



**FACTORES DETERMINANTES EN EL RENDIMIENTO DE LOS
ESTUDIANTES DE LA REGIÓN PACÍFICO-COLOMBIANA EN LAS PRUEBAS
SABER-11**

PROYECTO DE GRADO DE PREGRADO

MICHAEL STEVEN MORALES CEBALLOS

DIRECTOR DEL PROYECTO

CRISTIAN CAMILO HOYOS

UNIVERSIDAD ICESI

FACULTAD DE CIENCIAS ADMINISTRATIVAS Y ECONÓMICAS

PROGRAMA DE ECONOMÍA Y NEGOCIOS INTERNACIONALES

SANTIAGO DE CALI

2019

CONTENIDO

1.	INTRODUCCIÓN	6
1.1.	Importancia de la Educación	6
1.2.	¿Cómo está la región Pacífico en educación?.....	7
1.3.	Propósito del estudio	9
2.	OBJETIVO	11
3.	MARCO TEÓRICO	12
3.1.	Literatura internacional	12
3.2.	Literatura nacional	13
4.	DATOS	18
4.1.	Tratamiento de datos	18
4.2.	Variables	18
4.3.	Estadísticas Descriptivas	19
5.	METODOLOGÍA	25
5.1.	Especificación del modelo	25
6.	RESULTADOS	26
6.1.	Variables personales.....	26
6.2.	Variables familiares	27
6.3.	Variables del plantel.....	29
6.4.	Problemas econométricos.....	30
7.	CONCLUSIONES Y RECOMENDACIONES	32
8.	ANEXOS	33
9.	BIBLIOGRAFÍA	36

Listado de Anexos, Tablas y Gráficos

1. Tablas

Tabla 1. Lista de variables por categorías

2. Gráficos

Gráfico 1. Distribución de departamentos por periodo

Gráfico 2. Distribución de sexo por departamentos

Gráfico 3. Distribución de ubicación geográfica de los colegios

Gráfico 5. Distribución de puntaje global por periodos

3. Anexos

Anexo 1. Tabla de regresión

RESUMEN

La educación en la región pacífico-colombiana es un tema de interés general. Por lo tanto, en el presente trabajo se aborda este tema desde el rendimiento en los departamentos Cauca, Chocó, Nariño y Valle en cuanto a las pruebas de Estado SABER 11. Teniendo en cuenta que desde el 2015 estas pruebas arrojan un resultado global, por medio de una regresión múltiple lineal se buscó hallar cuáles eran las características que hacían que a un estudiante de la región tuviera un mayor desempeño en las pruebas.

Las variables se obtuvieron de la base de datos del Instituto Colombiano para el Fomento de la Educación Superior en Colombia (ICFES) en las cohortes semestrales entre 2017-I a 2018-II. Todas las variables fueron estadísticamente significativas, aunque no todas las categorías dentro de cada una de las variables. Finalmente, se encontró que la dedicación a la lectura, al internet, la ubicación geográfica del colegio y si este es bilingüe están entre las principales características que implican buenos resultados en esta prueba.

Por último, se hace una invitación a todos los lectores del documento a seguir investigando desde sus áreas de conocimiento como impactar positivamente la educación en la región pacífico de Colombia, teniendo en cuenta que los estudios actuales se basan en una coyuntura pasada de las pruebas y no en el puntaje global como la variable dependiente de los diferentes modelos y estudios que se realizan.

Palabras claves: Educación, Saber 11, Regresión Lineal Múltiple, Colombia.

ABSTRAC

Education in the Pacific-Colombian region is a topic of general interest. Therefore, this paper addresses this issue from the perspective of performance in the departments of Cauca, Chocó, Nariño and Valle in terms of the SABER 11 State tests. Taking into account that since 2015 these tests give a global result, by means of a linear multiple regression we sought to find what were the characteristics that made a student in the region perform better in the tests.

The variables were obtained from the database of the Colombian Institute for the Promotion of Higher Education in Colombia (ICFES) in the semester cohorts between 2017-I and 2018-II. All variables were statistically significant, although not all categories within each variable. Finally, it was found that dedication to reading, the internet, the geographic location of the school and whether the school is bilingual are among the main characteristics that imply good results in this test.

Finally, all readers of the document are invited to continue investigating from their areas of knowledge how to positively impact education in the Pacific region of Colombia, bearing in mind that current studies are based on a past juncture of the tests and not on the overall score as the variable dependent on the different models and studies that are conducted.

Key words: Education, SABER 11, Multiple Linear Regression, Colombia.

1. INTRODUCCIÓN

1.1. Importancia de la Educación

Es importante entender que en el contexto mundial “la educación es un importante motor del desarrollo, así como uno de los instrumentos más eficaces para reducir la pobreza y mejorar la salud, y lograr la igualdad de género, la paz y la estabilidad” (Banco Mundial, 2017). Dada esta importancia, muchos estudios se han realizado en torno a este tema. Por ejemplo, el Banco Mundial que es un actor fundamental en la promoción de la educación a nivel global, realiza constantemente mediciones e investigaciones con el ánimo de ofrecer datos que sirvan para hacer distintos análisis tanto globales como por sectores demográficos.

Además, se debe considerar que la educación brinda herramientas y capacidades que llevan a mejorar las expectativas de calidad de vida de las personas.

La educación conlleva beneficios considerables y sistemáticos en materia de ingresos, y contrarresta el aumento de la desigualdad. En el caso de las personas, promueve el empleo, los ingresos, la salud y la reducción de la pobreza. En el caso de las sociedades, contribuye al desarrollo económico a largo plazo, promueve la innovación, fortalece las instituciones y fomenta la cohesión social. (Banco Mundial, 2017).

Aparte de aspectos de calidad de vida, la educación a nivel mundial es un derecho que ha cobrado gran importancia, tanto que las diferentes naciones en el escenario internacional, en especial las que son estados sociales de derecho, han volcado sus políticas de gobierno en torno a sus sistemas de educación nacional. Una muestra de esto es la radicación de la educación como un derecho dentro de la declaración de los derechos humanos en su artículo 26 de 1948 (Vincezi & Tudesco, 2009). Es por lo anterior, que tanto a nivel mundial como en Colombia la educación ha

sido un tema sujeto de constantes investigaciones que tratan de encontrar evidencias para poder dar soporte a las diferentes políticas públicas que buscan impactar esta área en las comunidades.

1.2. ¿Cómo está la región Pacífico en educación?

Para comenzar el análisis de la educación en la región pacífica de Colombia, se podría prever que existe una notable disparidad entre los datos de la región. Por ejemplo, si comparamos en cuestión de Educación Superior, el Valle del Cauca cuenta con cuatro universidades acreditadas por el ICETEX como Instituciones con Acreditación de Alta Calidad mientras que el Cauca solo cuenta con una y Chocó junto con Nariño no cuentan con alguna (Estos datos son obtenidos de la publicación oficial del ICETEX de Instituciones con Acreditación de Alta Calidad).

En cuanto a la educación básica y media hay que aclarar que desde el año 2012 es gratuita en el sector público. Esto es con el ánimo de cumplir el Pacto Internacional de Derechos Económicos Sociales y Culturales -PIDESC- en el cual Colombia hace parte de los Estados firmantes (Colprensa, 2011). Además, a partir del 2018 se implementó en los colegios oficiales la jornada única de estudio en donde se busca aumentar el tiempo dedicado a las actividades académicas y mejorar la calidad educativa de los establecimientos, entre otros objetivos. (Ministerio de Educación Nacional, 2018).

Sin embargo, aunque en el papel garantizar la educación de calidad en el país ha sido un eje a seguir para los diferentes gobiernos, en la región del pacífico no se ha hecho mayores cambios. Una de las razones que se adjudican a esta falta de cumplimiento con estas promesas ha sido la corrupción que, aunque afecta a todo el país, es muy notoria en esta región. “Buenaventura, Tumaco y Quibdó, principales ciudades de esta región del país, de acuerdo con el promedio de sus instituciones educativas, se ubicaron generalmente por debajo de la media nacional.” (Mayolo, 2010). Mayolo argumenta que la mala administración tanto local como nacional de los recursos

públicos ha provocado que los planteles carezcan de plantas físicas y herramientas idóneas para el desarrollo cognitivo.

