

**DESIGUALDAD EN EL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES  
EN COLOMBIA**

**AUTOR:**

**ELISA MEJÍA VÉLEZ**

**MANUELA GIRALDO SIERRA**

**ASESOR:**

**MÓNICA OSPINA**

**EAFIT**

**ECONOMÍA**

**MEDELLÍN**

**4 DE NOVIEMBRE DE 2014**

1	RESUMEN .....	5
2	INTRODUCCIÓN .....	5
3	PLANTEAMIENTO DEL PROBLEMA .....	7
4	OBJETIVO PRINCIPAL.....	7
4.1	OBJETIVOS ESPECÍFICOS .....	7
5	ESTADO DEL ARTE .....	8
5.1	Calidad educativa vs Equidad educativa.....	8
5.2	Estudios que utilizan los resultados de los exámenes estatales para medir la desigualdad en el rendimiento académico .....	11
5.3	SABER 11 vs SIMCE .....	13
5.4	Coeficiente de Gini .....	14
6	METODOLOGÍA.....	15
7	DATOS.....	24
8	ANÁLISIS DE RESULTADOS .....	26
8.1	Desigualdad en los puntajes de las pruebas Saber 11 .....	28
8.2	Descomposición del Índice Gini en subpoblaciones.....	30
8.2.1	Segmentación por género.....	30
8.2.2	Segmentación por condición de ruralidad.....	31
8.2.3	Segmentación por estrato socioeconómico .....	32
8.2.4	Segmentación por tipo de establecimiento (público-privado).....	34
9	CONCLUSIONES.....	35
10	BIBLIOGRAFÍA.....	37

11	ANEXOS.....	39
	Anexo 1 Variables Utilizadas para la Construcción del INSE .....	39
	Anexo 2 Resultado de la Cuantificación de las Variables Cualitativas .....	40
	Anexo 3 Distribución del Índice de Nivel Socioeconómico .....	42
	Ilustración 1 Coeficiente Gini en Colombia para el período objeto de estudio.....	14
	Ilustración 2 Scree Test.....	21
	Ilustración 3 Gini nivel socioeconómico.....	27
	Ilustración 4 Gini puntaje (Lenguaje) .....	28
	Ilustración 5 Gini Puntaje (Matemáticas) .....	29
	Ilustración 6 Segmentación por género .....	31
	Ilustración 7 Segmentación por condición de ruralidad .....	32
	Ilustración 8 Segmentación por estrato socioeconómico.....	33
	Ilustración 9 Segmentación por tipo de establecimiento.....	34
	Tabla 1 Coeficiente Gini para el año 2012 en Latinoamérica .....	14
	Tabla 2 Estadística INSE 2012.....	17
	Tabla 3 Contribución de cada variable al INSE .....	18
	Tabla 4 Matriz de Correlación de las Variables .....	20
	Tabla 5 Test de KMO .....	20
	Tabla 6 Género .....	24

Tabla 7 Área en la que vive.....	25
Tabla 8 Estrato .....	25
Tabla 9 Tipo de Establecimiento .....	25
Tabla 10 Nivel de estudios del padre y de la madre.....	25
Tabla 11 Ingreso Familiar.....	26
Tabla 12 Matriz de Resultados Promedio vs Gini.....	26
Tabla 13 Índice Gini de Nivel socioeconómico y media del INSE .....	27
Tabla 14 Gini resultados puntaje Saber 11 Lenguaje.....	28
Tabla 15 Gini resultados puntaje Saber 11 Matemáticas .....	29

## **1 RESUMEN**

En la presente investigación se busca evaluar que tan desiguales son los resultados académicos de los estudiantes del grado 11 en Colombia para el período 2008-2012. Para esto se construyó un Gini de puntajes de las pruebas Saber 11 para cada año del período evaluado y se compara con el Gini de nivel socioeconómico estimado a partir de información de los estudiantes sobre sus condiciones socioeconómicas tomada del formulario de inscripción al examen. Los resultados obtenidos muestran que la desigualdad en términos de puntaje es menor que la desigualdad en términos socioeconómicos, y aunque es relativamente baja se encuentra acompañada de puntajes muy bajos en las poblaciones más vulnerables. Los resultados más desiguales y con menor puntaje promedio se dan en el área de matemáticas en los estratos socioeconómicos 1 y 2. Según estos resultados, la política educativa debería estar dirigida a reducir la desigualdad en el rendimiento de los alumnos aumentando el rendimiento promedio de estos y no a costa de puntajes más bajos, ya que un sistema educativo que logre menor desigualdad implica una sociedad más homogénea, donde todas las personas puedan tener las mismas oportunidades y se cierren brechas de oportunidades e ingresos.

## **2 INTRODUCCIÓN**

El presente trabajo pretende medir la distribución de los resultados en las pruebas SABER 11 en Colombia durante el período 2008-2012, en general para todos los estudiantes de colegios públicos y privados, y para diferentes subpoblaciones según las características de los estudiantes. En especial, se hace énfasis en la relación existente entre el coeficiente Gini de puntajes y el coeficiente Gini de ingresos y se descompone el coeficiente Gini de puntajes para diferentes subpoblaciones con el fin de analizar en cuáles grupos poblacionales se evidencia mayor desigualdad en los puntajes.

La literatura se ha enfocado en analizar los resultados de pruebas estandarizadas en términos absolutos, sin embargo medidas relativas, como la desigualdad de los puntajes, son igualmente importantes a la hora de hacer recomendaciones de

políticas en términos de la calidad de la educación. Por ejemplo, aun cuando dos países o dos colegios tengan promedios iguales en términos de puntajes, es posible que la dispersión de sus resultados sea muy diferente, implicando intervenciones diferentes a la hora de mejorar la calidad de educación. En general, se busca no solo que los estudiantes tengan altos puntajes en promedio, si no que la dispersión de los resultados entre estudiantes con diferentes condiciones socioeconómicas disminuya durante el ciclo escolar, mostrando que el sistema educativo es capaz de cerrar las brechas entre estudiantes, generando así igualdad de oportunidades para la población en términos de capital humano.

Por otro lado, la literatura en economía de la educación ha mostrado consistente evidencia empírica de la relación positiva entre el nivel socioeconómico de los estudiantes y su efecto en los resultados en pruebas estandarizadas de conocimiento. Sin embargo, no existe mucha investigación que evalúe la distribución del desempeño académico al interior de diferentes grupos de población y su relación con la distribución de los ingresos. Se espera que sistemas educativos más eficientes logren disminuir la dispersión en términos de puntajes en comparación con la dispersión del nivel socioeconómico de los estudiantes, es decir, un buen sistema educativo debería cerrar brechas en el rendimiento académico de sus estudiantes, lo cual se puede ver reflejado en oportunidades homogéneas al graduarse, para el acceso a educación superior y al mercado laboral, aún si las condiciones iniciales de los estudiantes son desiguales.

Este trabajo mide y analiza la desigualdad en los puntajes y su relación con la desigualdad en el nivel socioeconómico, pero no identifica efectos de causalidad, y se limita a dar evidencia sobre la primera hipótesis, donde se relacionan los coeficientes de desigualdad de puntajes y de nivel socioeconómico de los estudiantes. Los resultados son usados para dar recomendaciones de focalización de política pública en términos de cuáles son los grupos de población con mayores niveles de desigualdad en términos de puntajes.

### **3 PLANTEAMIENTO DEL PROBLEMA**

En la literatura de economía de la educación hay suficiente evidencia de cuales características de los estudiantes o del colegio afectan positivamente los resultados en pruebas estandarizadas, sin embargo, no conocemos que tan desigual es la población en términos de los puntajes, y para qué grupos de población se presentan mayores niveles de desigualdad. Se espera que mejores sistemas educativos tengan no solo altos promedios en los resultados, sino resultados menos dispersos, garantizando así que la educación sea una herramienta que permita cerrar la brecha entre los estudiantes.

Dada la importancia de medir la dispersión de los resultados académicos dentro de la población estudiantil, se propone responder a la pregunta de qué tan desigual es el rendimiento de los estudiantes en comparación con la desigualdad en su nivel socioeconómico. Esto es importante en términos de política pública, ya que menor desigualdad de resultados en puntajes puede implicar poblaciones con capital humano más homogéneo, y por tanto, menor desigualdad de ingresos en el futuro.

### **4 OBJETIVO PRINCIPAL**

Analizar la desigualdad en los puntajes del examen SABER 11 para diferentes subgrupos de población y analizar su relación con la desigualdad en ingresos de los padres de los estudiantes.

#### **4.1 OBJETIVOS ESPECÍFICOS**

- Hacer una revisión de literatura sobre estudios que hayan abordado el tema de la desigualdad en el rendimiento académico de los estudiantes a nivel nacional e internacional.
- Estimar el indicador Gini para puntajes de los estudiantes en la Prueba SABER 11 y estimar un indicador multivariado de las características socioeconómicas de los estudiantes para estimar el indicador Gini por nivel socioeconómico de los estudiantes.

- Hacer descomposiciones del Gini de puntajes por características de la población que permitan capturar los efectos en los puntajes de las pruebas.
- Comparar la evolución del GINI de los resultados de las pruebas SABER 11 y el GINI de ingresos de los padres en el período 2008-2012.

## **5 ESTADO DEL ARTE**

### **5.1 Calidad educativa vs Equidad educativa**

Dada la importancia de que exista un buen sistema educativo en una sociedad, que garantice el aprendizaje y la formación de capital humano con acceso a oportunidades en el ámbito educativo y laboral, la labor investigativa tiene un papel importante en determinar no solo si las instituciones están realizando una labor eficiente y si está mejorando la calidad y cobertura educativa, sino también si la forma en que se evalúa el desempeño de estas instituciones es el correcto y por ende si las políticas educativas están bien dirigidas o no. Al respecto se analizan temas no sólo de eficiencia y calidad sino también de equidad y justicia social.

Gamboa, Casas & Piñeros (2003) abordan el problema sobre la carencia que existe en los procedimientos para clasificar y evaluar los establecimientos educativos. Debido a que el nivel socioeconómico es diferente para cada niño que entra a estudiar a una institución educativa, el desempeño de estos establecimientos no se puede medir a partir del resultado promedio en las pruebas estatales sino que se debería medir teniendo en cuenta las brechas sociales que existen entre los estudiantes que asisten a los diferentes establecimientos, para evaluar si teniendo en cuenta los niveles socioeconómicos de estos, las instituciones están logrando formar capital humano con menos desigualdad y con igualdad de oportunidades en el mercado laboral. La forma como se evalúan las instituciones educativas es finalmente la materia prima para medir si la política pública está bien dirigida o no. Este estudio complementa el propósito de esta tesis, pues medir la desigualdad en los puntajes de manera complementaria de los resultados promedio busca

principalmente una mejor medición del rendimiento de los estudiantes y por ende de la labor del sistema educativo.

A pesar del reconocimiento general de que la mayor equidad no sólo es deseable en sí misma sino también por ser instrumental para alcanzar una sociedad como un todo más próspera en el largo plazo, las sociedades difícilmente toman decisiones socialmente óptimas que reflejen una mayor equidad. Esto se da por intereses políticos particulares o por falta de información en términos de desigualdad, que no refleja la verdadera problemática de situaciones específicas, como la educación en este caso. (Mélendez, 2012)

En el momento en que un docente se debe enfrentar a un grupo heterogéneo de estudiantes, debería adecuar la manera de enseñar a las estructuras cognitivas de cada niño de manera que alumnos con características muy distintas logren resultados académicos similares. Esto se puede lograr por ejemplo incentivando a los mejores egresados de pedagogía para que trabajen en escuelas de bajo rendimiento promedio.

