

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS

**APORTES EDUCATIVOS EN PERSPECTIVA: ANÁLISIS
COMPARATIVO MEDIANTE EL MÉTODO DE VALOR AGREGADO
EN COLEGIOS PARTICULARES SUBVENCIONADOS Y
PARTICULARES PAGADOS**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

AUTOR

ROBERTO ANDRÉS HERNÁNDEZ GARCÍA

PROFESOR GUÍA

BERNARDO PINCHEIRA

PROFESOR CO-REFERENTE

RODOLFO SALAZAR

VALPARAÍSO, 25 DE JULIO 2025



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): Memoria o trabajo de título; Tesis de Postgrado;

Título del trabajo: Aportes educativos en perspectiva: Análisis comparativo mediante el método del valor agregado en colegio particular subvencionados y particulares pagados.

Nombre del candidato(a): Roberto Andrés Hernández García

Carrera / Grado: Ingeniero Civil Industrial

Campus: Casa Central Valparaíso ; **Departamento:** Departamento de Industrias

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Bernardo Pincheira, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL

El trabajo **NO contiene información que amerite confidencialidad** y puede ser publicado de inmediato en repositorio con acceso abierto.

El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (embargo) por:

6 meses; 12 meses; 2 años; 3 años; 5 años; 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 25/07/25

; Firma:

Estudiante o Candidato(a):

Fecha: 25/07/25

; Firma:

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.



Tabla de Contenidos

1	Problema de Investigación	9
2	Objetivos	11
2.1	Objetivo General	11
2.2	Objetivos Específicos	11
3	Marco Teórico	12
3.1	El Sistema Educativo Chileno: Antecedentes y Evolución	12
3.2	Una de las Pruebas Estandarizadas en Chile: el SIMCE	13
3.3	Factores Socioeconómicos y Rendimiento Académico	14
3.4	Formación Docente y su Impacto en el Aprendizaje	15
3.5	Teorías Educativas Relacionadas	17
3.6	Valor Agregado en Educación	18
3.6.1	<i>Metodologías Para el Cálculo del Valor Agregado</i>	18
3.6.2	<i>Ventajas del Enfoque de Valor Agregado</i>	19
4	Descripción de Datos	20
4.1	Fuente de Datos	20
4.2	Base de Datos Originales	21
5	Consolidación de Datos	23
5.1	Selección de Variables de Base de Datos del SIMCE	23
5.1.1	<i>Base de Datos de los Alumnos</i>	23



5.1.2	<i>Base de Datos de los Apoderados</i>	24
5.1.3	<i>Base de Datos de los Profesores</i>	24
5.1.4	<i>Base de Datos de los Establecimientos</i>	24
5.2	Selección de Variables de Base de Datos del MINEDUC	25
5.3	Unión de Base de Datos	25
5.4	Limpieza de las Bases de Datos	26
5.4.1	<i>Bases de Datos del MINEDUC</i>	26
5.4.2	<i>Bases de Datos del SIMCE</i>	27
5.5	Creación de Variables	28
5.5.1	<i>Justificación de la Categorización de las Variables</i>	29
6	Metodología	32
6.1	Creación del Modelo	32
6.1.1	<i>Variables del Modelo</i>	33
6.1.1.1	Variables Dependientes.	33
6.1.1.2	Variables Independientes.	33
6.1.1.2.1	<i>Variables de Control del Establecimiento.</i>	33
6.1.1.2.2	<i>Variables de Control del Estudiante.</i>	34
6.1.1.2.3	<i>Variables de Control del Contexto Familiar y Socio-económico.</i>	34
6.1.2	<i>Fórmula del Modelo</i>	36
6.1.3	<i>Descripción de los Modelos</i>	37
6.1.4	<i>Creación de Grupos de Estudio</i>	39



7	Estadísticas Descriptivas	41
7.1	Variables Dependientes	41
7.2	Variables Independientes	42
7.2.1	<i>Tipo de Colegio</i>	42
7.2.2	<i>Nivel de Estudio Profesor</i>	43
7.2.3	<i>Puntajes SIMCE</i>	44
7.2.4	<i>Promedio General</i>	44
7.2.5	<i>Género del Estudiante</i>	45
7.2.6	<i>Vive con Figura Materna</i>	46
7.2.7	<i>Vive con Figura Paterna</i>	47
7.2.8	<i>Nivel de Estudio Figura Materna</i>	48
7.2.9	<i>Nivel de Estudio Figura Paterna</i>	49
7.2.10	<i>Nivel Socioeconómico</i>	50
7.2.11	<i>Número de Libros en el Hogar</i>	51
8	Resultados	52
8.1	Modelo de Lenguaje	53
8.1.1	<i>SIMCE 2010-2014</i>	53
8.1.2	<i>SIMCE 2011-2015</i>	56
8.1.3	<i>Aplicación de T-Test Pruebas de Lenguaje</i>	58
8.1.3.1	Grupo 1: Part. Subvencionados a Part. Subvencionados vs. Grupo 3: Part. Subvencionados a Part. Pagados.	58



