Issues and Ideas* 78 (5 mayo de 2005), págs. 218-223. ISSN: 0009-8655. DOI: 10.3200/tchs.78.5.218-223.
- [3] Su Jung Choi, Jin Chul Jeong y Seoung Nam Kim. «Impact of vocational education and training on adult skills and employment: An applied multilevel analysis». En: *International Journal of Educational Development* 66 (abr. de 2019), págs. 129-138. ISSN: 07380593. DOI: 10.1016/j.ijedudev.2018.09.007.
- [4] J. E Coleman y col. «Equality of Educational Opportunity». En: *Us Department of Health, Education, and Welfare: U.S Government Printing office* (1966).
- [6] Lutz Erbring y Alice A. Young. «Individuals and Social Structure: Contextual Effects as Endogenous Feedback». En: *Sociological Methods & Research* 7 (4 1979), págs. 396-430. ISSN: 15528294. DOI: 10.1177/004912417900700404.
- [9] Duncan Greg J. y col. «Supplemental Material for School Readiness and Later Achievement.» En: *Developmental Psychology* 43 (2007). ISSN: 0012-1649. DOI: 10.1037/[0012-1649.43.6.1428].supp.
- [10] M. Jiménez. «Competencia social: intervención preventiva en la escuela.» En: *Infancia y Sociedad*. 24 (2000), págs. 21-48.
- [11] P. R. Kaufman. «Rural Poor have Less Access to Supermarkets, Large Grocery Stores». En: *Rural Development Perspectives* 13 (3 1999), págs. 19-25.

- [12] Nizamettin Koc y Bekir Celik. «The Impact of Number of Students per Teacher on Student Achievement». En: *Procedia - Social and Behavioral Sciences* 177 (abr. de 2015), págs. 65-70. ISSN: 18770428. DOI: 10.1016/j.sbspro.2015.02.335.
- [13] Jen Jia Lin, Yun Chieh Huang y Chia Li Ho. «School accessibility and academic achievement in a rural area of Taiwan». En: *Children's Geographies* 12 (2 abr. de 2014), págs. 232-248. ISSN: 14733285. DOI: 10.1080/14733285.2013.812308.
- [14] Mariem Liouaeddine y Mohammed Bijou. «Multilevel modeling an innovative tool for analyzing clustered data: application to education and its projection on environmental management». En: *E3S Web of Conferences* 319 (2021), pág. 01008. DOI: 10.1051/e3sconf/202131901008.
- [15] Oliver Lüdtke y col. «Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling». En: *Contemporary Educational Psychology* 34 (2 abr. de 2009), págs. 120-131. ISSN: 0361476X. DOI: 10.1016/j.cedpsych.2008.12.001.
- [16] G. F. Nemet y A. J. Bailey. «Distance and Health Care Utilization among the Rural Elderly». En: *Social Science & Medicine* 50 (9 2000), págs. 1197-1208.
- [17] Joseph Sunday Owoeye y Philias Olatunde Yara. «School location and academic achievement of secondary school in Ekiti State, Nigeria». En: *Asian Social Science* 7 (5 2011), págs. 170-175. ISSN: 19112017. DOI: 10.5539/ass.v7n5p170.
- [18] Dylan L. Robertson y Arthur J. Reynolds. «Family profiles and educational attainment». En: *Children and Youth Services Review* 32 (8 ago. de 2010), págs. 1077-1085. ISSN: 01907409. DOI: 10.1016/j.childyouth.2009.10.021.
- [19] Samuel Robison y col. «Correlates of educational success: Predictors of school dropout and graduation for urban students in the Deep South». En: *Children and Youth Services Review* 73 (feb. de 2017), págs. 37-46. ISSN: 01907409. DOI: 10.1016/j.childyouth.2016.11.031.
- [20] Liliana Rojas-Guyler, Judy Murnan y Randall R. Cottrell. «Networking for Career-Long Success: A Powerful Strategy for Health Education Professionals». En: *Health Promotion Practice* 8 (3 2007), págs. 229-233. ISSN: 15248399. DOI: 10.1177/1524839907302735.
- [22] E. Talen. «School, Community, and Spatial Equity: An Empirical Investigation of Access to Elementary Schools in West Virginia». En: *Annals of the Association of American Geographers* 91 (3 2001), págs. 465-486.