Hay evidencia de que en los países con mejores estándares de calidad educativa, como Finlandia, Corea del Sur y Japón, los estudiantes con peor rendimiento académico alcanzan niveles suficientes para seguir avanzando en su estudio, lo que indica que la educación está logrando generar oportunidades homogéneas para estudiantes con condiciones iniciales heterogéneas. Es así como se debería medir la eficiencia de las instituciones educativas, más allá de los resultados promedios. (Ramírez, 2007)

Los resultados educativos en Colombia no sólo son deficientes en relación a las pruebas internacionales PISA, donde está ubicado en los rangos inferiores (por debajo de países Latinoamericanos como Chile, México y Argentina) (OECD, 2012), sino que están distribuidos desigualmente entre la población. Los estudiantes de los colegios públicos, de los niveles socioeconómicos bajos y los de las zonas rurales tienen peor desempeño con respecto a los colegios privados, los de niveles de

ingreso altos y los que habitan en zonas urbanas. Asimismo, existen diferencias en calidad entre regiones. La presencia de diferencias en la calidad de la educación que reciben los individuos, genera brechas en calidad de vida y brechas en ingresos que profundizan y reproducen las desigualdades sociales (Barrera M. D., 2014). Es importante entonces determinar los factores que inciden en que estas brechas continúen aumentando. Algunos estudios realizados sobre el tema en Colombia, destacan factores como la calidad de los docentes, la capacidad institucional, la jornada escolar y sobretodo el nivel socioeconómico: el estatus socioeconómico de los estudiantes ha sido conocido a través del tiempo como un factor relevante en el desempeño académico de los mismos. Se ha demostrado que existe una relación directa entre los resultados académicos y su nivel socioeconómico. (Coleman, 1996) Colombia al ser un país tan desigual en términos de ingreso, hace que este factor sea relevante a la hora de hacer estudios de resultados académicos, y se espera que la desigualdad en educación sea mayor respecto a países que son más equitativos en su distribución de ingresos.

Un ejemplo de la desigualdad reflejada en este tipo de pruebas es el estudio realizado por Gamboa & Waltenberg (2011), el cual tiene como objetivo evaluar la inequidad en las oportunidades de las personas en 6 países de Latinoamérica para los resultados académicos, teniendo como base las pruebas PISA (Program for International Student Assessment). Según los resultados encontrados, los hacedores de políticas públicas en estos 6 países: Argentina, México, Colombia, Brasil, Chile y Uruguay, deberían ser más creativos y buscar nuevas políticas orientadas a disminuir las desigualdades tanto en acceso como en calidad de la educación. Es relevante además incluir planes para mejorar las condiciones de aprendizaje de los menos favorecidos y que puedan en un futuro tener las mismas oportunidades. Este estudio es relevante pues la educación tiene un valor intrínseco en el futuro de la sociedad; viéndolo desde el punto de vista macroeconómico, la calidad de la educación, medida en puntajes, parece tener una relación positiva y dominante con el crecimiento económico de una región o país. Los autores encuentran que en Latinoamérica existe una desigualdad alta y preocupante en los

resultados de las pruebas PISA, y esta desigualdad depende del año, país, y materia que se esté evaluando, variando del 1% al 25%. Para el caso de Colombia, se encuentran puntajes promedio de más de 100 puntos por debajo del nivel promedio tomado por la OECD, lo cual resulta preocupante, a pesar de tener la menor desigualdad entre los 6 países, para matemáticas y lectura.

Otro ejemplo es Santos (2007), quien estudia los determinantes y la distribución de los resultados académicos, medidos con el puntaje obtenido en las pruebas PISA 2000 en lectura y matemática. Se concluye que existe una brecha de género a favor de los hombres en matemática y que características como el tamaño de las aulas, si la escuela es privada o pública y las características familiares de los estudiantes influyen en la distribución de los resultados de las pruebas y deberían ser tomados en cuenta para el diseño de política pública. En el presente estudio se pretende también analizar los resultados para subpoblaciones según género, tipo de establecimiento (público o privado) y características socioeconómicas y ver cómo es su distribución.

Para el propósito de esta investigación se menciona a continuación algunos estudios realizados en Chile y en Colombia, en los cuales se analiza la distribución de los puntajes académicos obtenidos en las pruebas estatales. Luego se procede a hacer una breve comparación entre las pruebas SABER 11 aplicadas en Colombia y las pruebas SIMCE aplicadas en Chile y por último se hace una breve descripción del indicador Gini utilizado en el presente estudio para medir la desigualdad en el rendimiento académico.

## **5.2 Estudios que utilizan los resultados de los exámenes estatales para medir la desigualdad en el rendimiento académico**

El estudio en el cual este trabajo se apoya principalmente, realizado por la Agencia de Calidad de la Educación del Gobierno de Chile, mide la distribución de los

resultados SIMCE<sup>1</sup> y analiza los niveles de desigualdad en el rendimiento de los estudiantes en comparación con la desigualdad en ingreso de sus padres. En este estudio se calcula el índice Gini a partir de las declaraciones de ingreso del cuestionario de padres y apoderados en los niveles de 4° Básico y II Medio<sup>2</sup> (Educación, 2013). Los resultados de este estudio muestran una clara relación lineal negativa entre el rendimiento de los alumnos (medido como el promedio de los resultados de las pruebas Simce en Lenguaje, Matemáticas y Ciencias Sociales) y la desigualdad de los puntajes en ambos niveles, que se acentúa para los grupos socioeconómicos más desventajados, es decir, *“a mayor nivel del grupo socioeconómico, mayor el puntaje Simce y menor la desigualdad de esos puntajes”*. (Agencia de Calidad de la Educación, 2013)

En otro estudio para Chile, Ramírez (2007) analiza la distribución del rendimiento en matemáticas en alumnos de 4° básico a nivel de regiones, provincias, comunas, escuelas, salones y alumnos dentro de los salones. Se encuentra que la mayor dispersión de resultados en el SIMCE se encuentra dentro de las aulas de clase, lo cual evidencia un método de enseñanza que no tiene en cuenta los distintos niveles de rendimiento de los alumnos. Aunque no tiene mucho sentido buscar que no haya diferencias entre el desempeño de alumnos de un mismo colegio, si se debe buscar reducir la desigualdad entre escuelas; cuando esta es mínima se puede afirmar que los estudiantes tienen similares oportunidades de aprendizaje independiente de la escuela a la que asistan y las diferencias de rendimiento dentro de las escuelas pueden deberse a características específicas de cada alumno, como su motivación y esfuerzo por aprender. Sus resultados permiten entender como la política pública debe estar direccionada, dependiendo de dónde específicamente se encuentre la

---

<sup>1</sup> Prueba nacional en Chile que se utiliza para evaluar los resultados de aprendizaje de los estudiantes del país que cursan los niveles evaluados (4° básico y II medio.)

<sup>2</sup> Los niveles 4° Básico y II Medio en Chile son homólogos en Colombia a los niveles 4° de básica primaria y 10° de educación media respectivamente.

desigualdad, para poder implementar métodos adecuados ya sea relacionados con los docentes y sus estrategias o con las condiciones personales de cada estudiante.

En el caso específico de Colombia, Gamboa (2014) analiza cómo el haber asistido o no a preescolar, tiene una influencia en las brechas de aprendizaje que se presentan en la educación básica primaria y secundaria. Se encuentra que la diferencia promedio en el resultado de la prueba SABER 11 entre los estudiantes que asistieron a preescolar y los que no es del 5% y aumenta a 17% en matemáticas si se mide teniendo en cuenta el tiempo de permanencia en preescolar. Estos resultados evidencian la necesidad de enfocar la política pública a una mejora en la eficiencia y equidad de la educación, más allá de buscar mejorar los resultados promedio, se necesita lograr que los estudiantes cuenten con las mismas oportunidades.

### **5.3 SABER 11 vs SIMCE**

La Ley 1324 de 2009, convierte al ICFES en el Instituto Colombiano para la Evaluación de la Calidad de la Educación y le asigna la función técnica de desarrollo e implementación de los exámenes de Estado (SABER 11 y PRO) y lo habilita para desarrollar y ofrecer otras evaluaciones. Las pruebas Saber 11 tiene, entre otros, el objetivo de ofrecer información que sirva como referente estratégico para el establecimiento de políticas educativas nacionales, territoriales e institucionales. (Nacional, 2014)

SIMCE es la prueba nacional de Chile que se utiliza para evaluar los resultados de aprendizaje de los estudiantes del país. Se realiza una vez al año a los estudiantes que cursan cuarto básico, y año por medio a aquellos que cursan octavo básico y segundo medio. Los nivel 4° Básico y II Medio en Chile son homólogos en Colombia a los niveles 4° de básica primaria y 10° de educación media respectivamente. La prueba evalúa el desempeño en Lenguaje, Matemáticas, Ciencias (Ciencias Naturales y Ciencias Sociales) y tiene como propósito obtener y proveer información sobre logros de aprendizaje que sirva a las instituciones y al gobierno para tomar

decisiones para el continuo mejoramiento de la calidad de la educación. (Meckes, 2014)

#### 5.4 Coeficiente de Gini

El coeficiente de Gini es un indicador sintético con el cual se busca dar un valor absoluto a la desigualdad en la variable que se esté evaluando (en este caso puntajes en el examen SABER 11), lo cual permite hacer comparaciones en el tiempo y entre individuos y poblaciones.

Un índice de desigualdad mide como se distribuye una variable entre varios individuos, de esta forma se puede usar para medir la distribución no solo del ingreso sino de otras variables como resultados académicos que puede ser útil para analizar la raíz de muchos problemas sociales.

A continuación se observa la medida de desigualdad en ingresos para Colombia y Latinoamérica.

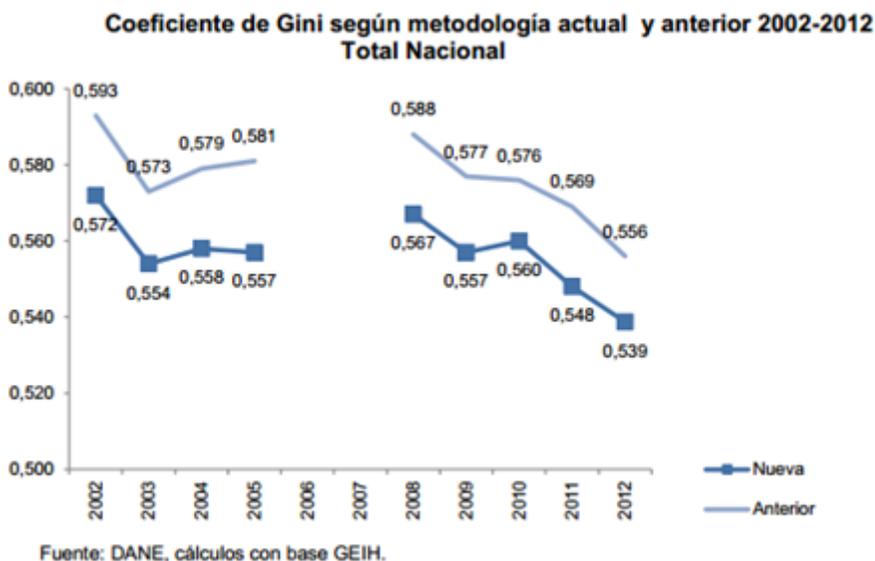
Tabla 1 Coeficiente Gini para el año 2012 en Latinoamérica

	2010-2012
ARGENTINA*	0,475
BOLIVIA	0,472
BRASIL	0,567
CHILE	0,516
COLOMBIA	0,536
COSTA RICA	0,504
MÉXICO	0,492
PERÚ	0,449
VENEZUELA	0,405

Fuente: Estadísticas de la CEPAL-Naciones Unidas

\*Los datos para Argentina corresponden a la región urbana al ser los únicos disponibles.

Ilustración 1 Coeficiente Gini en Colombia para el período objeto de estudio



## 6 METODOLOGÍA

La metodología de este estudio se basa en la construcción de dos índices Gini para cada uno de los años del período objeto de estudio, uno para medir la desigualdad en puntajes de la Prueba SABER 11 y otro para medir la desigualdad en el nivel socioeconómico. Se utilizan los dos Gini debido a la relación que existe entre el nivel socioeconómico de un estudiante y su rendimiento escolar, este factor incide directamente en las condiciones con las que cuenta un estudiante para estudiar, por ejemplo el lugar de estudio y el tiempo que le puede dedicar; por este motivo este estudio busca hacer una comparación entre los dos niveles de desigualdad para evaluar si la educación logra cerrar brechas de rendimiento académico aun cuando las condiciones socioeconómicas iniciales de los estudiantes son desiguales.