8.1.3.2	Grupo 2: Part. Pagados a Part. Pagados vs. Grupo 4: Part. Pagados a Part. Subvencionados.	60
8.1.4	<i>Conclusiones Modelo de Lenguaje</i>	61
8.2	Modelo de Matemática	64
8.2.1	<i>SIMCE 2010-2014</i>	64
8.2.2	<i>SIMCE 2011-2015</i>	67
8.2.3	<i>Aplicación de T-Test Pruebas de Matemática</i>	69
8.2.3.1	Grupo 1: Part. Subvencionados a Part. Subvencionados vs. Grupo 3: Part. Subvencionados a Part. Pagados.	70
8.2.3.2	Grupo 2: Part. Pagados a Part. Pagados vs. Grupo 4: Part. Pagados a Part. Subvencionados.	71
8.2.4	<i>Conclusiones Modelo de Matemática</i>	71
8.3	Matriz de Correlaciones	74
9	Conclusiones	75
10	Limitaciones	77
11	Anexos	79
	Anexo A: Variables utilizadas desde la base de datos de alumnos provista por el SIMCE.	79
	Anexo B: Variables utilizadas de la bases de datos de apoderados provista por SIMCE.	81



Anexo C: Variables utilizadas de la bases de datos de los profesores provista por SIMCE.	84
Anexo D: Variables utilizadas de la bases de datos de los establecimientos educacionales provista por SIMCE.	85
Anexo E: Variables utilizadas de la bases de datos del MINEDUC.	86
Anexo F: Variables creadas para el modelo econométrico	87
Anexo G: Resultado prueba de heterocedasticidad.	89
Anexo H: Matriz correlación representada por números, prueba Lenguaje año 2010-2014.	90
Anexo I: Matriz correlación representada por números, prueba Matemática año 2010-2014.	91
Anexo J: Matriz correlación representada por números, prueba Lenguaje año 2011-2015.	92
Anexo K: Matriz correlación representada por números, prueba Matemática año 2011-2015.	93



Lista de Tablas

1	Información sobre la base de datos de los alumnos.	21
2	Información sobre la base de datos de los apoderados.	21
3	Información sobre la base de datos de los profesores.	22
4	Información sobre la base de datos de los establecimientos educacionales.	22
5	Información sobre la base de datos del MINEDUC.	22
6	Estadísticas descriptivas de los puntajes de Lectura y Matemática (2014 y 2015).	41
7	Estadísticas descriptivas de los puntajes de Lectura y Matemática (2010 y 2011).	44
8	Estadísticas descriptivas para la variable Promedio General.	44
9	Tabla resumen regresiones prueba de Lenguaje, año 2010-2014.	53
10	Tabla resumen regresiones prueba de Lenguaje, año 2011-2015.	56
11	Resultados T-test prueba de Lenguaje G1 vs G3, año 2010–2014.	59
12	Resultados T-test prueba de Lenguaje G1 vs G3, año 2011–2015.	59
13	Resultados T-test prueba de Lenguaje G2 vs G4, año 2010–2014.	60
14	Resultados T-test prueba de Lenguaje G2 vs G4, año 2011–2015.	60
15	Tabla resumen regresiones prueba de Matemática, año 2010-2014.	64
16	Tabla resumen regresiones prueba de Matemática, año 2011-2015.	67
17	Resultados T-test prueba de Matemática G1 vs G3, año 2010-2014.	70
18	Resultados T-test prueba de Matemática G1 vs G3, año 2011-2015.	70
19	Resultados T-test prueba de Matemática G2 vs G4, año 2010-2014.	71
20	Resultados T-test prueba de Matemática G2 vs G4, año 2011-2015.	71
21	Variables utilizadas desde la base de datos de alumnos provista por el SIMCE.	79