Libros

- [5] Jacques Delors. *Informe a la UNESCO de la Comisión Internacional sobre la educación para el siglo XXI, presidida por*. Editorial Santillana, 1996, págs. 1-46. ISBN: 84-294-4978-7.
- [7] J. J. Hox. *Applied multilevel analysis*. TT-publikaties, 1994. ISBN: 9080107328.

- [8] Joop J. Hox. *Multilevel Analysis Techniques and applications*. Segunda. Routledge, 2010. ISBN: 0-203-85227-3.
- [21] *Schools Count: World Bank Project Designs and the Quality of African Primary Education*. 59.^a ed. World Bank, 1996. URL: <http://hdl.handle.net/10986/9974>.

Anexos

Anexo 1. Efectos aleatorio β_0

Figura 5: Correlación entre la distancia que afrontan los estudiantes a sus colegios y el logro académico desagregado por la proporción de profesores mayores a 40 años

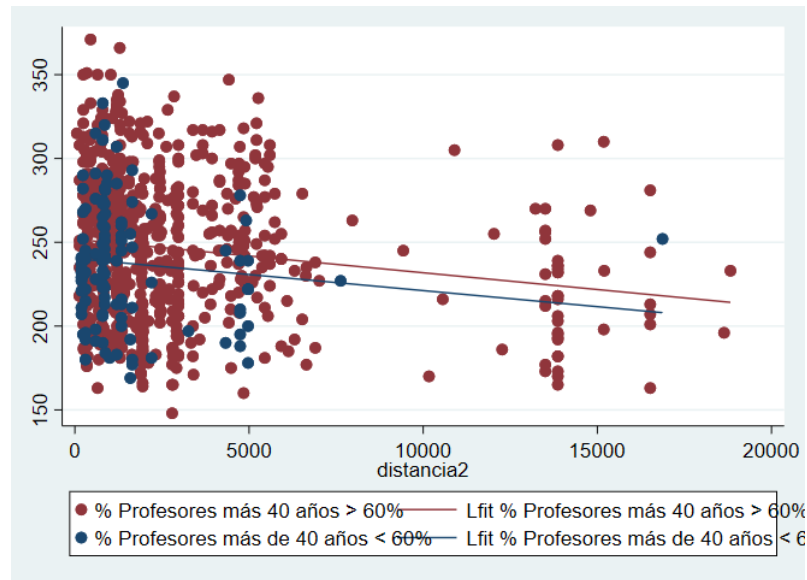
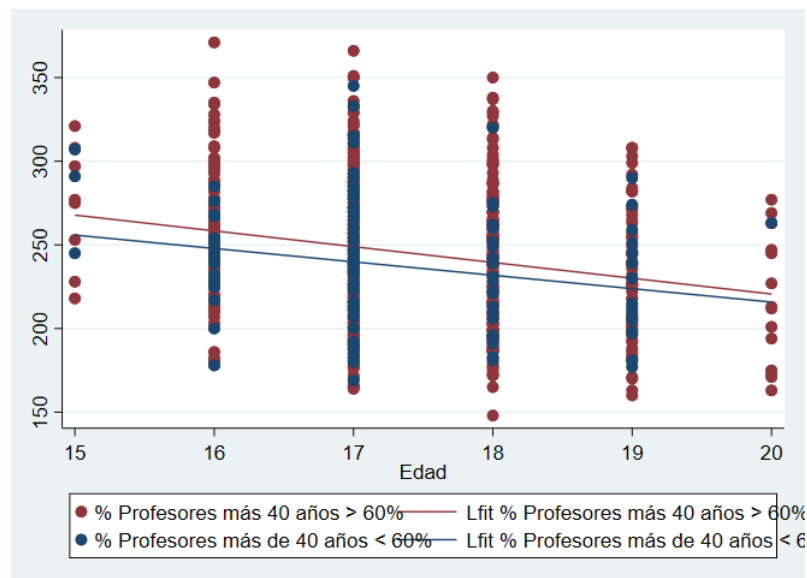


Figura 6: Correlación entre Edad de los estudiantes y el logro académico desagregado por la proporción de profesores mayores a 40 años