El Coeficiente de Gini para los puntajes de la Prueba SABER 11 en el área de Lenguaje y Matemáticas se construye con la siguiente ecuación:

$$GINI(X) = -2Cov\left(\frac{X}{\mu(X)}, (1 - F(X))\right) \quad (1),$$

Donde:

X: puntaje en la Prueba SABER 11.

$\mu(X)$ : media de la variable puntaje.

F(X): función de distribución acumulada de la variable puntaje en la Prueba SABER 11.

- El Coeficiente de Gini para el nivel socioeconómico se construye con la misma ecuación (1), donde:

X: Índice de nivel socioeconómico.

$\mu(X)$ : media de la variable nivel socioeconómico.

F(X): función de distribución acumulada de la variable nivel socioeconómico.

Para la variable de nivel socioeconómico de los hogares se construirá un índice multivariado replicando la metodología utilizada por el ICFES (diseñada por el Programa Nacional de Desarrollo Humano del DNP) en la cual se utiliza información obtenida del formulario de inscripción al Examen para estimar el Índice de Nivel Socioeconómico (INSE). Las variables que se toman representan los siguientes cuatro componentes: vivienda, capital humano, servicios y hacinamiento. (VER ANEXO 1)

Para la construcción del INSE se deriva un índice ponderado de las variables a través del procedimiento de *Componentes Principales*. Como en su mayoría, las variables tomadas para la construcción del índice son cualitativas, estas se deben transformar en numéricas por medio de la técnica de *Cuantificación Óptima*<sup>3</sup> para

---

<sup>3</sup> Método estadístico utilizado para convertir variables cualitativas en numéricas por medio de la asignación de valores a cada una de las categorías de cada variable. Los resultados generados son óptimos debido a que los valores de cada categoría presentan entre sí máxima separación. Se asigna una puntuación  $\alpha_j$  a la categoría j de la variable X, definida por:

poder utilizar el método de Componentes Principales; para el propósito de este estudio se toman los coeficientes obtenidos por el ICFES debido a que se trabajará con la misma base de datos y por lo tanto es adecuado asumir los mismos coeficientes (VER ANEXO 2).

Después de realizar la transformación se obtiene el índice a partir de la primera componente principal de un sistema de variables numéricas y de variables cualitativas cuantificadas. Se construye de tal forma que tome valores mayores o iguales a cero y a medida que se aleja de cero representa mejores condiciones de vida del hogar. El componente principal de las variables se obtiene por medio del software STATA con el comando *pca* (Principal Components Analysis) y se construye el índice con el comando *predict*, el cual después se estandariza con la ecuación:

$$INSE \text{ Estandarizado} = \frac{(INSE - \text{Min } INSE)}{(\text{Max } INSE - \text{Min } INSE)} * 100 \quad (2)$$

La información estadística del INSE se observa en la tabla 2:

**Tabla 2 Estadística Descriptiva del INSE 2008-2012**

Variable	Obs	Media	Desv Estándar	Min	Max
<b>INSE 2008</b>	417.205	35,5271	18,3537	0,00	100
<b>INSE 2009</b>	466.552	35,3721	18,5555	0,00	100
<b>INSE 2010</b>	502.145	36,2177	19,0853	0,00	100
<b>INSE 2011</b>	507.983	37,7810	18,9306	0,00	100
<b>INSE 2012</b>	510.976	37,2800	18,7400	0,00	100

Fuente: Elaboración propia.

Al obtener las componentes principales se descartan las variables de nivel de SISBEN y alfabetismo del padre y de la madre debido a que no aportan variabilidad al índice o hay muchas observaciones sobre las que no hay información. Finalmente

---


$$\lambda \alpha_j = \sum_{k=1}^K \frac{f_{jk}}{f_j} \beta_k, \text{ siendo } \lambda \text{ un factor de escala. (Vásquez, Ramírez, \& Camardiel, 2001)}$$

las variables utilizadas para construir el INSE son: Nivel educativo del padre y de la madre, Ocupación del padre y de la madre, Zona de residencia (rural o urbana), Estrato Socioeconómico, Hacinamiento, Ingresos del Hogar, Material del piso de la vivienda, Número de personas que conforman el Hogar y Dotación de bienes de uso doméstico.

Al aplicar el método de Componentes Principales para cada año se obtienen los siguientes resultados:

Para el 2008, 2009, 2010, 2011 y 2012, la primera componente principal de las variables explica el 37,82%, 39,79%, 41,94%, 42,14% y 40,47% respectivamente de la varianza total de este (índice). La contribución de cada variable al índice se observa en la tabla 3:

**Tabla 3 Contribución de cada variable al INSE tras la aplicación de PCA**

<b>Variable</b>	<b>2008</b>	<b>2009</b>	<b>2010</b>	<b>2011</b>	<b>2012</b>
<b>Área en la que vive</b>	0,1504	0,1578	0,1561	0,1660	0,1807
<b>Educación Madre</b>	0,3756	0,3679	0,3435	0,3448	0,3550
<b>Educación Padre</b>	0,3771	0,3662	0,3604	0,3609	0,3683
<b>Estrato</b>	0,3769	0,3733	0,3569	0,3585	0,3643
<b>Hacinamiento</b>	0,2066	0,2077	0,1998	0,1970	0,2020
<b>Nivel de Ingresos</b>	0,3939	0,3852	0,3719	0,3719	0,3789
<b>Material de los Pisos</b>	-0,2563	-0,2519	-0,2347	-0,2355	-0,2408
<b>Ocupación Madre</b>	0,2885	0,2792	0,3339	0,3303	0,2714
<b>Ocupación Padre</b>	0,3100	0,2963	0,3394	0,3369	0,3337
<b>Número de Personas en el Hogar</b>	0,1035	0,1065	0,0982	0,0994	0,1037
<b>Dotación del Hogar</b>	0,3139	0,3664	0,3615	0,3601	0,3672

Fuente: Elaboración propia.

En el Anexo 3 se observa en cada histograma la distribución del INSE, para el año 2012 se observa un sesgo a la izquierda, lo cual es de esperarse pues la mayoría de personas en Colombia viven bajo condiciones socioeconómicas desfavorables.

Una vez se desagrega el INSE por subpoblaciones se observa que para estudiantes de colegios privados tiene una distribución más uniforme, mientras que el de estudiantes de colegios públicos está completamente sesgado hacia la izquierda, pues estos tienen peores condiciones socioeconómicas; tanto el área rural como urbana tienen un sesgo hacia la izquierda, siendo el área rural mucho más concentrada en los niveles más bajos; por último el estrato 1 muestra un sesgo hacia la izquierda y el estrato 6 un sesgo hacia la derecha, lo cuál sería de esperarse.

Para verificar que las variables seleccionadas son adecuadas para construir el índice y que el método de *PCA* es adecuado, se saca la matriz de correlación (Tabla 4) de las variables y se hace el test de Kaiser-Meyer-Olkin - *KMO*<sup>4</sup> (Tabla 5). Según Tabachnick & Fidell (2007), para que las variables sean adecuadas para aplicar el método, las correlaciones deben ser al menos de 0,3, en la matriz se observa que no existe multicolinealidad (no hay correlación por encima de 0,9) y existe suficiente correlación entre las variables, excepto para las variables hacinamiento, número de personas en el hogar y área en la que vive, sin embargo, al realizar el test de *KMO*, los valores indican que existe correlación suficiente y es correcto utilizar estas variables para la construcción del índice; el resultado del test para número de personas en el hogar es un poco bajo en comparación con las otras variables pero sigue estando por encima de 0,5 que es lo deseable. (B.G Tabachnick, 2007)

---

<sup>4</sup> Según Kaiser (1974) se da la siguiente calificación a los valores de *KMO*:

0.00 to 0.49 Unacceptable

0.50 to 0.59 Miserable

0.60 to 0.69 Mediocre

0.70 to 0.79 Middling

0.80 to 0.89 Meritorious

0.90 to 1.00 Marvelous

**Tabla 4 Matriz de Correlación de las Variables incluidas en el INSE**

	Educación Madre	Educación Padre	Estrato	Hacinamiento	Nivel de Ingresos	Material de los Pisos	Ocupación Madre	Ocupación Padre	Núm Personas en el Hogar	Dotación del Hogar	Área en la que vive
Educación Madre	1,0000										
Educación Padre	0,6763	1,0000									
Estrato	0,5119	0,5180	1,0000								
Hacinamiento	0,2411	0,2280	0,2953	1,0000							
Nivel de Ingresos	0,5603	0,5642	0,6301	0,2743	1,0000						
Material de los Pisos	-0,2860	-0,2862	-0,4076	-0,1836	-0,3643	1,0000					
Ocupación Madre	0,3957	0,4027	0,3172	0,1791	0,3705	-0,1941	1,0000				
Ocupación Padre	0,3928	0,5113	0,3923	0,1821	0,4467	-0,2257	0,6810	1,0000			
Número de Personas en el Hogar	0,1273	0,1125	0,1419	0,2814	0,0983	-0,0824	0,0852	0,0819	1,0000		
Dotación del Hogar	0,4188	0,4119	0,5155	0,2995	0,5315	-0,3588	0,3672	0,3390	0,1065	1,0000	
Área en la que vive	0,2123	0,2124	0,2219	0,1043	0,2223	-0,1437	0,1215	0,1458	0,0563	0,2433	1,0000

Fuente: Elaboración propia.

Como se observa en el resultado del test, los valores están cercanos a 1, lo cual es ideal ya que se considera que por encima de 0,5 es buena la adecuación muestral para un análisis de Componentes Principales (Universidad Complutense Madrid, 2014). Este resultado significa que las variables están suficientemente correlacionadas entre sí y el método es adecuado.

**Tabla 5 Test de KMO para evaluar la adecuación muestral**

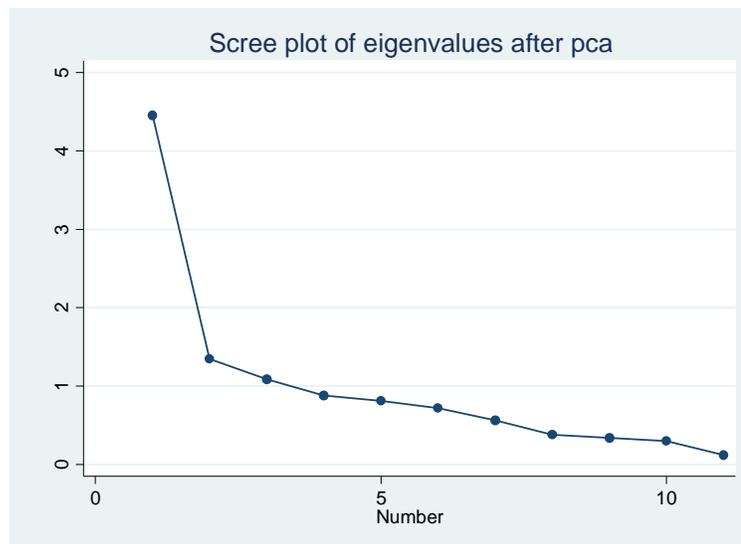
Variable	kmo
Área en la que vive	0,9573
Educación Madre	0,8320
Educación Padre	0,8350
Estrato	0,8990
Hacinamiento	0,8395
Nivel de Ingresos	0,8931
Material de los Pisos	0,9202
Ocupación Madre	0,8661
Ocupación Padre	0,8872

<b>Número de Personas en el Hogar</b>	0,6712
<b>Dotación del Hogar</b>	0,9144
<b>Overall</b>	0,8713

Fuente: Elaboración propia

Con el propósito de verificar hasta que componente se debe considerar para la construcción del índice se realiza un *scree test* en STATA, en la Ilustración 2 se observa un punto para cada uno de los eigenvalues de los factores. El punto en el cual la pendiente disminuye notablemente indica hasta que componente considerar, en este caso se debe tener en cuenta solo la primer componente principal.