Continuando con el análisis de la educación de la región pacífica, otro indicador es la cobertura que se tiene frente al total de niños y jóvenes del país. Según Viveros (2017) del total de la cobertura nacional, tan sólo el 11,3% representa la región pacífica respecto al total nacional; el valle del cauca es el 33,8% respecto a la región. Junto con tasas de deserción estudiantil de carreras técnicas o tecnológicas que llegando a ser el 56,9% en la región. (Viveros, 2017)

Dado que la educación es “necesaria para promover la salud de los individuos y las comunidades... debe favorecer el acceso a la información, el desarrollo de las habilidades para la vida... y el empoderamiento de los individuos y las comunidades para actuar en defensa de su salud.” (Vincezi & Tudesco, 2009). Es por esta razón que los estudios sobre educación son tan relevantes para los tomadores de decisiones en cuestión de políticas públicas. Debido a que es una forma objetiva y centralizada de enfocar los recursos de la nación a programas que estén encaminados a contribuir de forma positiva y eficiente a las diferentes regiones del país.

Tal como lo dice Gómez (2002) las políticas públicas de educación se integra por el conjunto de planes, programa, estrategias y acciones gubernamentales que buscan hacer un cambio dentro del sistema con una finalidad única. En la mayoría de los casos sobre educación, se quiere un impacto positivo buscando la vinculación de organizaciones gubernamentales y privadas para la consecución de un objetivo en común. Además, en materia de políticas públicas, en cualquiera de sus áreas de acción, éstas “involucran conflictos entre los distintos actores de la comunidad”. (Espinoza, 2009)

Un ejemplo de lo anterior es el programa “Ser Pilo Paga”, que permitió que jóvenes de escasos recursos pudieran acceder a Instituciones de Calidad Superior del país. Este programa es una política pública del gobierno de turno que se basa en el puntaje de las pruebas SABER 11 como un requisito para la admisión en las universidades. Un desalentador indicador es que, del total del país, la región pacífica tan solo aportó el 15% de los pilos, y aún más desalentador que el Chocó solo aportó el 0.16% para el programa. Además, según El Tiempo (2016) “la región está 20 puntos por debajo del resultado de las pruebas SABER 11 respecto al promedio nacional y Chocó es el departamento con menor desempeño de la región”. De los municipios que enfrentan gran disparidad con respecto a la media nacional son Quibdó, Buenaventura y Tumaco con más de 30 puntos de diferencia.

Es por las anteriores cifras acerca de la precaria situación sobre las condiciones de la calidad de la educación de la región pacífica que el presente trabajo se va a concentrar en hallar cuáles son los factores asociados a los estudiantes que determinan los mejores resultados en las pruebas SABER 11.

1.3. Propósito del estudio

Para realizar este estudio la población objetivo serán los estudiantes de la región pacífica de Colombia que presentaron la prueba SABER 11 durante los años 2017 y 2018. Dado que a partir del 2015 esta prueba ya arroja un puntaje global, ésta será la variable dependiente de una regresión lineal múltiple por método Mínimo Cuadrados Ordinarios, tratando de seguir el modelo propuesto por Vegas y Petrow (2008) con variables que se categorizan como: personales, familiares y de plantel. Además, por cuestiones de medición del ICFES estas cuatro cohortes son las que presentan las mismas variables de interés.

En síntesis, en este trabajo se buscará los principales elementos que influyen en el rendimiento en las pruebas SABER 11, en la región pacífica colombiana conformada por los departamentos de Chocó, Cauca, Nariño y Valle del Cauca.

2. OBJETIVO

Determinar cuáles son los factores asociados a los estudiantes de la región pacífico-colombiana que más influye en el rendimiento de las pruebas de Estado SABER 11, usando la base de datos del Instituto Colombiano para la Evaluación de la Educación (ICFES).

2.1 OBJETIVOS ESPECÍFICOS

- Analizar y depurar las bases de datos de los resultados de las pruebas de estado SABER
- Emplear el método de regresión lineal múltiple para detectar las variables que influyen en el rendimiento en las pruebas de estado SABER 11.
- Realizar test para probar la no existencia de posibles problemas econométricos.
- Hacer conclusiones y recomendaciones con los datos hallados en el estudio

3. MARCO TEÓRICO

Existe una serie de sustentaciones teóricas del por qué las variables personales y familiares del estudiante como las del plantel en donde estudian generan una causalidad en los resultados de pruebas como la que realiza el Estado colombiano llamada SABER 11. Esta revisión bibliográfica se hará analizando primero los autores internacionales que han escrito sobre el tema y después los documentos escritos por colombianos. A su vez, se irá resaltando las variables independientes comunes para dar sustentación a sus diferentes trabajos. Lo anterior, con el fin de sintetizar los principales trabajos que existan relacionados con el tema a abordar.

3.1.Literatura internacional

Desde que Becker (1975) y Schultz (1981) hicieron aportes sobre la educación como una variable que influye en su teoría del capital humano, los diferentes trabajos que concentran sus esfuerzos por definir los factores que determinan el rendimiento académico de los estudiantes han aumentado en todo el mundo, tanto para niveles de educación básica como profesional.

Sin embargo, se podría considerar que el primero en interesarse en realidad en el análisis de estos factores determinantes de rendimiento fue Coleman en su informe para el presidente estadounidense Johnson “*Equality of Educational Opportunity*” en 1996 en donde toma los datos para Estados Unidos e intenta analizar la influencia de variables como: gasto por alumno, docentes por alumno, disponibilidad de libros, entre otras; dejándolas como variables a nivel personal, familiar y del plantel. Encuentra que las variables del plantel o colegio no son significativas para explicar los resultados en las pruebas. Por lo anterior, despertó interés en otros autores que, desde la evidencia empírica, notaban que la conclusión de Coleman era opuesta a lo que en realidad estaba ocurriendo.

Es así, como autores de diferentes partes del mundo plantean modelos con las variables socioeconómicas y familiares, que según Coleman influyen en el rendimiento de los estudiantes, no son las únicas en explicar estos resultados. Uno de ellos es Escardíbul (2008) quien usa variables personales del plantel y el esfuerzo, motivacionales, familiares y el uso de TICS. Ejemplos de estas son: actitud hacia la escuela, esfuerzo, mínimo de horas, clase social del padre, escolaridad de la madre; también usa variables de la escuela como los recursos disponibles, criterios de admisión y procesos de enseñanza; en cuanto a las TICS las variables son uso de un ordenador en diferentes etapas de la vida, el uso de computador dentro del hogar y la confianza que siente con el internet para realizar las tareas. Además, y aunque mencionaba que aún hay muchas similitudes en los diferentes centros que no permite obtener muchas variables significativas, encuentra que la autonomía del plantel y el reclutamiento de los docentes influyen de manera positiva en los resultados de las pruebas PISA para el caso de España.

Delpatro (1999) argumentó sobre las diferencias que existen entre las categorías de variables estableciéndolas como niveles. Estos niveles eran: estudiante, escuelas y jurisdicciones. Un aporte en su trabajo fue el contrastar las diferencias cuando hacen el análisis multinivel sin tener en cuenta las jurisdicciones. Encontró que el efecto del sector social en el rendimiento medio de la escuela es de mayor intensidad en las escuelas privadas. El PIB de las jurisdicciones posee un impacto superior al gasto privado en los dos años de estudio. Lo anterior, nos muestra como una corriente de investigadores empezaron a contrariar los resultados de Coleman en cuanto a la influencia de los planteles educativos en los resultados de las pruebas.

3.2.Literatura nacional

Uno de los primeros estudios que se realizó en el país sobre educación fue “Los Insumos Escolares en la Educación Secundaria y su Efecto Sobre el Rendimiento Académico de los

Estudiantes: Un estudio en Colombia” Realizado por Piñeros y Rodríguez (1998) en un informe para el Banco Mundial. Ellos usan el análisis multinivel para caracterizar las variables según: Nivel de la escuela, nivel del curso y nivel del estudiante. Sin embargo, desagregan esas variables para también tener en cuenta categorías como: características del alumno, del hogar, del aprendizaje, de los docentes y de los planteles. Para su estudio, usan la población de 1997 que realizaron la prueba Saber 11 (en su tiempo AC 972) en donde emplean dos modelos de regresión: uno llamado “vacío” con tres regresiones y otro “ampliado” con dos niveles. Para el primer modelo, destaca que hay una congruencia con la mayoría de los estudios en donde aseguran que la calidad de la educación de los colegios del sector privado es mayor a la ofrecida por los colegios oficiales. No obstante, cuando muestra los resultados del modelo ampliado encuentran que, eliminando el efecto de la situación socioeconómica, los colegios oficiales alcanzan puntajes superiores en todas las pruebas presentadas. Por lo que deja a la variable NSE (estratificación económica) como una variable importante para el estudio.