22	Continuación de Tabla 21	80
23	Variables utilizadas de la bases de datos de apoderados provista por SIMCE.	81
24	Continuación de Tabla 23	82
25	Continuación Tabla 24	83
26	Variables utilizadas de la bases de datos de los profesores provista por SIMCE. . .	84
27	Variables utilizadas de la bases de datos de los establecimientos educacionales pro- vista por SIMCE.	85
28	Variables utilizadas de la bases de datos del MINEDUC.	86
29	Variables creadas para el modelo econométrico.	87
30	Continuación Tabla 29	88
31	Resultado prueba de heterocedasticidad mediante el test de Breusch-Pagan.	89
32	Matriz correlación representada por números, prueba Lenguaje año 2010-2014. . .	90
33	Matriz correlacion representada por números, pueba Matemática año 2010-2014. .	91
34	Matriz correlacion representada por números, pueba Lenguaje año 2011-2015 . . .	92
35	Matriz correlacion representada por números, pueba Matemática año 2011-2015. .	93



Lista de Figuras

1	Frecuencia por Tipo de Colegio y Año Escolar.	42
2	Frecuencia por categoría para la variable Nivel de Estudio Profesor.	43
3	Frecuencia por categoría para la variable Género del Estudiante.	45
4	Frecuencia por categoría para la variable Vive con Figura Materna.	46
5	Frecuencia por categoría para la variable Vive con Figura Paterna.	47
6	Frecuencia por categoría para la variable Nivel de Estudio Figura Materna.	48
7	Frecuencia por categoría para la variable Nivel de Estudio Figura Paterna.	49
8	Frecuencia por categoría para la variable NSE.	50
9	Frecuencia por categoría para la variable Nivel de Número de Libros en el Hogar.	51

1. Problema de Investigación

En el contexto de las políticas educativas chilenas, uno de los principales desafíos ha sido reducir las brechas de calidad entre los distintos tipos de establecimientos escolares. Sin embargo, aún persisten dudas respecto a la efectividad real de estas medidas, especialmente cuando se analizan las diferencias de rendimiento académico observadas entre colegios municipales, particulares subvencionados y particulares pagados.

Una herramienta clave para abordar esta problemática es el concepto de valor agregado educativo, entendido como el aporte efectivo que un colegio genera en el aprendizaje de sus estudiantes, controlando sus condiciones iniciales. Este enfoque permite ir más allá de los puntajes absolutos en pruebas estandarizadas, al centrarse en el progreso individual del alumno a lo largo del tiempo, y resulta fundamental para distinguir si las diferencias de rendimiento responden realmente a la calidad del establecimiento o a factores externos.

Aunque estas diferencias han sido ampliamente documentadas, no siempre queda claro si reflejan la calidad de la enseñanza impartida o si, por el contrario, responden a variables como el nivel socioeconómico, el capital cultural del hogar o el contexto previo del estudiante. La literatura ha demostrado que el contexto socioeconómico es uno de los principales determinantes de las desigualdades educativas en Chile, al influir directamente en el tipo de colegio al que puede acceder el estudiante y, con ello, en las oportunidades de aprendizaje disponibles (Bellei, 2013; Gutiérrez y Carrasco, 2021).

Asimismo, se ha identificado el “factor colegio” como un elemento significativo en el rendimiento académico. Investigaciones muestran que estudiantes de escuelas públicas que acceden mediante becas a colegios particulares pagados tienden a obtener mejores resultados que aquellos

que permanecen en el sistema público, incluso compartiendo condiciones socioeconómicas similares (Anand et al., 2009). Esto sugiere que el tipo de establecimiento podría tener un efecto propio sobre los resultados de aprendizaje.

No obstante, otros estudios enfatizan el rol de variables adicionales como el desarrollo del pensamiento crítico (Quaresma, 2017) o el nivel educativo de los padres (Torche, 2005), indicando que el entorno familiar y cultural también desempeña un papel clave en el rendimiento escolar, más allá de la estructura del sistema educativo.

En este contexto, esta investigación busca estimar y comparar el valor agregado entregado por los colegios particulares pagados y particulares subvencionados en Chile, utilizando como base los puntajes SIMCE en Lenguaje y Matemática de 4° y 8° básico. A través de la construcción de un modelo econométrico que incorpora variables individuales, familiares y escolares, se pretende identificar si existen diferencias significativas en el aporte educativo de estos tipos de establecimientos, una vez controlado el rendimiento inicial de los estudiantes.

Con ello, se espera responder a las siguientes preguntas:

- ¿Cuál es el valor agregado en términos de rendimiento académico de los estudiantes en colegios particulares pagados, comparado con los particulares subvencionados, en pruebas como el SIMCE?
- ¿Cómo influyen los factores socioeconómicos en los puntajes obtenidos en estas pruebas?
- ¿Qué variables contribuyen de manera significativa a las diferencias en el desempeño académico de los estudiantes según el tipo de colegio?