Anexo 2. Efectos aleatorio β_i

Figura 7: Correlación entre el INSE y el logro académico desagregado por la proporción de profesores mayores a 40 años

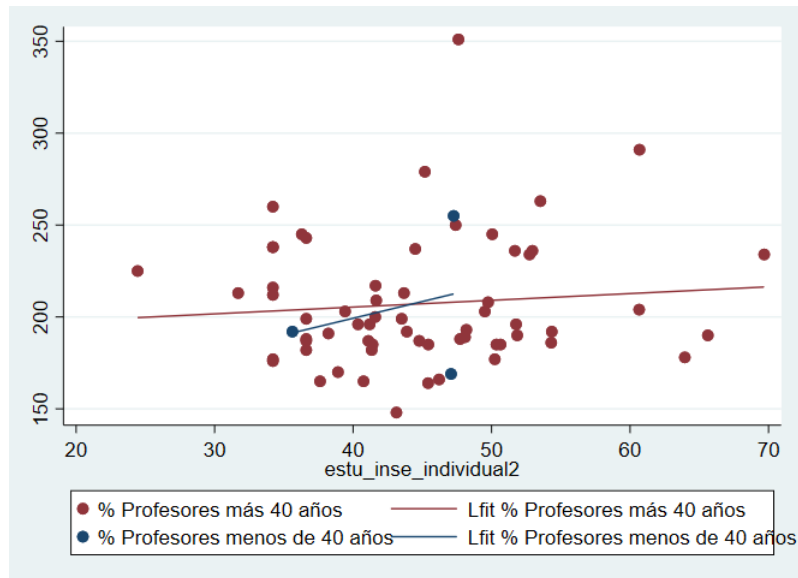
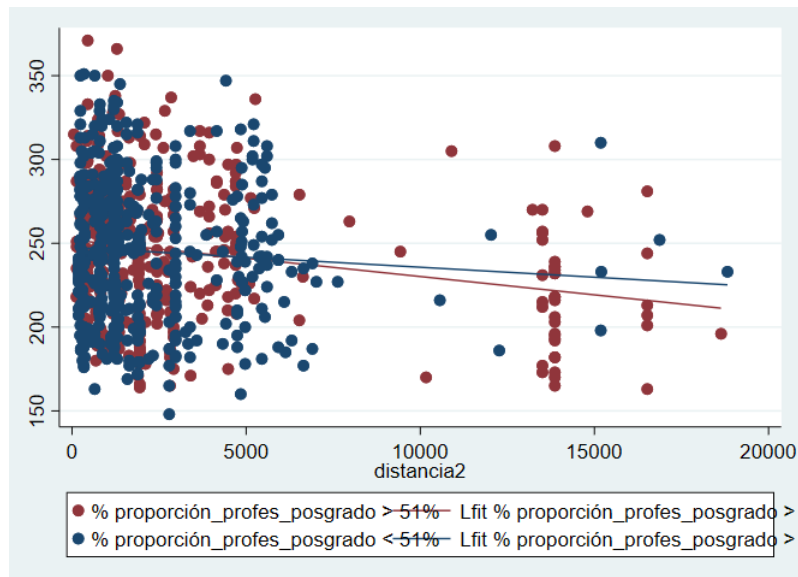


Figura 8: Correlación entre la distancia que afrontan los estudiantes a sus colegios y el logro académico desagregado por la proporción de profesores mayores a 40 años



Anexo 3. Pruebas de los modelos estimados por MCO

Figura 9: Prueba de normalidad modelo 1

Skewness/Kurtosis tests for Normality					
Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	joint chi2(2)	Prob>chi2
rs1	831	0.0052	0.2433	8.81	0.0122

Figura 10: Prueba de normalidad modelo 2

Skewness/Kurtosis tests for Normality					
Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	joint chi2(2)	Prob>chi2
rs2	831	0.0061	0.3828	8.07	0.0177

Figura 11: Prueba de heterocedasticidad modelo 1

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity		
Ho: Constant variance		
Variables: fitted values of PuntajeGlobal		
chi2(1)	=	6.18
Prob > chi2	=	0.0129

Figura 12: Prueba de heterocedasticidad modelo 2

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity		
Ho: Constant variance		
Variables: fitted values of PuntajeGlobal		
chi2(1)	=	7.37
Prob > chi2	=	0.0066