Otro de los documentos nacionales en los que se ha contrastado las variables sobre el plantel es “Determinantes de la Calidad de la Educación en Colombia” por parte de Gaviria & Barrientos (2001). En este texto realizado para FEDESARROLLO utilizan como población los estudiantes que realizaron las pruebas SABER 11 en el año 1999. Utilizaron un modelo de regresión lineal donde la variable dependiente es el puntaje del individuo. Para las variables independientes, utiliza un vector de características propias del individuo como sexo, edad y la educación de los padres (aquí sólo usa dos niveles, el nivel familiar está determinado dentro de las variables individuales¹) y otro vector para las características propias del plantel. Entre los

¹ Sin embargo, separa las variables personales por dos grupos: unas relacionadas con el nivel socioeconómico y otras relacionadas con la cantidad y calidad de las interacciones entre padres e hijos.

principales resultados, ellos destacan que los individuos que trabajan, en promedio, obtienen menores resultados y que los estudiantes con padres que ocupan cargos profesionales obtienen mejores resultados. Sin embargo, dentro de sus hallazgos descubren que, una vez controlado las incidencias de los factores socioeconómicos, los colegios públicos no inciden positivamente en los resultados.

A modo de ejemplo, actualmente podemos ver casos colombianos de estudios basados en la teoría de Coleman que han permitido evidenciar la presencia de intervención en los colegios públicos por parte del sector privado por medio de fondos para la financiación. Dado lo anterior, se han omitido las variables de los planteles para explicar los resultados de pruebas como matemáticas y lenguaje; como es el caso de Chica, Galvis y Ramírez (2011). En este mismo, encontraron que hay un mayor conjunto de variables que pueden explicar los resultados en matemáticas como la presencia de un computador en casa o la jornada en la que los estudiantes realizan sus estudios.

Por otra parte, el caso de (López, Posada, & Henao, 2014) quienes hacen uso de un análisis multinivel para explorar los diferentes niveles que representan los estudiantes agrupados en diferentes planteles educativos. Además, como en otros modelos que usan esta misma metodología parten de los supuestos de homoscedasticidad e independencia debido a que los individuos de un mismo colegio son más parecidos entre ellos que a individuos en otros colegios. Uno de los resultados obtenidos es la relación entre los resultados de la prueba con la situación económica tienen una relación positiva; para esto, usa el estrato socioeconómico como proxy de la situación económica. Otras variables que se tuvieron en cuenta y fueron determinantes fueron: acceso a nuevas tecnologías, diferencia entre colegio público o privado, cantidad de empleados

administrativos por estudiante en los colegios y el respaldo con el que cuentan los colegios de parte de actores gubernamentales.

Del mismo modo, (Orejuela, 2013) hace un análisis para el caso colombiano desde modelos jerárquicos lineales, en los cuales tiene en cuenta aproximadamente treinta y dos variables para explicar los resultados de las competencias de matemáticas y lectura en las pruebas de estado del país. Para poder cumplir con su análisis, usaron el Análisis de Componentes Principales debido a la interdependencia que hay en los diferentes niveles que existen en su análisis. De lo anterior, encontró que en los colegios públicos el factor no académico que más influye en los resultados es el índice socioeconómico del estudiante; en cambio, para los colegios privados, lo que más influye en el índice socioeconómico de los colegios. Dejando claro, que en el caso colombiano existen disparidades entre los colegios oficiales y no oficiales. Otro hallazgo de su estudio, es que no encuentra una diferencia estadísticamente significativa en el género de los estudiantes como determinante del rendimiento de las Pruebas de Estado Saber 11.

Otro aporte a la bibliografía nacional lo hace (Estrada, 2017) por medio de modelos lineales jerárquicos en una regresión por cuantiles, además de un modelo multinivel para describir el comportamiento del estudiante promedio. Para esto, organizó las variables en tres características: personales, familiares y del colegio. Otro aspecto para tener en cuenta durante su estudio es lo que llama heterogeneidad municipal el cual se le asocia con fuertes brechas para las regiones del país. Entre sus resultados, se encuentra la diferencia entre los puntajes obtenidos para los estudiantes que viven en las zonas metropolitanas del país (Bogotá, Cali, Medellín) siendo estos de categoría A con un desarrollo sostenido; además, encuentran también diferencias significativas a favor de los estudiantes de colegios privados en comparación con los de colegios oficiales.

Entre otros trabajos nacionales están el de Carvajal (2004) con un modelo de Componentes Principales y encontrando que hay independencia entre el rendimiento en matemáticas de los estudiantes de ingenierías de la UTP y los resultados que obtienen en las pruebas ICFES en las áreas de ciencias, física y matemáticas. Mina (2004) con un modelo de panel sin efectos, efectos fijos y efectos aleatorios para determinar el puntaje promedio de las categorías del ICFES por medio de características educativas; del plantel y socioeconómicas. Finalmente, Vallejo (2004) con modelos estadísticos descriptivos para hallar un patrón en la evolución de las notas académicas de la Universidad de los Andes.

4. DATOS

4.1. Tratamiento de datos

Como se mencionó al principio del documento, los datos fueron obtenidos del ICFES y corresponden a los resultados para los estudiantes de bachillerato de los cuatro departamentos de la región pacífico-colombiana de la prueba SABER 11.

Para poder hacer un análisis desagregado de las variables del núcleo familiar que inciden en los resultados del examen, se optó por tomar las cohortes desde el 2017-1 hasta el 2018-2. Principalmente porque estas cuatro cohortes contienen las mismas variables que son de interés para el estudio y que en su mayoría para años anteriores, no fueron capturadas por el ICFES.

4.2. Variables

Se tuvieron en cuenta las siguientes variables presentadas en la **Tabla 1** las cuales están categorizadas en tres: variables personales, familiares y del plantel. Estas se escogieron siguiendo la literatura mencionada en la sección anterior y según disponibilidad en la base de datos.