Ilustración 2 Scree Test

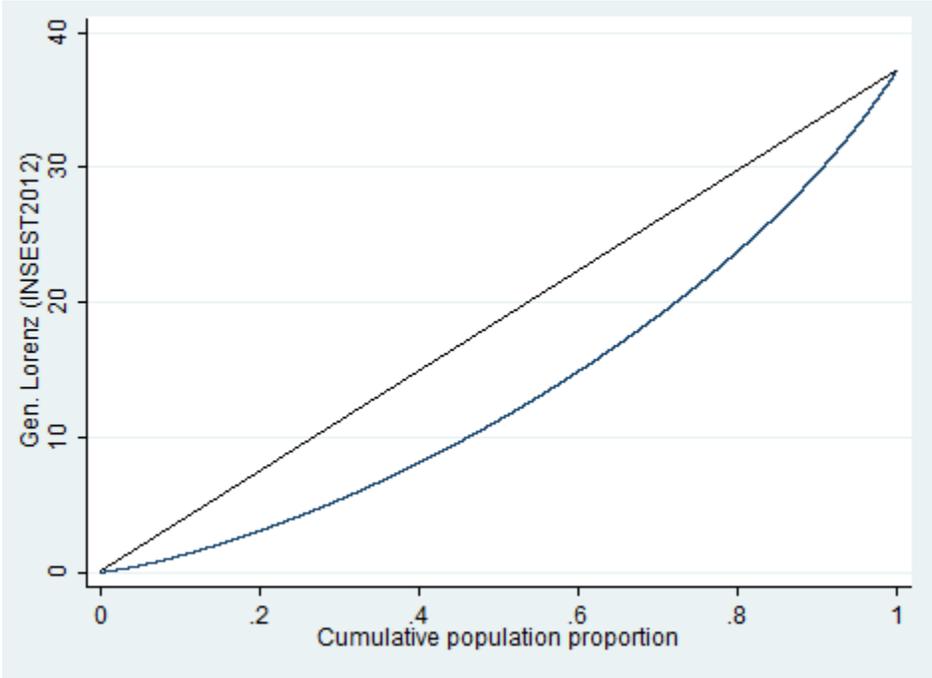


Fuente: Elaboración propia.

Una vez construido el índice se calcula un indicador Gini para el nivel socioeconómico y otro para los puntajes de Lenguaje y Matemáticas con el comando *sgini* para cada año, y se procede a comparar el cambio en el Gini de nivel socioeconómico con el cambio en Gini de puntajes.

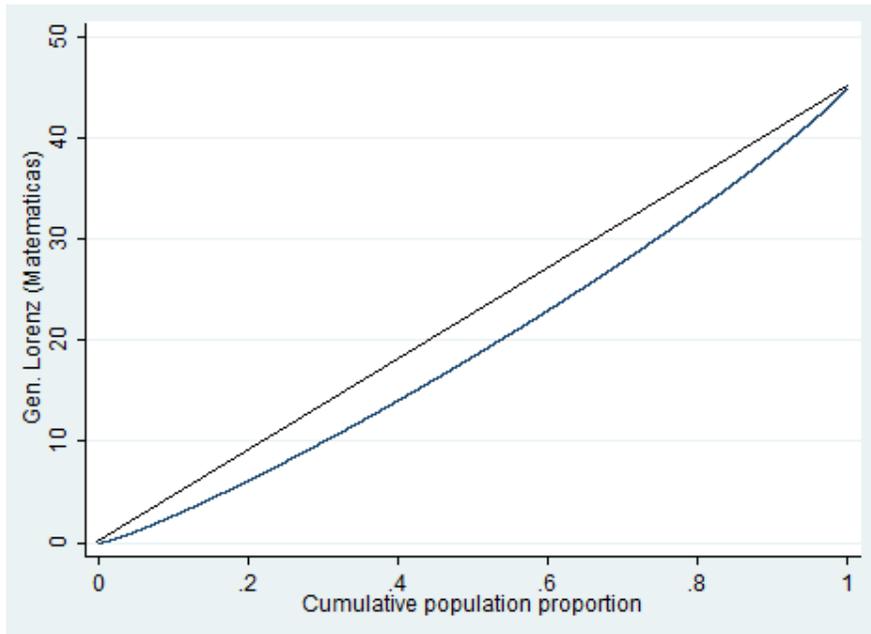
Las curvas de Lorenz correspondientes al Gini para el INSE, nivel de puntajes en matemáticas y nivel de puntajes en lenguaje para el año 2012 se observan en las Ilustraciones 3, 4 y 5 respectivamente.

Ilustración 3 Curva de Lorenz para el Índice de Nivel Socioeconómico del año 2012



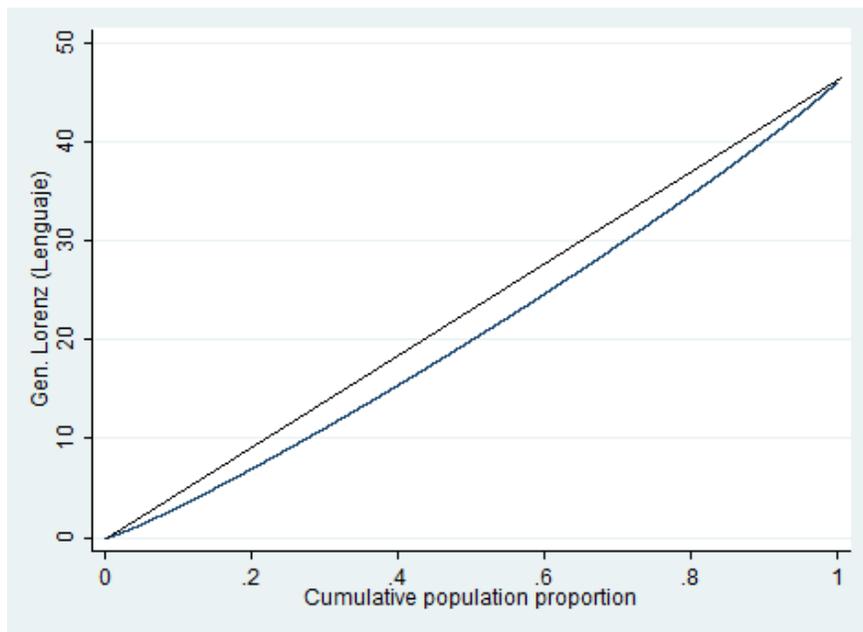
Fuente: Elaboración propia.

Ilustración 4 Curva de Lorenz para los puntajes en el área de Matemáticas en la prueba SABER 11 del año 2012



Fuente: Elaboración propia.

Ilustración 5 Curva de Lorenz para los puntajes en el área de Lenguaje en la prueba SABER 11 del año 2012



Fuente: Elaboración propia.

Finalmente se realiza un ejercicio de desagregación del Gini de puntajes en subpoblaciones de alumnos según género, condición de ruralidad, estrato y dependencia administrativa (si el colegio es público o privado) y se analizan los resultados utilizando gráficas de dispersión de Gini de puntajes versus promedio de los resultados de las pruebas.

## 7 DATOS

Los datos usados para estimar los coeficientes de desigualdad son los resultados el examen SABER 11 en Colombia en el período 2008-2012, así como las respuestas a las preguntas sobre información socioeconómica incluidas en el formulario de inscripción al examen. Las variables utilizadas para construir el coeficiente de Gini de nivel socioeconómico se pueden observar en el ANEXO 1 y se toman los puntajes obtenidos por los estudiantes en todo el país en Lenguaje y Matemáticas, para la construcción del Gini de puntajes. La base de datos construida para este estudio contiene 55 variables y 2´759.269 observaciones (en esta base sólo se incluyen los resultados de los estudiantes que presentan la prueba por primera vez).

En las Tablas 6, 7, 8, 9, 10 y 11 se observa la distribución de la muestra según las categorías para cada variable:

**Tabla 6 Distribución de la muestra según Género**

Género	Freq.	Porcentaje
Femenino	1.510.112	54,73%
Masculino	1.249.155	45,27%
Total	2.759.267	100,00%

Fuente: Elaboración propia

**Tabla 7 Distribución de la muestra según Área en la que vive**

Área en la que vive	Freq.	Porcentaje
Rural	575.016	20,84%
Urbano	2.183.571	79,16%
Total	2.758.587	100,00

Fuente: Elaboración propia

**Tabla 8 Distribución de la muestra según Estrato**

Estrato	Freq.	Porcentaje
1	1.007.172	36,59%
2	1.028.311	37,36%
3	537.511	19,53%
4	111.904	4,07%
5	42.731	1,55%
6	25.069	0,91%
Total	2.752.698	100,00%

Fuente: Elaboración propia

**Tabla 9 Distribución de la muestra según Tipo de Establecimiento**

Tipo de Establecimiento	Freq.	Porcentaje
Privado	776.472	28,14%
Público	1.982.777	71,86%
Total	2.759.249	100,00%

Fuente: Elaboración propia

**Tabla 10 Distribución de la muestra según Nivel de estudios del padre y de la madre**

Nivel de Estudios	Padre		Madre	
	Freq.	Porcentaje	Freq.	Porcentaje
Educ Profesional Completa	274.893	9,97%	272.244	9,87%
Educ T&T Completa	107.384	3,89%	133.258	4,83%
Educ T&T Incompleta	32.193	1,17%	42.768	1,55%
Ninguno	105.280	3,82%	65.916	2,39%
No sabe	168.051	6,09%	81.158	2,94%
Postgrado	57.974	2,10%	54.233	1,97%
Primaria Completa	501.554	18,18%	508.388	18,43%

Primaria Incompleta	516.320	18,72%	462.860	16,78%
Secundaria Completa	600.607	21,77%	668.448	24,23%
Secundaria Incompleta	394.301	14,29%	469.284	17,01%
Total	2.758.557	100,00%	2.758.557	100,00%

Fuente: Elaboración propia

**Tabla 11 Distribución de la muestra según Ingreso Familiar**

Ingreso Familiar	Freq.	Porcentaje
Entre 1 y 2 SM	1.197.240	43,40%
Entre 2 y 3 SM	394.407	14,30%
Entre 3 y 5 SM	183.765	6,66%
Entre 5 y 7 SM	61.266	2,22%
Más de 7 SM	73.211	2,65%
Menos de 1 SM	848.508	30,76%
Total	2.758.397	100,00%

Fuente: Elaboración propia

## 8 ANÁLISIS DE RESULTADOS

Los resultados encontrados fueron analizados según la Tabla 12, donde se ve que la situación óptima es aquella donde existe un mayor nivel promedio y una menor desigualdad (cuadrante superior, izquierdo).

**Tabla 12 Matriz de Resultados Promedio vs Gini de Puntaje en la Prueba SABER 11**

<b>Resultado promedio</b>	Resultado ideal, donde hay un mayor puntaje promedio y menor desigualdad.	Puntaje promedio alto y desigualdad alta. Existen variables que pueden afectar los resultados de los estudiantes.
	Puntaje promedio bajo y desigualdad baja, esta situación se puede mejorar con políticas públicas enfocadas en mejorar la calidad de la educación.	Puntaje promedio bajo y desigualdad alta. Aquí pueden haber variables externas que están afectando los resultados.
<b>Gini</b>		

Fuente: Elaboración propia

En el cuadrante inferior, izquierdo, las políticas públicas deberán estar enfocadas a mejorar la calidad de la educación, pues ni siquiera los estudiantes más destacados están logrando puntajes altos y por esto la desigualdad es baja. Para los dos cuadrantes de la derecha, pueden existir factores económicos y sociales que llevan a una mayor desigualdad.