Tabla 1. Lista de variables por categorías

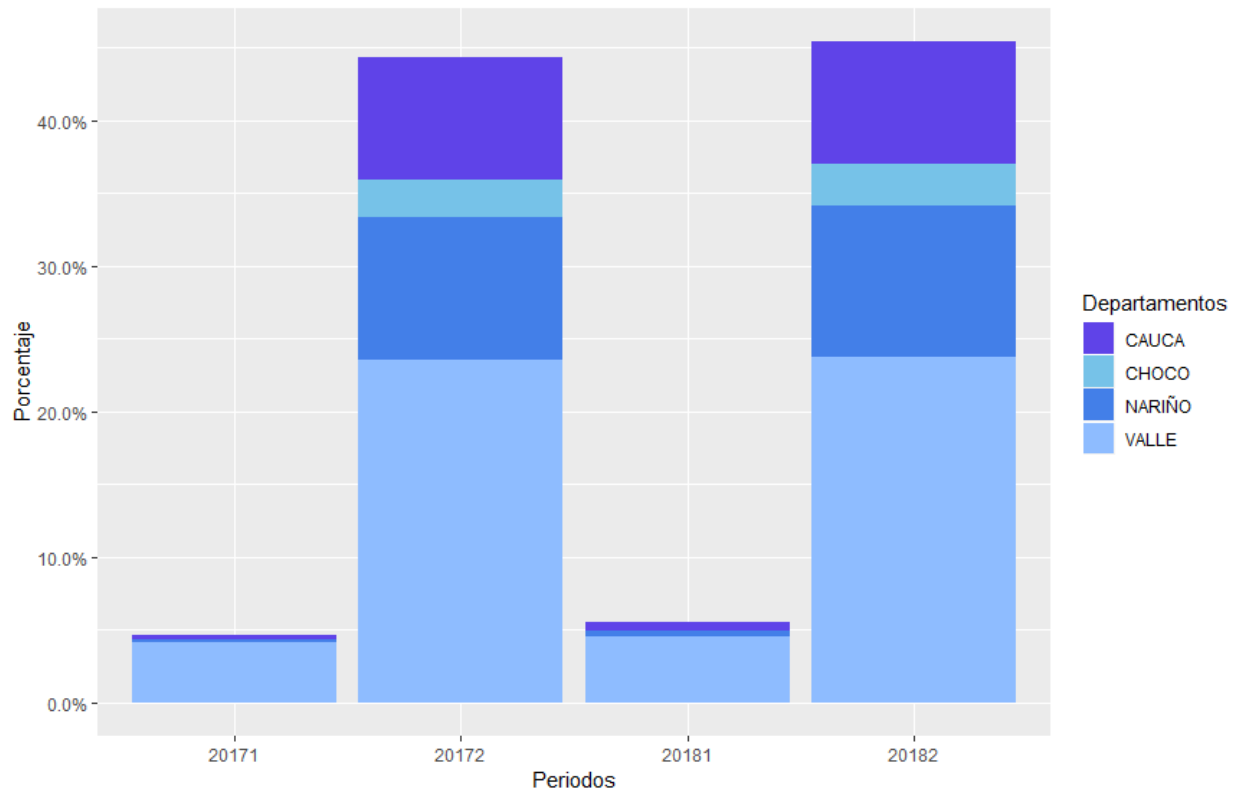
Personales	Familiares	Plantel
Genero	Personas_Hogar	Genero_Establ
Trabaja	Cuartos_Hogar	Naturaleza
Periodo	Edu_Padre	Bilingue
Tiene_Etnia	Edu_Madre	Caracter
Lectura_Diaria	Trabajo_Padre	Sede_Principal
Ded_Internet	Trabajo_Madre	Area_Ubicacion
punt_global	Computador	Jornada
Edad	Horno	Dpto_Ubicacion
-	Servicio_tv	-
-	Automovil	-
-	Motocicleta	-

Fuente: Elaboración propia

4.3. Estadísticas Descriptivas

Para empezar, la muestra tiene un total de 165.284 observaciones distribuidas en cuatro departamentos Cauca, Chocó, Nariño y Valle con una participación de 17.7%, 5.4%, 20.7% y 56.1% respectivamente. Sin embargo, como se puede observar en el **Gráfico 1**, para los períodos 2017-1 y 2018-1, el departamento del Chocó no presenta resultados y Cauca y Nariño tienen poca participación de la muestra, esto está explicado por qué en el primer semestre del año no presentan las pruebas los colegios públicos.

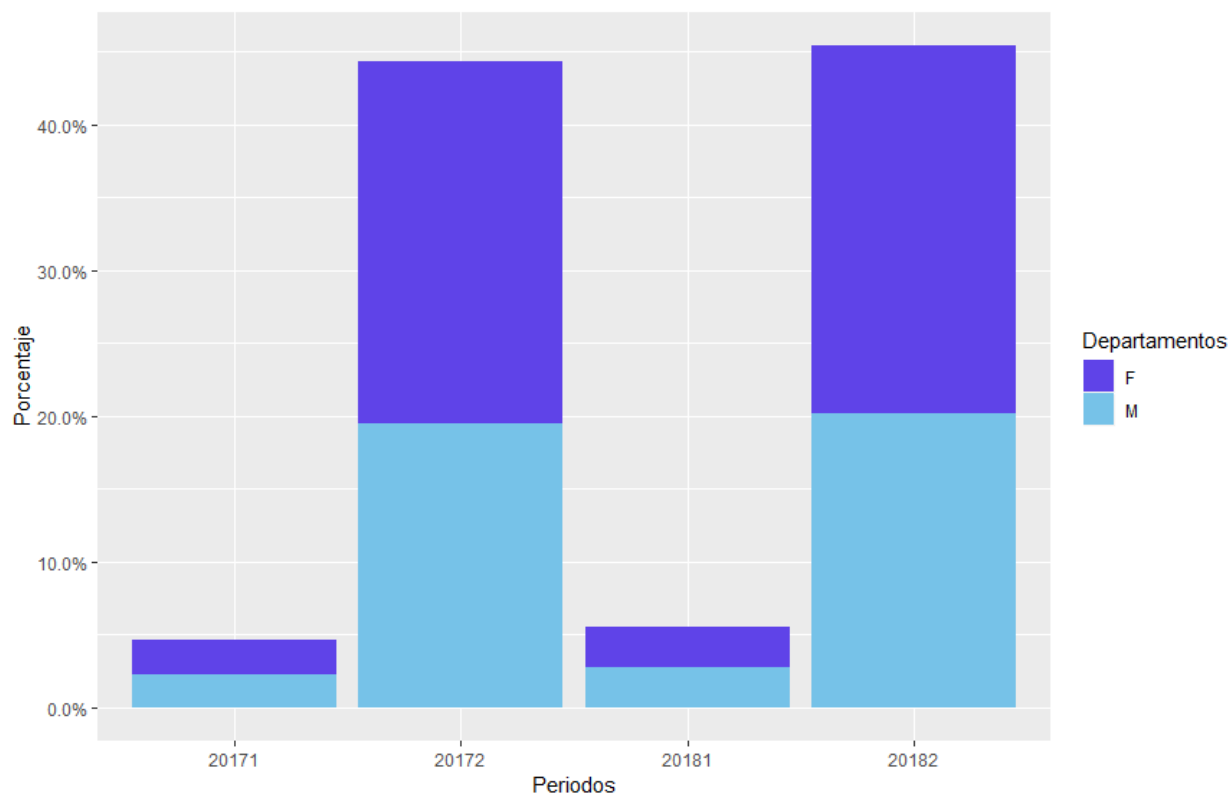
Gráfico 1. *Distribución de departamentos por periodo*



Fuente: Elaboración propia

En cuanto a las variables personales, tal como se puede observar en el **Gráfico 2** el 55.2% de la muestra indicó ser del género femenino. Asimismo, el 80% de la población manifestó no tener ninguna etnia. El 65.3% manifestó no estar trabajando en el momento de realizar la prueba. En cuanto a dedicación al internet, solo el 7.3% manifestaron no navegar en internet; en el otro extremo de la muestra, el 19% manifestaron dedicarle más de tres horas al internet. Por último, y contrario a la dedicación en internet, solo 4.3% de los estudiantes afirmaron leer más de 2 horas y 17.9% dijeron no leer por entretenimiento.

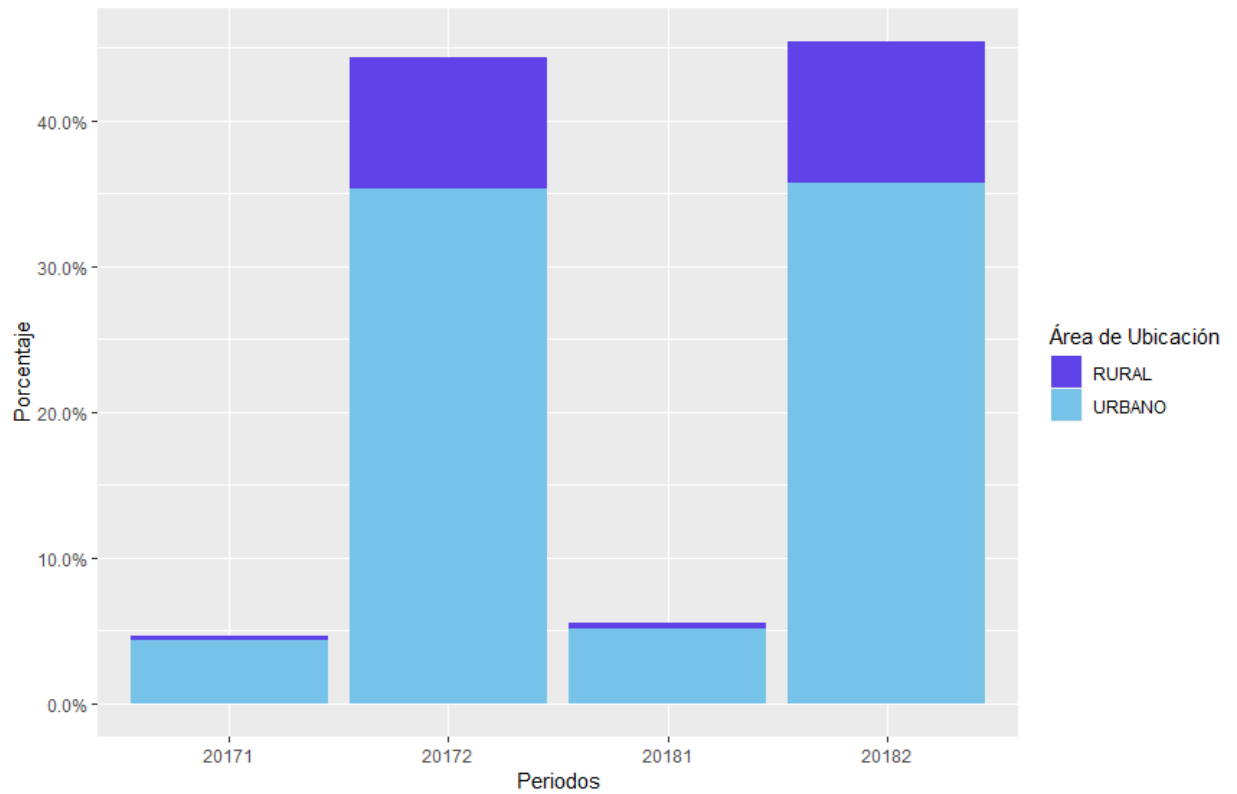
Gráfico 2. Distribución de sexo por departamentos