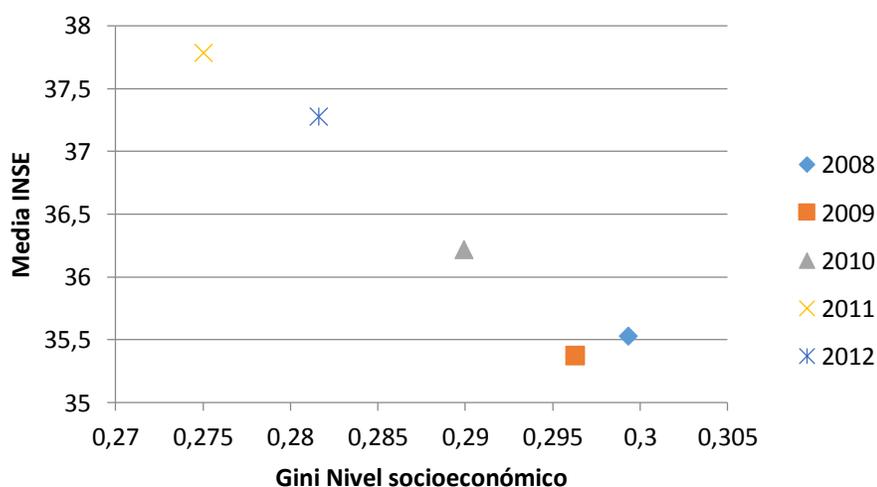
Los datos que se muestran en la Tabla 13 representan los resultados encontrados para el índice Gini de condiciones socioeconómicas de los estudiantes que presentaron las pruebas para cada uno de los años a estudiar. Se puede ver que es un índice que refleja una desigualdad en los ingresos más o menos constante y alta aunque no en la misma medida en estimaciones de Gini de ingresos para el total de la población colombiana, el cual se ubica en un 0.53 aproximadamente.

**Tabla 13** Índice Gini de Nivel socioeconómico y media del INSE

ESTADÍSTICO	2008	2009	2010	2011	2012
<b>GINI NIVEL SOCIOECONÓMICO</b>	0.2993	0.2963	0.2899	0.275	0.2816
<b>MEDIA INSE</b>	35.5271	35.3721	36.2177	37.781	37.277

Fuente: Elaboración propia

**Ilustración 6** Gini Nivel Socioeconómico



Fuente: Elaboración propia

En la Ilustración 3 se puede observar una mejoría en el comportamiento del Gini de nivel socioeconómico y el INSE, aunque es leve, se ve que a través del tiempo está disminuyendo la desigualdad, con excepción del último año (2012), al igual que está mejorando el INSE, lo que refleja mejores condiciones de vida para los estudiantes. El dato del año 2012 puede estar dado por una situación económica a nivel internacional que no ha mejorado, y por una disminución en el crecimiento que se venía teniendo en los años anteriores.

### 8.1 Desigualdad en los puntajes de las pruebas Saber 11

En la tabla 14 se muestra el Gini de resultados de las pruebas Saber 11 para Lenguaje y Matemáticas, allí se puede ver claramente que este índice está muy por debajo del INSE visto anteriormente. Esta alrededor de 0.09 para lenguaje y 0.12 para matemáticas, lo que sería de esperarse, pues generalmente las personas tienden a tener más dificultades y menores resultados en esta área del conocimiento. Adicionalmente se muestran los puntajes promedio para cada uno de los años, y en la misma medida se ve que hay un mejor desempeño en el área de lenguaje. Es importante destacar que los puntajes del año 2011 tienen un comportamiento atípico respecto a los demás años. No hay una explicación lógica a este hecho, pero se ve claramente un menor resultado acompañado de una mayor desigualdad para todas las pruebas.

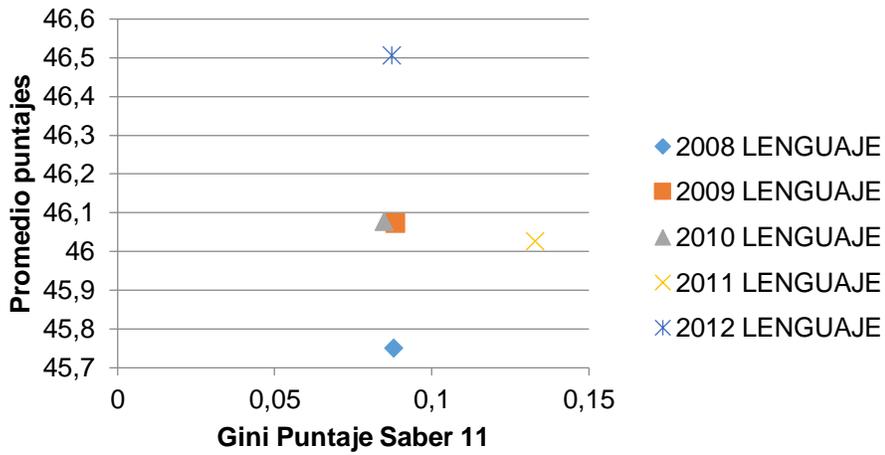
Se nota una tendencia creciente en los índices de GINI para las dos áreas de conocimiento, a pesar que las medias de los puntajes parecen ir aumentando.

Tabla 14 Gini para Puntaje Saber 11 en el área de Lenguaje

ESTADÍSTICO	LENGUAJE				
	2008	2009	2010	2011	2012
<b>GINI</b>	0.0879	0.0884	0.0848	0.1329	0.0873
<b>PROMEDIO</b>	45.751	46.0719	46.0773	46.0261	46.5062

Fuente: Elaboración propia

**Ilustración 7** Gini para Puntaje Saber 11 en el área de Lenguaje



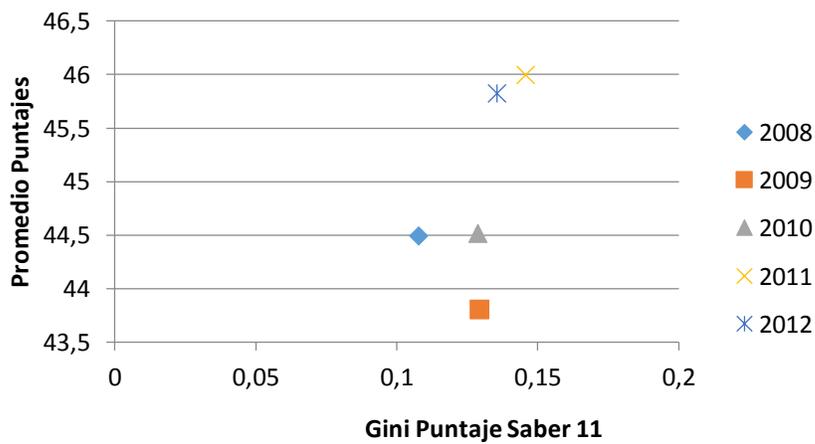
Fuente: Elaboración propia

**Tabla 15** Gini para Puntaje Saber 11 en el área de Matemáticas

ESTADÍSTICO	MATEMÁTICAS				
	2008	2009	2010	2011	2012
<b>GINI</b>	0.1078	0.1293	0.1288	0.1457	0.1356
<b>PROMEDIO</b>	44.4923	43.8047	44.5181	45.9948	45.8218

Fuente: Elaboración propia

**Ilustración 8** Gini para Puntaje Saber 11 en el área de Matemáticas



Fuente: Elaboración propia

Adicionalmente en las Ilustraciones 4 y 5 se ve una mejoría en el promedio de los estudiantes, a pesar que la desigualdad permanece más estable en este período. Esto nos podría indicar según la Tabla 12, que el gobierno está trabajando por que el desempeño de los estudiantes mejore, y en el largo plazo podríamos llegar a un nivel en el que haya una baja desigualdad acompañada de un buen promedio en los resultados de las pruebas, lo que sería el caso ideal.

## **8.2 Descomposición del Índice Gini en subpoblaciones**

A pesar de encontrar que la desigualdad en puntajes de las Pruebas Saber 11 es relativamente baja, se quiso profundizar y comprobar en qué sectores de la población puede haber una mayor desigualdad y pueden ser más vulnerables, y es precisamente en estos sectores donde deberían estar enfocadas las políticas públicas educativas para reducir la brecha y garantizar una mejor calidad y cobertura educativa.

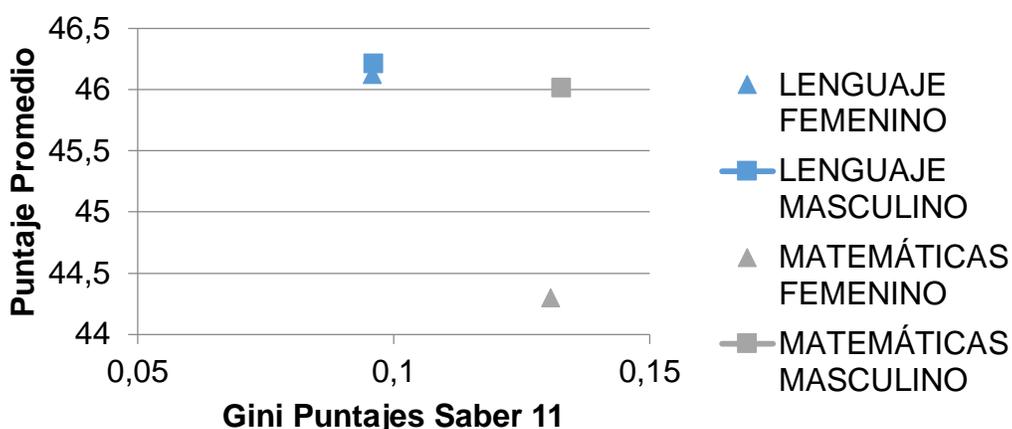
Para esto se hizo una descomposición del Gini de puntajes según el género, estrato socioeconómico, ruralidad y si la institución es pública o privada. Estos resultados se compararon con el promedio de resultados de los estudiantes para garantizar que una menor desigualdad en una subpoblación no esté relacionada con un menor desempeño promedio. Para esto, la situación ideal sería para aquellos que en las gráficas que se presentarán a continuación estén situados en el extremo izquierdo (menor desigualdad y mayor resultado promedio).

### **8.2.1 Segmentación por Género**

La primera segmentación que se hizo fue según el género de los estudiantes, aquí se puede ver que en lenguaje existe un comportamiento muy similar para ambos géneros tanto en desigualdad como en puntaje promedio. Igualmente se observa en la Ilustración 6 un menor puntaje promedio del género femenino acompañado de una menor desigualdad en las pruebas de matemáticas.

Para el género masculino se presenta una baja desigualdad acompañado de un bajo puntaje promedio en ambas asignaturas, mientras que para el género femenino se muestra un desempeño más disperso, hay una mayor brecha entre las pruebas de matemáticas y lenguaje.

Ilustración 9 Segmentación del Gini de puntajes por Género

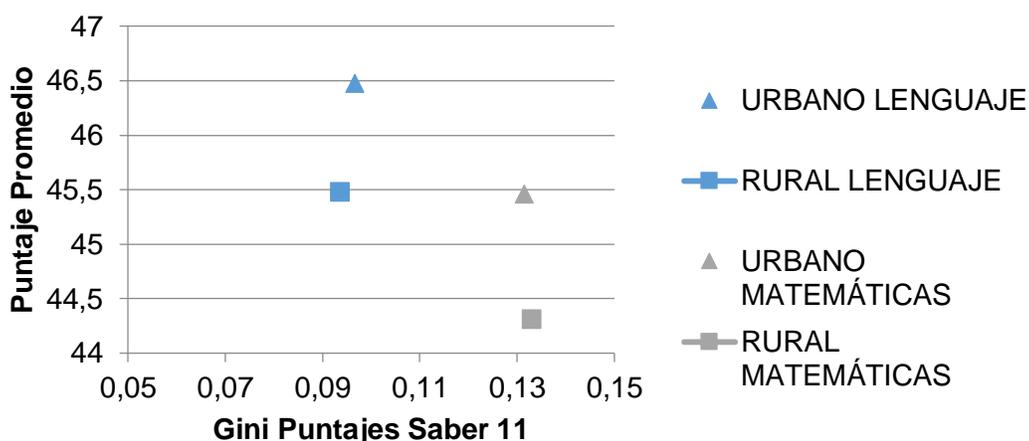


Fuente: Elaboración propia

### 8.2.2 Segmentación por condición de ruralidad

Cuando se estudian los resultados de acuerdo a la condición de ruralidad del establecimiento donde se encuentran los estudiantes se observa una desigualdad muy similar para cada una de las pruebas. Se ve un mejor desempeño promedio en el área urbana, lo cual es coherente con la teoría, pues allí se encuentran más colegios privados, una mayor inversión, y mejores establecimientos en general.