Fuente: Elaboración propia

En cuanto a las variables del plantel, 80.6% los estudiantes que realizaron la prueba, estudian en colegios ubicados en un área urbana. El 66.5% estudian en jornada de la mañana, seguidos del 10.7% que estudian en la tarde y el 8% tienen jornada completa; el resto, estudian de noche, los sábados o en jornada única. El 96.9% de los estudiantes provienen de colegios mixtos y el 74.2% estudiaban en colegios de naturaleza oficial (públicos). Por último, tan solo el 3.6% de los estudiantes provienen de colegios bilingües.

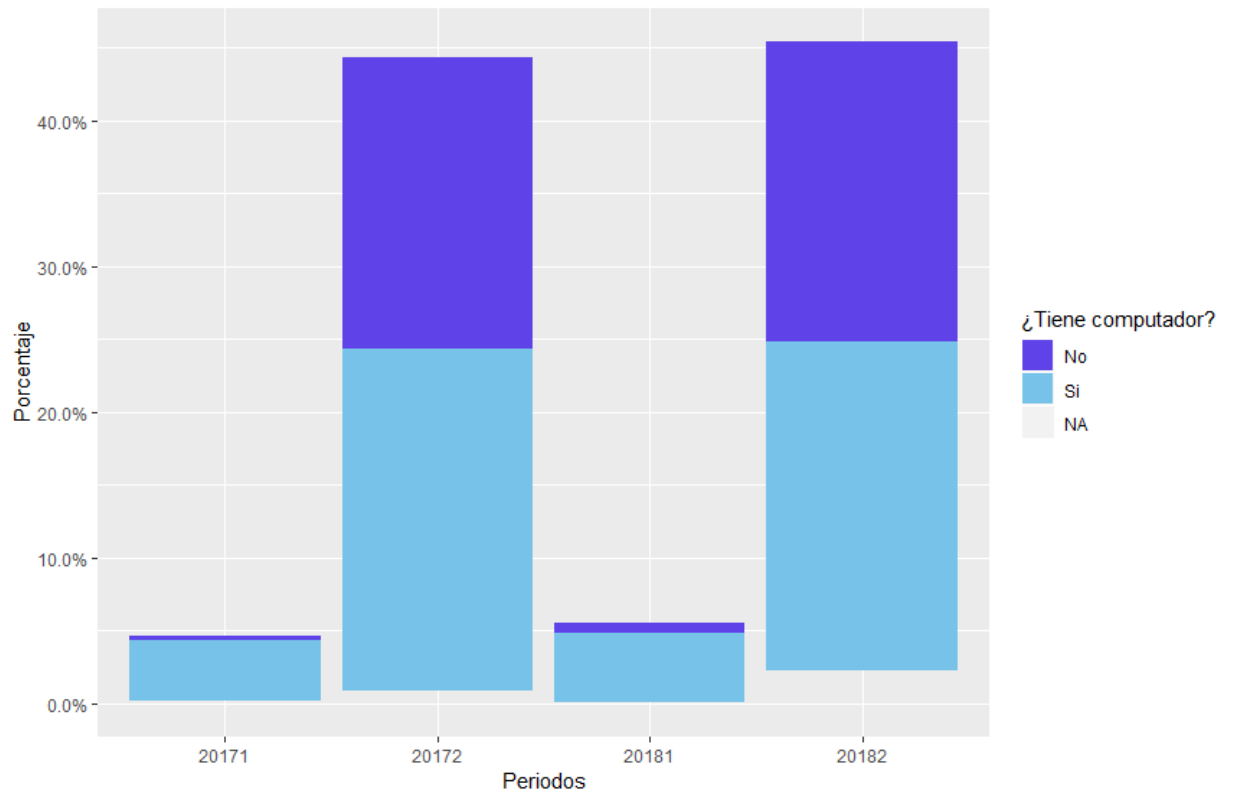
Gráfico 3. *Distribución de ubicación geográfica de los colegios*



Fuente: Elaboración propia

En lo que refiere a las variables familiares, aproximadamente el 50% de la muestra vive en un hogar conformado entre 3 a 4 personas, seguido de un 30% que viven entre 5 a 6 personas; tan solo el 3.6% vive en hogares con 9 o más personas. Además, es alarmante ver que el 43.1% de los estudiantes manifestaron no tener un computador en el hogar. Entre otras observaciones, el 57.6% no tienen un horno (ya sea microondas o de gas), el 73.2% tiene servicio de T.V en sus hogares y el 78.1% manifestaron que no hay un automóvil en su hogar.

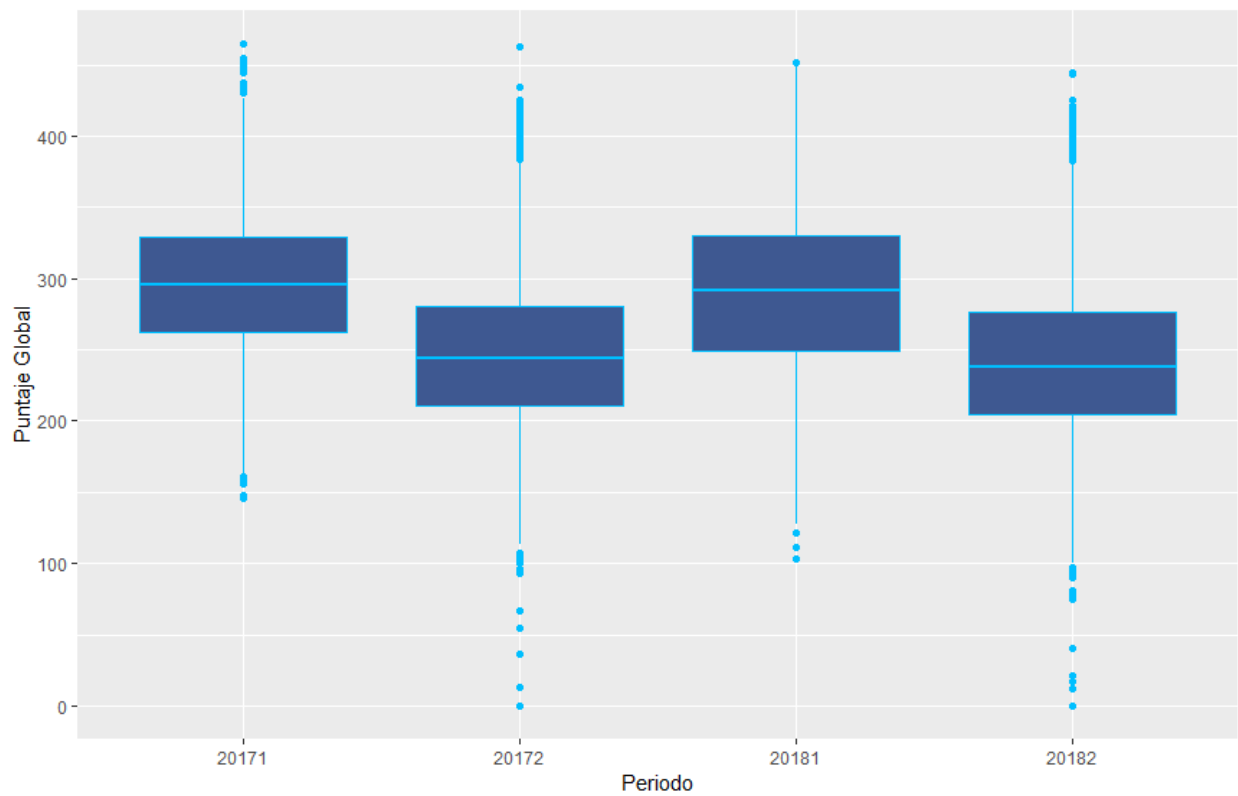
Gráfico 4. Porcentaje de evaluados que tienen computador en su hogar