Ilustración 10 Segmentación del Gini de puntajes por Condición de Ruralidad



Fuente: Elaboración propia

Para la zona rural se presenta una menor desigualdad de puntajes, acompañada de un menor desempeño promedio. Este es el caso del cuadrante inferior izquierdo, donde el gobierno podría intervenir con mejoras a la calidad y cobertura de la educación para solucionar esta problemática. Este caso no es deseable, pero puede tener una solución si se interviene efectivamente. Se puede inferir que los establecimientos de educación no están cumpliendo su función en el área rural, pues todos llegan con condiciones muy similares y ningún está obteniendo una posición destacada pues no se ve desigualdad, ni desempeño superior.

### 8.2.3 Segmentación por estrato socioeconómico

Se realizó el mismo ejercicio anterior, para observar el comportamiento de los resultados y el Gini para cada estrato socioeconómico del país, y con esto se encontró que en los estratos 1, 2 y 3 (niveles más bajos) existen menores resultados

promedio y mayor desigualdad. Para los niveles 4, 5 y 6 existen mejores resultados, y un Gini más bajo.

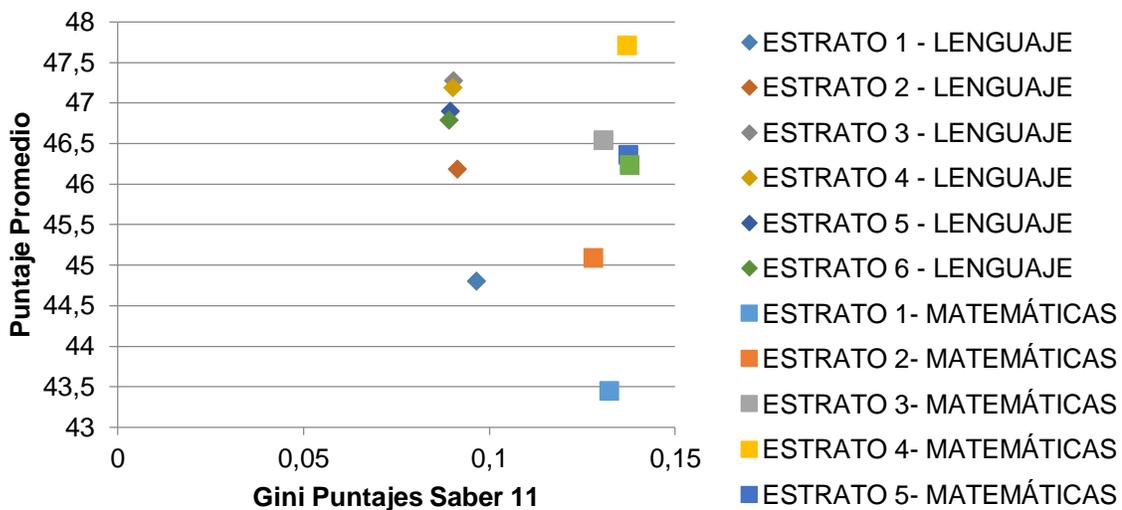
En el estrato 1 se evidencian los resultados promedio más bajos y la mayor desigualdad, aquí es donde mayores dificultades tienen los estudiantes y más intervención de las políticas públicas debería haber para mejorar el desempeño académico y empezar a cerrar brechas que pueden beneficiar a los estudiantes y al desarrollo del país.

En el estrato 2, a pesar que se observa la menor desigualdad en matemáticas, esta se da a costa de un bajo rendimiento, esta situación no es deseable, pero puede ser corregida a través de estímulos a la calidad de la educación.

Para el estrato 3, se observa un mejor desempeño tanto en promedio como en desigualdad respecto a los 2 niveles socioeconómicos anteriores.

Para el estrato 5 y 6 se observa un resultado promedio similar, pero están por debajo del estrato 4. En desigualdad, el área de lenguaje tiene el menor Gini, pero el área de matemáticas tiene una desigualdad más alta respecto a los otros estratos.

Los resultados encontrados para los estratos 4, 5 y 6 pueden estar explicados por nivel socioeconómico, por el tipo de establecimiento al que estas personas asisten, que en general es privado, y en particular para el estrato 4 puede estar dado por una necesidad puntual de obtener mejores resultados para ingresar a una universidad pública o acceder a una beca. Para corroborar esta hipótesis habría que realizar investigaciones posteriores.



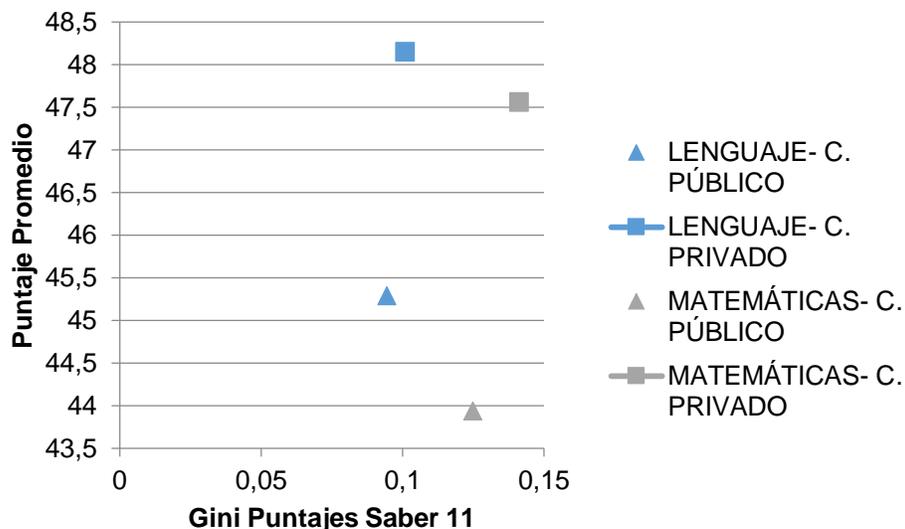
Fuente: Elaboración propia

### 8.2.4 Segmentación por tipo de establecimiento (público-privado)

Se realizó la segmentación del Gini según el tipo de establecimiento al que asistieron los estudiantes (público o privado) y se encontró que los estudiantes que asistieron a un establecimiento público tuvieron menores resultados en ambas áreas de estudio (matemáticas y lenguaje) y menor desigualdad.

Adicional se puede decir que existe una menor brecha en la prueba de lenguaje tanto en los establecimientos públicos como privados sobre la prueba de matemáticas.

Ilustración 12 Segmentación del Gini de puntajes por Tipo de establecimiento



Fuente: Elaboración propia

## 9 CONCLUSIONES

En este documento se evalúa un índice de Gini para los puntajes de las pruebas Saber 11 en el período 2008-2012. Se encuentra que la desigualdad de puntajes de estas pruebas es menor que el Gini de nivel socioeconómico también construido en este trabajo.

Posteriormente se hace una desagregación del Gini de puntajes según 4 criterios: género del estudiante, condición de ruralidad, estrato socioeconómico y tipo de establecimiento. Este Gini es comparado con el puntaje promedio de cada uno de los grupos poblacionales, para corroborar que una menor desigualdad no se dé a costa de un menor puntaje promedio en las pruebas. Se encuentra que para cada uno de los grupos, la condición más desfavorable se da en el área de matemáticas, donde se obtiene un menor puntaje promedio y mayor desigualdad.

A pesar de que no hay una relación directa entre los puntajes de los estudiantes y la desigualdad, se ve claramente que para el caso colombiano, la menor desigualdad se da acompañada de un menor nivel promedio en las pruebas. Esto

no es un resultado deseable, pues se espera que la menor desigualdad en puntajes, a pesar de la alta desigualdad en nivel socioeconómico, este dada por una mejor calidad de la educación y unas políticas que estén enfocadas en disminuir estas brechas, pero el sistema no está logrando que los estudiantes obtengan mejores resultados ni más homogéneos.

También es importante recalcar que a pesar de que la desigualdad encontrada no es muy grande en relación con la desigualdad de nivel socioeconómico, si vale la pena estudiar los resultados y hacer un seguimiento futuro, para poder ver su evolución y construir políticas públicas que puedan ayudar a cerrar estas brechas, para garantizar las mismas oportunidades a todos los estudiantes independientemente de sus antecedentes familiares (nivel socioeconómico). Sería importante realizar un seguimiento futuro de estos resultados porque la desigualdad es una forma para analizar los resultados de una variable, en este caso educación, y mirar cómo es su comportamiento como un conjunto. Una mayor desigualdad es perjudicial en casi cualquier aspecto, es por esto que el sistema educativo debería estar enfocado en acciones que mejoren no solo el nivel de resultados promedio sino también enfocarse en políticas para reducir la desigualdad en los resultados.

Colombia en particular es un país que ha tenido niveles altos de desigualdad en términos de ingreso, y esto ha llevado a desequilibrios en muchos otros aspectos, como la educación. A partir del estudio se puede decir que se ve una leve mejoría en los resultados y la desigualdad para los puntajes de las pruebas Saber 11, sin embargo se sigue evidenciando un bajo rendimiento en las pruebas, acompañado de una baja desigualdad para casi todos los casos estudiados, lo que nos indica que el gobierno está trabajando de alguna manera en políticas públicas enfocadas a la educación, pero aún se debe trabajar para disminuir la desigualdad en educación siempre acompañado de una mejora en los puntajes promedio de los estudiantes, a través de mejoras en la calidad de la educación media y básica.

## 10 BIBLIOGRAFÍA

B.G Tabachnick, & L. (2007). *Using multivariate statistics*. Boston: Pearson/Allyn & Bacon.

Barrera, M. D. (2014). *La educación básica y media en Colombia: Retos en equidad y calidad* . Bogotá.

Coleman, J. S. (1996). *Equality of Educational Opportunity*. Washington.

Gamboa, L. F. (2014). *Pre-School contributions to future achievements*. Bogotá.

Gamboa, L. F., & Waltenberg, F. D. (2011). *Inequality of opportunity in educational achievement in Latin America: Evidence from PISA 2006-2009*.

Gamboa, L. F., Casas, A. F., & Piñeros, L. J. (2003). *La teoría del valor agregado: Una aproximación a la calidad de la educación en Colombia*. Bogotá.

Vásquez, M., Ramírez, G., & Camardiel, A. (2001). Un procedimiento para la cuantificación de variables cualitativas basado en los principios de la promediación recíproca y el análisis de homogeneidad. *Revista Venezolana de Análisis de Coyuntura*.

Meckes, L. (24 de Octubre de 2014). *SIMCE: Sus propósitos y características*. Obtenido de SIMCE: Sus propósitos y características: <http://www.isei-ivei.net/cast/pub/simce/simce%20general.pdf>

Mélendez, F. H. (2012). *Desigualdad de Resultados y Oportunidades en Colombia: 1997- 2010*. Bogotá.

Nacional, M. d. (25 de Julio de 2014). *Ministerio de Educación Nacional*. Obtenido de <http://www.mineduacion.gov.co/1621/w3-article-244735.html>

OECD. (2012). *Pisa 2012 results: what makes schools successful? Resources, policies and practices (Vol IV)*. Paris.

Ramírez, M. J. (2007). *Diferencias dentro de las salas de clases: Distribución del rendimiento en matemáticas*.

Santos, M. E. (2007). *Calidad de la educación en Argentina: determinantes y distribución utilizando los resultados de PISA 2000*. Buenos Aires.

Agencia de Calidad de la Educación . (2013). *Gini al SIMCE: Una aplicación del índice de Gini a los puntajes de las pruebas*. . Santiago de Chile.

CEPAL. (2012). *Cepal*. Obtenido de <http://www.cepal.org>

Departamento Nacional de Planeación . (2002). *El nuevo Índice de Condiciones de Vida*. Bogotá.