Fuente: Elaboración propia

Finalmente, tal como se puede observar en el **Gráfico 3**, las distribuciones de los puntajes globales varían según el período en que se presentan las pruebas. A modo de hipótesis, podríamos decir que esto se debe a que los estudiantes de los colegios privados tienen una probabilidad más alta de obtener mejores resultados. Como se puede observar, el valor mínimo en los periodos 2017-1 y 2018-1 están por encima de los 100 puntos en el puntaje global; mientras que para los periodos 2017-2 y 2018-2 presentan datos, aunque atípicos, que se acercan mucho a cero.

Gráfico 5. *Distribución de puntaje global por periodos*



5. METODOLOGÍA

Para la metodología del trabajo se usará un Modelo de Regresión Lineal Múltiple tal como se mostrará en la siguiente subsección. Además, para elegir las variables que se incluyeron en el modelo se hizo uso de dos criterios. El primero es que las variables disponibles en la base de datos que se ciñeran a la revisión bibliográfica tanto nacional como internacional.

El segundo criterio fue por medio de un proceso de Stepwise usando el programa estadístico (R Graham J, McNeney B and Seillier-Moiseiwitsch F, 2005) para poder escoger las mejores variables explicativas dentro del modelo. Para esto, se desarrolló stepwise en los dos sentidos hacia adelante (forward) y hacía atrás (backward). Básicamente, ambos criterios se basan en el valor P de cada una de las variables independientes y las va agregando al modelo o quitándolas del mismo, respectivamente, hasta llegar al mejor modelo posible. Cabe aclarar que dista mucho de la minería de datos, porque las variables que principalmente se presentaron en la base de datos están sustentadas por el marco teórico.

5.1. Especificación del modelo

MODELO

$$R_i = \beta_0 + \beta_1 V_{1i} + \beta_2 V_{2i} + \dots + \beta_{33} V_{33i} + \varepsilon_i$$

Donde:

R_i = El puntaje global que obtiene un individuo en la región pacífico-colombiana

V_i = Es el vector de variables dentro del modelo mencionadas en el apartado de Datos.

ε_i = Es el error estándar del modelo.

$i = 1, 2, 3, \dots, 165.284$

6. RESULTADOS

6.1. Variables personales

El primer resultado determinante en esta investigación, es que todas las variables incluidas en el modelo fueron estadísticamente significativas, sin embargo, no todas las categorías lo fueron, dicho de otra forma, cuando se hace la transformación en variables dummy de cada una de las categorías de las variables como por ejemplo educación del padre o de la madre, las categorías educación del padre técnica o tecnológica completa y profesional incompleta para la madre no son estadísticamente significativas.

En cuanto a las variables personales, en promedio ser hombre en la región representa se esperan 11 puntos más en el puntaje global con respecto a ser mujer. También el modelo nos arrojó que las personas que tienen etnias, en promedio obtienen un resultado con 23 puntos menos, comparado con los individuos que manifiestan no pertenecer a alguna etnia. La dedicación en lectura diaria, congruentemente con lo que a priori se esperaba, los estudiantes que manifestaron dedicar más tiempo a la lectura esperan en promedio mejores resultados. Del mismo modo, los estudiantes que manifestaron no dedicarle tiempo a la lectura voluntaria, tienen en promedio 5 puntos menos en sus resultados con respecto a los que leen 30 minutos diarios.

Con respecto a la dedicación en internet, aquí no funciona de una forma tan clara como con la dedicación a la lectura. Tomando como base los estudiantes que dedican 30 minutos o menos tiempo al internet, todos los que dedican más tiempo obtienen resultados mayores. Sin embargo, en este caso, cuando manifiestan dedicarle más de tres horas diarias a internet, la tendencia se rompe y obtiene menores resultado.

Por último, en cuanto a las variables personales los estudiantes que manifiestan trabajar, en promedio obtienen 9 puntos menos en el puntaje global que los que no trabajan. Lo cual es de

esperarse porque tienen menos tiempo para dedicarse a sus estudios. Además, en la región pacífico, en promedio, por cada año más que se tenga al momento de presentar la prueba, los estudiantes tendrán en promedio una disminución de 1.5 puntos en el resultado global. Esto es de interés, por ejemplo, para los estudiantes que cursan sus estudios nocturnos o que retoman sus clases después de la mayoría de edad.

6.2. Variables familiares

Dentro de las variables familiares la educación del padre y de la madre del estudiante, la tendencia clara o notoria en los resultados. Es decir, no es tan sencillo como que a medida que se vaya completando etapas educativas, los hijos obtienen estadísticamente mejores o peores resultados tomando como referencia los hijos de padres con estudios profesionales completos. La única tendencia que se cumple y es de esperarse, es que los hijos de padres y madres con posgrados obtienen mejores resultados. Sin embargo, el que los padres tengan estudios profesionales incompletos, contrario a lo que se esperaría, genera que los hijos tengan levemente mejores resultados. Por el lado de la madre, esta categoría no es estadísticamente significativa.

Con base en lo anterior, haciendo comparaciones entre las mismas instancias de estudio, la educación de los padres tiene un más positivo en los resultados globales que la madre a excepción de que las madres no tengan ningún estudio o tenga una carrera técnica o tecnológica completa. Para estos casos, los aportes de la educación de las madres son superiores en el puntaje global de la prueba.

A la hora de hablar de los trabajos de los padres y madres, dado que se pueden catalogar los diferentes trabajos de modo ordinal, mas no cardinal, haré el análisis comparando el trabajo de mamá vs papá. Por ejemplo, el trabajo de la madre influye en los resultados mejores para casi todas las mismas categorías de trabajo comparando con el padre. Sin embargo, cuando los

evaluados no saben en que trabaja la mamá, o que esta sea pensionada o que esté en un puesto general, por lo general, sus resultados tendrán un puntaje inferior comparada con las mismas características para los padres. Cabe aclarar, que ambos resultados están siendo tomados con la base a que los padres y madres trabajen como pescador, agricultor o jornalero.

Otras variables familiares incluidas en el modelo y que son estadísticamente significativas nos arrojaron los siguientes resultados que podrían considerar los esperados. Las personas que tienen computador en su hogar obtienen en promedio un puntaje superior por 7 puntos en el resultado global y los que tiene automóvil en su hogar 2 puntos más. Sin embargo, dentro de los hallazgos en este estudio es que tener horno, servicio de tv y motocicleta en el hogar, estadísticamente hace que los resultados globales sean inferiores a las personas que dentro de la región no tienen ninguno de estos artículos en su hogar.

Siguiendo con el análisis, la hora de hablar de la acomodación dentro de los hogares, es decir, con cuantos cuartos cuenta y con cuantas personas comparte el hogar, encontramos que la tendencia con respecto a personas por hogar es que, a mayor número de personas, menor es el resultado esperado en el puntaje global de la prueba, siendo que los individuos que viven con nueve personas, su efecto marginal decreciente se vuelve negativo. Es decir, el único escenario en donde es mejor en un hogar de 1 o 2 personas, es con 9 personas o más.

Por último, con respecto a los cuartos por hogar, tomando como base tener 5 cuartos en el hogar, tener 4 o menos cuartos hace que haya un mejor resultado y tener 6 o más hace que el resultado esperado sea menor. Sin embargo, entre uno a cuatro cuartos no hay una tendencia estable. El estudio nos indica que los hogares con dos cuartos son los que mejores resultados obtendrán. Esto lo podríamos sustentar de modo que, a mayores cuartos, más personas y según el

número de personas a más personas, menor puntaje. Por lo que estaría siendo coherentes ambos resultados.

6.3. Variables del plantel

Muchos de los trabajos analizados en el marco teórico hacen la diferenciación por niveles de las diferentes características que son intrínsecas a un individuo. Por eso, las variables que están relacionadas con el plantel educativo son importantes para ver cuánto aporta estudiar en un colegio u otro y cuáles son los criterios en los que una persona debería basarse a la hora de escoger.

En cuanto al género del plantel educativo, los colegios que son masculinos tienden a tener mejores resultados, en promedio 10 puntos por encima que los femeninos. Además, los colegios mixtos son los que peores resultados tienen tendiendo a ser 11 puntos por debajo del resultado global de los colegios femeninos.

Un hallazgo importante para nosotros es que, en nuestro estudio, un estudiante de un colegio oficial, o público como se conoce, tiene un resultado de un punto por encima de los colegios no oficiales o privados. Sin embargo, se debe hacer la salvedad que como se muestra en los datos, la participación de los estudiantes de colegios oficiales es notoriamente superior.

Con respecto al carácter del colegio, ser académico es la característica que brinda una probabilidad de obtener mejores resultados. Sin embargo, los técnicos o técnicos académicos están por debajo en 3 y 2 puntos respectivamente, mientras que los colegios a los cuales no aplica tienden a alejarse en 18 puntos por debajo de los colegios académicos. Esto nos deja como guía importante para buscar los colegios con al menos una de estas características. Además, encontramos que los estudiantes que asisten a la sede principal de su colegio obtienen 7 puntos

por encima en sus resultados. Como también encontramos que los colegios que están en áreas urbanas obtienen 7 puntos por encima de los colegios de áreas rurales.

Continuando con el análisis, encontramos que los estudiantes de colegios de jornada completa son los que mejores resultados tienen. Tomando esta característica como base, los que estadísticamente tienen resultados más inferiores son los sabatinos, seguido de los de jornada nocturna que tienen un resultado 40 y 37 puntos por debajo. Los colegios que mejores resultados obtienen seguidos de la jornada completa son los de jornada única, sin embargo, tienen 13 puntos por debajo.

Por último, teniendo como base los estudiantes del departamento del Cauca, los estudiantes del Chocó tienen 13 puntos por debajo y los de Nariño 14 puntos por encima de los primeros. El departamento del Valle tiende a tener 4 puntos por encima. Lo anterior, implicaría que, en promedio, a los estudiantes de Nariño les irá mejor en las pruebas de Estado, seguido de los del Valle del Cauca y Cauca y, por último, los que menos resultados tienen son los estudiantes ubicados en el departamento del Chocó.

6.4. Problemas econométricos

La multicolinealidad es uno de los principales y más graves problemas en la econometría. Debido a que, si el modelo llegase a presentarla en grado perfecto, no se podrían calcular los estimadores de los parámetros. Además, se conoce como algún grado de relación lineal entre las variables. Por lo anterior, se comprobó que no existiera este problema en grado perfecto. Una vez resuelto esto, los investigadores podemos vivir con multicolinealidad no perfecta, dado que, aunque los estimadores están relacionados, no hay otro que sea MELI, es decir con una varianza menor a los que hallamos.

Por otra parte, la heterocedasticidad es otro de los problemas más comunes de la econometría y se ve generalmente explicado por los datos, en particular, por los de corte transversal. Por lo anterior, y el hecho de que estamos trabajando con datos de corte trasversal, debemos suponer que hay riesgo de sufrir este problema, por lo cual, deberá realizarse las pruebas formales para descartarlo y poder continuar con el análisis. Una vez se realizó la prueba, se halló que nuestro modelo inicial presentaba problemas de heterocedasticidad.

Por lo anterior, se corrió el modelo de forma robusta para solucionar este problema y poder contar con unos estimadores MELI que nos permitiera obtener las mejores interpretaciones a cada una de las categorías analizadas en el presente trabajo. De este modo, se corroboró los problemas econométricos y en el caso de heterocedasticidad se corrigió para pasar a darle interpretación a cada uno de los resultados.

7. CONCLUSIONES Y RECOMENDACIONES

Dado los resultados obtenidos en el capítulo anterior, se puede decir que la región necesita que se sigan haciendo estudios en torno a los factores que determinan los resultados en las pruebas de estados para buscar, por medio de políticas públicas, disminuir la heterogeneidad en los resultados.

Por esto, hacemos un llamado a que cada profesional desde su respectiva área de conocimiento enfoque sus esfuerzos en la región y contribuyamos por medio de material académico unas bases sólidas para recomendaciones a los siguientes gobernantes de la región en aras a destacar a nivel nacional.

8. ANEXOS

Anexo 1. Tabla de regresión

Coefficients:	Estimate	
(Intercept)	8383.14661	***
GeneroM	11.19545	***
Periodo	-0.40115	***
Tiene_EtniaSi	-23.85907	***
Personas_Hogar3 a 4	4.45651	***
Personas_Hogar5 a 6	2.87030	***
Personas_Hogar7 a 8	0.43569	
Personas_Hogar9 o más	-4.01870	***
Cuartos_HogarCuatro	3.99167	***
Cuartos_HogarDos	10.87752	***
Cuartos_HogarSeis o mas	-1.69601	.
Cuartos_HogarTres	8.18076	***
Cuartos_HogarUno	7.50393	***
Edu_PadreEducación profesional incompleta	3.93867	***
Edu_PadreNinguno	-18.00369	***
Edu_PadreNo Aplica	-2.22462	.
Edu_PadreNo sabe	-3.50271	***
Edu_PadrePostgrado	20.50849	***
Edu_PadrePrimaria completa	-10.90220	***
Edu_PadrePrimaria incompleta	-10.55666	***
Edu_PadreSecundaria (Bachillerato) completa	-7.31821	***
Edu_PadreSecundaria (Bachillerato) incompleta	-10.08059	***
Edu_PadreTécnica o tecnológica completa	1.01901	
Edu_PadreTécnica o tecnológica incompleta	-3.67408	***
Edu_MadreEducación profesional incompleta	1.07224	
Edu_MadreNinguno	-17.82564	***
Edu_MadreNo Aplica	-20.79691	***
Edu_MadreNo sabe	-13.05274	***
Edu_MadrePostgrado	14.83282	***
Edu_MadrePrimaria completa	-14.65619	***
Edu_MadrePrimaria incompleta	-14.55208	***
Edu_MadreSecundaria (Bachillerato) completa	-10.17353	***
Edu_MadreSecundaria (Bachillerato) incompleta	-12.34802	***
Edu_MadreTécnica o tecnológica completa	-0.42258	
Edu_MadreTécnica o tecnológica incompleta	-3.76876	***
Trabajo_PadreEs dueño de un negocio grande, tiene un cargo de nivel directivo o gerencial	-2.64159	**
Trabajo_PadreEs dueño de un negocio pequeño (tiene pocos empleados o no tiene, por ejemplo tienda, papelería, etc	1.15178	*
Trabajo_PadreEs operario de máquinas o conduce vehículos (taxita, chofer)	-0.29702	
Trabajo_PadreEs vendedor o trabaja en atención al público	-2.76690	***
Trabajo_PadreNo aplica	0.90998	.
Trabajo_PadreNo sabe	-1.71305	**