ICFES. (2010). *Metodología de construcción del Índice de nivel socioeconómico de los estudiantes*. Bogotá.

Kaiser. (1974).

# 11 ANEXOS

## Anexo 1 Variables Utilizadas para la Construcción del INSE

VARIABLE	2007-2	2008-1	2008-2	2009-1	2009-2	2010-1	2010-2
Alfabetismo de padre	X	X	X				
Alfabetismo de madre	X	X	X				
Nivel educativo del padre	X	X	X	X	X	X	X
Nivel educativo de la madre	X	X	X	X	X	X	X
Ocupación del padre		X	X	X	X	X	X
Ocupación de la madre		X	X	X	X	X	X
Nro. de hermanos		X	X				
Max nivel educativo de algún hermano mayor	X	X					
Nivel educativo del hermano mayor con más educación			X	X			
Asistencia escolar de hermanos de 5 a 15 años	X	X	X				
Zona de residencia (cabecera municipal / resto)	X	X	X				
Estrato socioeconómico de la vivienda donde reside	X	X	X	X	X	X	X
Nivel SISBEN		X	X	X	X	X	X
Nro. total de personas que conforman el hogar	X	X	X	X	X	X	X
Nro total de cuartos de que dispone el hogar	X	X	X	X	X		
Nro total de cuartos en que duermen las personas del hogar	X	X	X	X	X	X	X
Material predominante en el piso de la vivienda	X	X	X	X	X	X	X
Uso o compartido o exclusivo del servicio sanitario con que cuenta el hogar				X			
Conexión del servicio sanitario a alcantarillado o pozo séptico				X	X		
Total ingresos mensuales del hogar (en SMLV)		X	X	X	X	X	X
Trabaja actualmente		X	X	X	X(*)	X(*)	X(*)
Nro. de horas / semana trabaja		X	X	X			
<b>DOTACIÓN</b>							
Tiene computador (SI / NO)	X						
Tiene computador (con / sin internet)		X					
<i>El hogar dispone de:</i>							
Computador			X	X	X	X	X
TV color			X	X			
Reproductor DVD			X	X	X	X	X
Automóvil particular			X	X	X	X	X
Motocicleta			X				
Celular			X		X	X	X
Teléfono fijo				X	X	X	X
Máquina lavadora de ropa				X	X	X	X
Horno microondas				X	X	X	X
Horno eléctrico o a gas				X	X	X	X
Nevera o enfriador				X	X	X	X
Estufa eléctrica o a gas				X			
<i>Además el hogar cuenta con:</i>							
Conexión a internet			X	X	X	X	X
Servicio cerrado de TV			X	X	X	X	X
Energía Eléctrica				X			
Acueducto				X			
Alcantarillado				X			
Recolección de basuras				X			

(\*) Se incorporó a las alternativas de respuesta de si trabaja, el número de horas (NO trabaja; Si, menos de 20 horas /semana; Si, más de 20 horas / semana)

Fuente: Departamento Nacional de Planeación-Programa de las Naciones Unidas para el Desarrollo. Programa Nacional de Desarrollo Humano. El nuevo índice de Condiciones de Vida. Bogotá. 2002.

En la tabla se presenta la lista de variables utilizadas por el ICFES para la construcción del Índice de Nivel Socioeconómico de los Estudiantes, la información es tomada del formulario de inscripción al examen SABER 11. El DNP selecciona variables en las que se evidencia una alta diferenciación entre la población y para algunos casos descarta variables que no adicionan variabilidad o que tienen un peso

poco significativo dentro del índice como por ejemplo variables con información sobre los hermanos del estudiante, bienes de uso doméstico y trabajo del estudiante.

#### Anexo 2 Resultado de la Cuantificación de las Variables Cualitativas

Variable	Categoría	Puntaje
AREA VIVE	Rural	-
	Cabecera municipal	3,3529
DOTACION HOGAR	Ninguno	-
	1 bien / servicio	0,0928
	2 bienes / servicios	0,9815
	3 bienes / servicios	1,7691
	4 bienes / servicios	2,6022
	5 bienes / servicios	3,5036
	6 bienes / servicios	4,4498
	7 bienes / servicios	5,4895
	8 bienes / servicios	6,6175
	9 bienes / servicios	7,7992
	10 bienes / servicios	9,2838
11 bienes / servicios	12,0843	
EDU_MADRE	Primaria incompleta	1,0113
	Primaria completa	2,5229
	Secundaria incompleta	2,9753
	Secundaria completa	4,2936
	Educación TóT incompleta	5,4317
	Educación TóT completa	7,1938
	Educación profesional incompleta	7,7796
	Educación profesional completa	9,8162
	Postgrado	11,7792

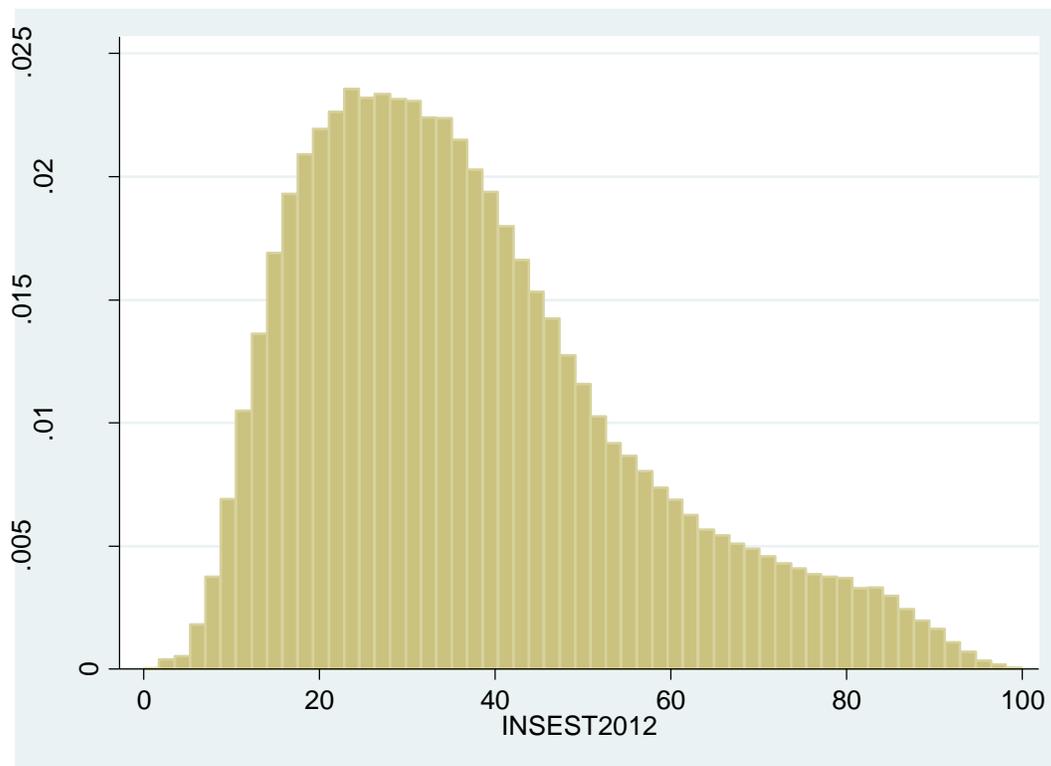
Fuente: ICFES. Dirección de Evaluación. Subdirección de Análisis y Difusión de Resultados. Metodología de Construcción del Índice de Nivel Socioeconómico de los Estudiantes. Bogotá. 2010.

EDU_PADRE	Primaria incompleta	0,9594
	Primaria completa	2,4341
	Secundaria incompleta	2,6427
	Secundaria completa	4,0784
	Educación TóT incompleta	5,6598
	Educación TóT completa	6,9408
	Educación profesional incompleta	7,6417
	Educación profesional completa	9,3260
	Postgrado	12,1162
ESTRATO	1	1,7577
	2	4,6632
	3	8,1191
	4	12,2935
	5	13,7841
	6	14,6955
HACINAM (presonas / cuarto)	5 y +	-
	4 - 4,99	0,8344
	3 - 3,99	2,0690
	2 - 2,99	3,5439
	1 - 1,99	5,0907
INGRESOS /MES HOGAR	Menos de 1 SM	-
	Entre 1 y menos de 2 SM	2,8131
	Entre 2 y menos de 3 SM	5,8911
	Entre 3 y menos de 5 SM	8,8103
	Entre 5 y menos de 7 SM	11,1658
	Entre 7 y menos de 10 SM	12,5574
	10 o más SM	13,6846
MAT_PISOS	Tierra, arena	-
	Cemento, gravilla, ladrillo	2,7152
	Madera burda, tabla o tablón	3,8510
	Madera pulida, baldosa, mármol, alfombra	7,6101
OCU_MADRE	Hogar	-
	Otra actividad u ocupación	0,7689
	Trabajador por cuenta propia	1,1739
	Obrero u operario empleado(a)	2,1912
	Pensionado(a)	4,2938
	Pequeño Empresario	4,5343
	Empleado(a) de nivel auxiliar o administrativo	4,5629
	Empleado(a) de nivel directivo	5,4493
	Profesional independiente	6,2384
	Empleado(a) de nivel técnico / profesional	6,6660
	Empleado con cargo como director(a) o gerente	7,2615
Empresario	7,8232	
OCU_PADRE	Hogar	-
	Trabajador por cuenta propia	0,2502
	Otra actividad u ocupación	0,9342
	Obrero u operario empleado(a)	1,6913
	Empleado(a) de nivel auxiliar o administrativo	3,5416
	Pequeño Empresario	3,5455
	Pensionado(a)	4,6070
	Profesional independiente	5,5021
	Empleado(a) de nivel técnico / profesional	7,0363
	Empleado(a) de nivel directivo	7,1727
	Empleado con cargo como director(a) o gerente	7,2963
Empresario	8,0475	
PERS_HOGAR	7 y +	-
	6	0,4973
	5	0,5092
	4	2,0759
	3	3,1840
	2	3,7158

Fuente: ICFES. Dirección de Evaluación. Subdirección de Análisis y Difusión de Resultados. Metodología de Construcción del Índice de Nivel Socioeconómico de los Estudiantes. Bogotá. 2010.

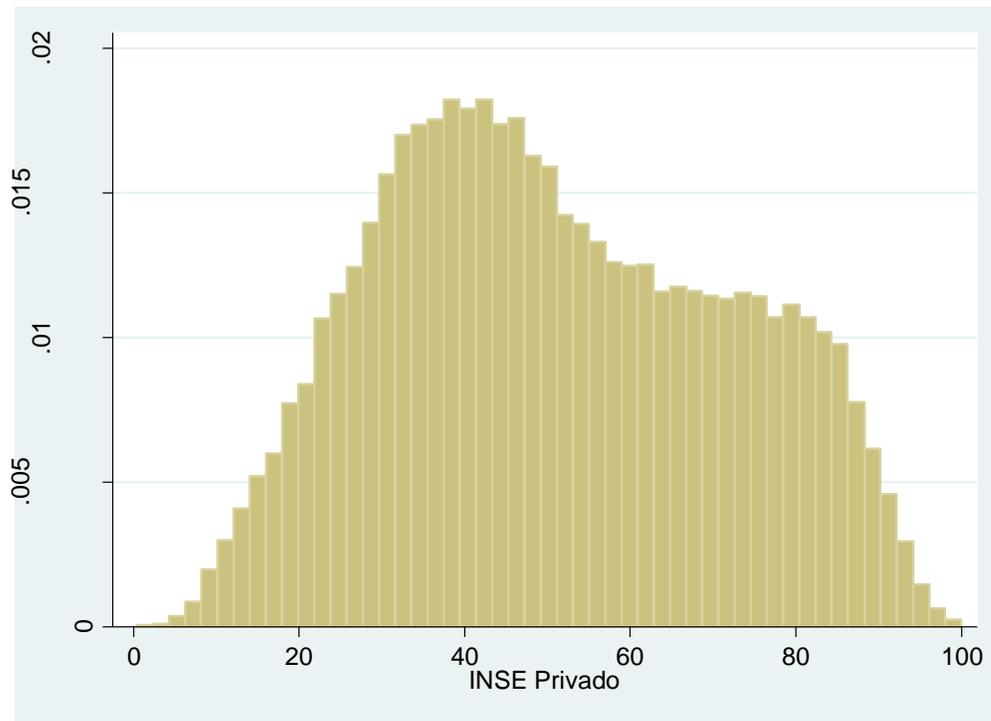
Anexo 3 Distribución del Índice de Nivel Socioeconómico

a)



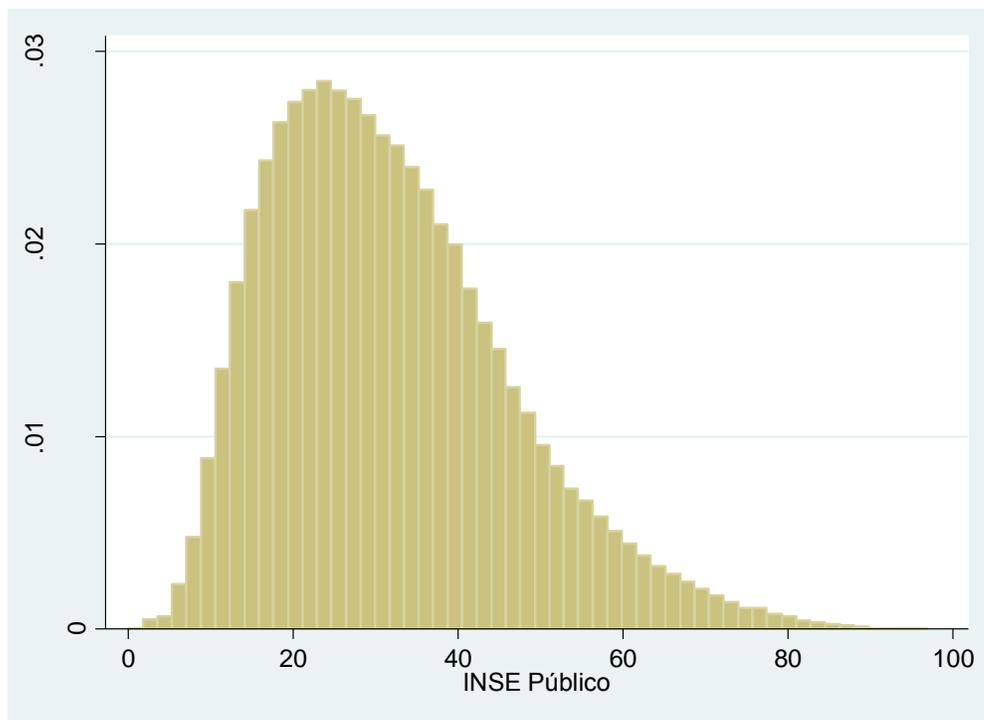
Fuente: Elaboración propia.

b)



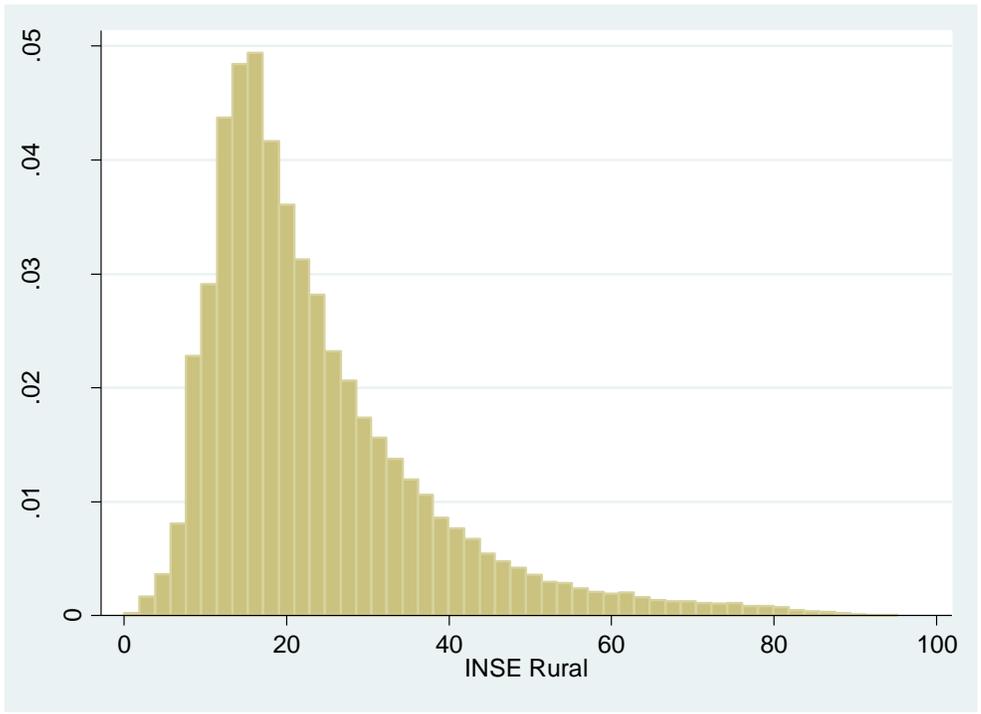
Fuente: Elaboración propia.

c)



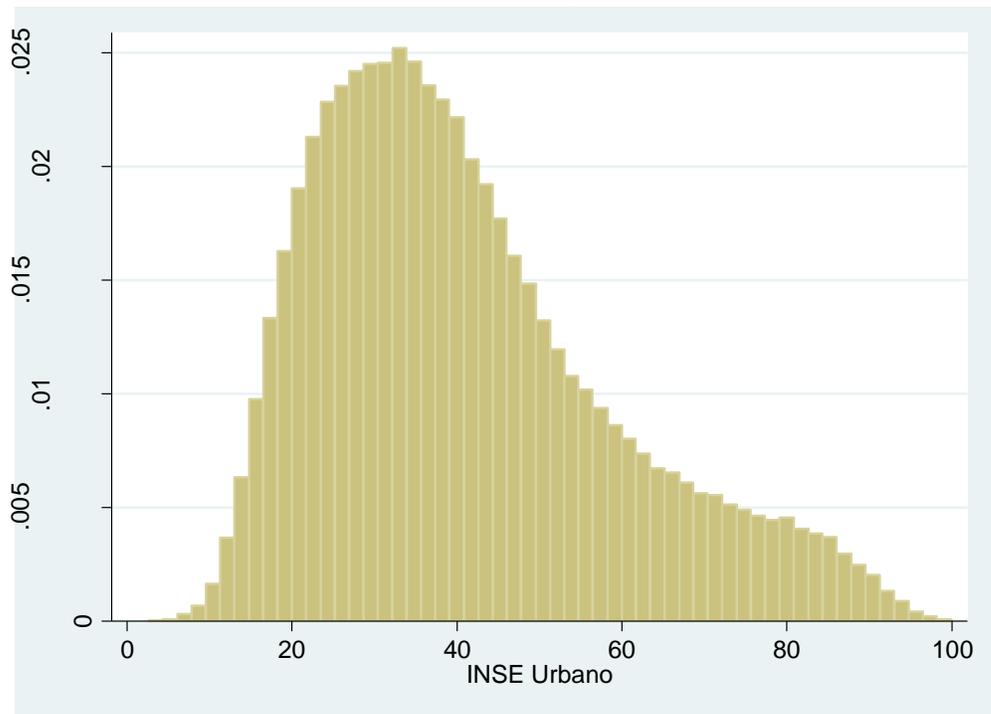
Fuente: Elaboración propia.

d)



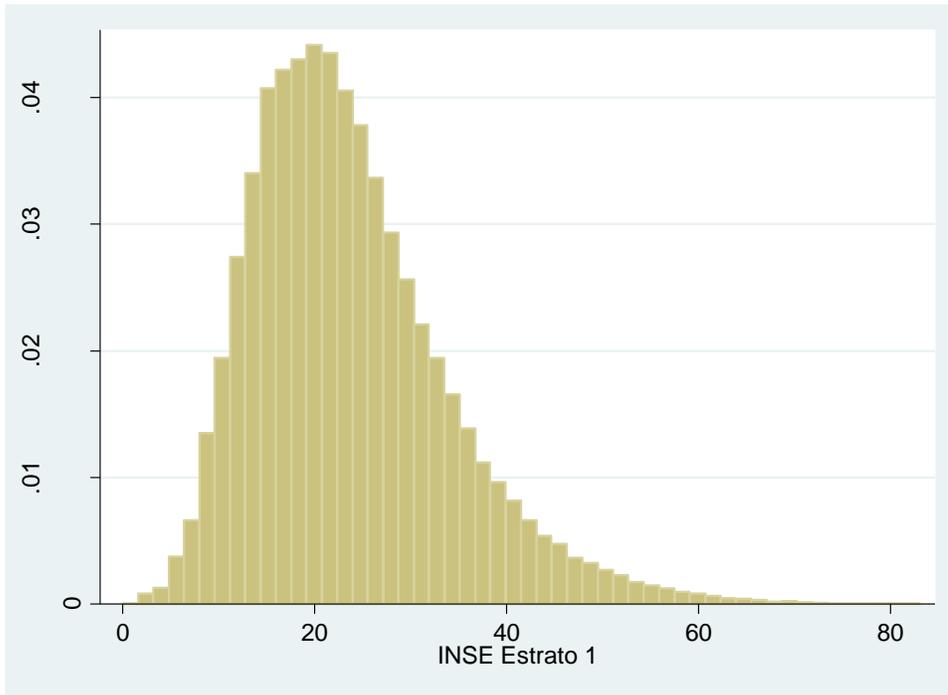
Fuente: Elaboración propia.

e)



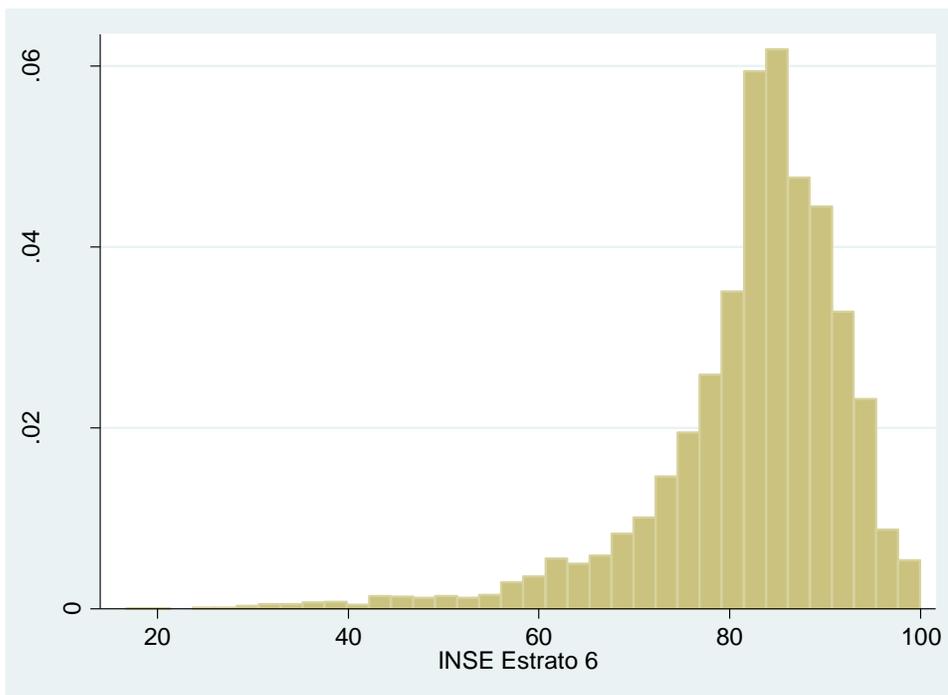
Fuente: Elaboración propia.

f)



Fuente: Elaboración propia.

g)



Fuente: Elaboración propia.