Trabajo_PadrePensionado	0.13397	
Trabajo_PadreTiene un trabajo de tipo auxiliar administrativo (por ejemplo, secretario o asistente)	-6.24727	***
Trabajo_PadreTrabaja como personal de limpieza, mantenimiento, seguridad o construcción	-3.58910	***
Trabajo_PadreTrabaja como profesional (por ejemplo médico, abogado, ingeniero)	1.68613	*
Trabajo_PadreTrabaja en el hogar, no trabaja o estudia	-9.68428	***
Trabajo_PadreTrabaja por cuenta propia (por ejemplo plomero, electricista)	1.19387	*
Trabajo_MadreEs dueño de un negocio grande, tiene un cargo de nivel directivo o gerencial	-4.10972	***
Trabajo_MadreEs dueño de un negocio pequeño (tiene pocos empleados o no tiene, por ejemplo tienda, papelería, etc)	4.88202	***
Trabajo_MadreEs operario de máquinas o conduce vehículos (taxita, chofer)	2.78792	*
Trabajo_MadreEs vendedor o trabaja en atención al público	2.57077	***
Trabajo_MadreNo aplica	-3.57013	***
Trabajo_MadreNo sabe	-6.62055	***
Trabajo_MadrePensionado	-0.92676	
Trabajo_MadreTiene un trabajo de tipo auxiliar administrativo (por ejemplo, secretario o asistente)	3.37652	***
Trabajo_MadreTrabaja como personal de limpieza, mantenimiento, seguridad o construcción	3.84698	***
Trabajo_MadreTrabaja como profesional (por ejemplo médico, abogado, ingeniero)	7.62908	***
Trabajo_MadreTrabaja en el hogar, no trabaja o estudia	4.54978	***
Trabajo_MadreTrabaja por cuenta propia (por ejemplo plomero, electricista)	4.07256	***
ComputadorSi	6.96164	***
HornoSi	-2.63040	***
Servicio_tvSi	-1.96409	***
AutomovilsSi	2.15168	***
MotocicletaSi	-3.61602	***
Lectura_DiariaEntre 1 y 2 horas	11.19463	***
Lectura_DiariaEntre 30 y 60 minutos	9.06743	***
Lectura_DiariaMás de 2 horas	18.78184	***
Lectura_DiariaNo leo por entretenimiento	-4.62183	***
Ded_InternetEntre 1 y 3 horas	4.69127	***
Ded_InternetEntre 30 y 60 minutos	1.69973	***
Ded_InternetMás de 3 horas	3.60282	***
Ded_InternetNo Navega Internet	-0.25518	
Genero_EstablMASCULINO	10.12457	***
Genero_EstablMIXTO	-11.20446	***
NaturalezaOFICIAL	0.77686	*
Bilingues	1.26354	.
CaracterNO APLICA	-18.24716	***
CaracterTÉCNICO	-3.09817	***
CaracterTÉCNICO/ACADÉMICO	-2.83076	***
Sede_Principals	6.90913	***
Area_UbicacionURBANO	7.05734	***
JornadaMAÑANA	-14.30117	***
JornadaNOCHE	-37.46159	***
JornadaSABATINA	-40.32525	***
JornadaTARDE	-16.39150	***
JornadaUNICA	-13.02087	***
Dpto_UbicacionCHOCO	-13.41310	***
Dpto_UbicacionNARIÑO	14.00710	***
Dpto_UbicacionVALLE	3.76904	***
Edad	-1.36550	***
TrabajaSi	-9.08320	***
Coefficients:		Estimate
(Intercept)	8383.14662	***
GeneroM	11.19546	***
Periodo	-0.40116	***
Tiene_EtniaSi	-23.85908	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	
Residual standard error: 40.03 on 117527 degrees of freedom	
Multiple R-squared: 0.3284, Adjusted R-squared: 0.3279	
F-statistic: 638.6 on 90 and 117527 DF, p-value: < 2.2e-16	

9. BIBLIOGRAFÍA

- Banco Mundial. (19 de Septiembre de 2017). *Banco Mundial*. Obtenido de <https://www.bancomundial.org/es/topic/education/overview>
- Barrientos, J. (2008). Calidad de la educación pública y logro académico en Medellín 2004-2006. Una aproximación por regresión intercuartil. *Lecturas de Economía*, 121-144.
- Bermudez, E. A. (27 de Junio de 2014). *EL ESPECTADOR*. Obtenido de En qué falla la educación del Pacífico Colombiano: <https://www.elespectador.com/noticias/educacion/falla-educacion-del-pacifico-colombiano-articulo-501105>
- Colprensa. (19 de Agosto de 2011). *El Colombiano*. Obtenido de Educación Pública básica y media en el país será gratuita a partir del 2012: http://www.elcolombiano.com/historico/educacion_publica_basica_y_media_en_colombia_a_sera_gratuita_a_partir_de_2012-BAEC_146304
- DANE. (2018). *Boletín Técnico. Pobreza Monetaria y Multidimensional en Colombia 2017*. Bogotá.
- DANE. (2018). *Gran Encuesta Integrada de Hogares (GEIH) - Fuerza Laboral y Educación 2017*. Bogotá.
- Delprato, M. (1999). *Determinantes del rendimiento educativo del nivel primario aplicando la Técnica de Análisis Multinivel*. Córdoba.

El Tiempo. (07 de Agosto de 2016). *EL TIEMPO*. Obtenido de A disminuir la brecha educativa en el Pacífico: <https://www.eltiempo.com/colombia/cali/disminuir-la-brecha-educativa-en-el-pacifico-37851>

Escardibul, J. O. (2008). *LOS DETERMINANTES DEL RENDIMIENTO EDUCATIVO EN ESPAÑA. UN ANÁLISIS A PARTIR DE LA EVALUACIÓN DE PISA-2006*. Barcelona: Investigaciones de Economía de la Educación.

Escobar, J. D. (2014). *Análisis de las competencias en matemáticas y lenguaje de los bachilleres*. Cali.

Espinoza, O. (2009). Reflexiones sobre los conceptos de "política", políticas públicas y política educacional. *Education Policy Analysis* .

Estrada, D. (2017). *El efecto del entorno de desarrollo municipal en la distribución de los resultados en matemáticas - Saber 11: una aproximación utilizo modelos jerárquicos lineales por cuartiles*. Cali.

Gaviria, A., & Barrientos, J. (2001). *Determinantes de la calidad de la educación en Colombia*. Bogotá: FEDESARROLLO.

GOMEZ, R. R. (2002). Continuidad y cambio de las políticas de educación superior. *Revista Mexicana de Investigación Educativa*, 133-154.

Gómez, S. M., Gutierrez, D. M., & Hasan, A. R. (2011). *Determinantes del rendimiento académico en Colombia: pruebas ICFES Saber 11°, 2009*. Medellín: Centro de Investigaciones Económicas y Financieras.

- J. Graham, B. M.-M. (2005). Stepwise detection of recombination breakpoints in sequence alignments. *Bioinformatics*, págs. 589 - 595.
- López, M., Posada, H. M., & Henao, E. A. (2014). *ENTORNO INSTITUCIONAL Y CALIDAD DE LA EDUCACIÓN: UN ANÁLISIS EXPLORATORIO DE LOS RESULTADOS DE LAS PRUEBAS ESTATALES PARA LOS COLEGIOS PÚBLICOS DE LA CIUDAD DE MEDELLÍN*. Medellín: Semestre Económico.
- Ministerio de Educación Nacional. (2018). *Lineamientos para la implementación de la Jornada Única en Colombia durante 2018*. Bogotá.
- Obregón, A. M. (2010 de Diciembre de 2010). *¿Cómo construir desarrollo en el Pacífico colombiano sin calidad educativa?* Obtenido de EL TIEMPO:
<http://blogs.eltiempo.com/afrocolombianidad/2010/12/08/como-construir-desarrollo-en-el-pacifico-colombiano-sin-calidad-educativa/>
- Orejuela, J. (2013). *Análisis del desempeño estudiantil en las pruebas de Estado para educación media en Colombia mediante modelos jerárquicos lineales*. Bogotá: Ingeniería, Vol 18, No 2.
- Piñeros Jiménez, L. J., & Rodríguez Pinzón, A. (1998). *Los Insumos Escolares en la Educación Secundaria y su Efecto Sobre el Rendimiento Académico de los Estudiantes: Un estudio en Colombia*. The World Bank.
- Schwartz, P. (06 de Mayo de 2014). Gary Becker o el enfoque económico. *Expansión*, pág. 54.

Vincezi, A. D., & Tudesco, F. (2009). La educación como proceso de mejoramiento de la calidad de vida de los individuos y de la comunidad. *Revista Iberoamericana de Educación*, 1 - 12.

Viveros, V. H. (26 de Mayo de 2017). *La Silla Vacía*. Obtenido de La Educación en el Pacífico es el camino: <https://lasillavacia.com/silla-llena/red-pacifico/historia/la-educacion-en-el-pacifico-es-el-camino-61099>