2. Objetivos

2.1 Objetivo General

Evaluar el valor agregado en el rendimiento académico de los estudiantes de enseñanza básica en Chile, comparando la evolución de sus puntajes SIMCE en Lectura y Matemática entre 4° y 8° básico, durante los años 2010 y 2011, en colegios particulares pagados y particulares subvencionados, con el fin de identificar diferencias significativas en el aporte educativo de estos tipos de establecimientos.

2.2 Objetivos Específicos

- Analizar el impacto de los factores socioeconómicos del entorno familiar en el rendimiento académico de los estudiantes, evaluando su incidencia en los puntajes SIMCE de Lenguaje y Matemática entre 4° y 8° básico.
- Evaluar la influencia del nivel educativo de los padres y de los docentes en el progreso académico de los estudiantes, identificando su efecto en el valor agregado observado en ambos tipos de establecimientos.
- Comparar el valor agregado en los aprendizajes entre colegios particulares pagados y particulares subvencionados, utilizando un modelo econométrico que controle por rendimiento previo, características individuales, familiares y escolares.

3. Marco Teórico

3.1 El Sistema Educativo Chileno: Antecedentes y Evolución

Desde la promulgación de la Ley Orgánica Constitucional de Enseñanza (LOCE) en 1990 (Ley N°18956, 1990), Chile adoptó un modelo educativo basado en la coexistencia de distintos tipos de establecimientos:

1. Colegios Municipales (o públicos): Totalmente financiados por el Estado y administrados por los municipios.
2. Colegios Particulares Subvencionados: Reciben financiamiento estatal, pero pueden complementar con un copago de las familias.
3. Colegios Particulares Pagados: Financiados exclusivamente por los apoderados, suelen ser accesibles solo para los sectores de mayores ingresos.
4. Escuelas de administración delegada y otras categorías especiales.

Este sistema fue diseñado bajo la premisa de que la competencia entre establecimientos y la libertad de elección de los padres contribuirían a mejorar la calidad de la educación. Sin embargo, investigaciones han demostrado que dicha estructura no solo no logró reducir las desigualdades existentes, sino que las profundizó, generando un sistema altamente segregado y fuertemente condicionado por el nivel socioeconómico de las familias (Elacqua, 2012).

Con el paso del tiempo, se implementaron políticas para mitigar estas brechas. Una de ellas fue la Ley de Subvención Escolar Preferencial (SEP) (Ley N°20.248, 2008), que introdujo recursos adicionales para establecimientos que atienden a estudiantes en situación de vulnerabilidad.

Posteriormente, la Ley de Inclusión Escolar (Ley N°20845, 2015) planteó reformas más estructurales. Esta normativa eliminó el copago en los colegios subvencionados y prohibió la selección arbitraria en el proceso de admisión, con el objetivo de garantizar igualdad de oportunidades en el acceso a la educación.

No obstante, diversos estudios señalan que, en ausencia de mecanismos de regulación robustos, la competencia escolar y la libertad de elección siguen favoreciendo a los sectores más acomodados, consolidando una estructura educacional profundamente segmentada según el nivel socioeconómico (Elacqua, 2012; Gutiérrez y Carrasco, 2021; Bellei, 2013).

3.2 Una de las Pruebas Estandarizadas en Chile: el SIMCE

Comprender cómo se evalúa el aprendizaje en el sistema escolar chileno es clave para analizar las dinámicas educativas del país. En este contexto, destaca el Sistema de Medición de la Calidad de la Educación (SIMCE), implementado desde 1988 como una prueba estandarizada nacional (Ministerio de Educación, s.f.).

El SIMCE evalúa aprendizajes en áreas como Lenguaje, Matemáticas, Ciencias Naturales e Historia y Geografía, y permite comparar el desempeño estudiantil entre tipos de establecimientos y regiones, constituyéndose en una herramienta fundamental para el diseño e implementación de políticas educativas.

Sin embargo, el SIMCE no está exento de críticas. Una de las principales limitaciones es la falta de control por variables contextuales, especialmente aquellas de carácter socioeconómico. Esta omisión puede derivar en una sobrevaloración de los resultados obtenidos por estudiantes de contextos más favorecidos, afectando la interpretación del verdadero aporte de los establecimientos al aprendizaje de los alumnos (Thieme et al., 2016).

Aunque el SIMCE es un instrumento valioso, su utilidad para reducir las brechas de equidad es limitada. Por ello, se ha propuesto complementar su uso con enfoques más integrales, como los modelos de valor agregado, que permiten una evaluación más justa del impacto real de las escuelas sobre sus estudiantes. De lo contrario, un uso inadecuado de los resultados del SIMCE podría reforzar desigualdades preexistentes, penalizando injustamente a instituciones que atienden a poblaciones más vulnerables (Thieme et al., 2016).

3.3 Factores Socioeconómicos y Rendimiento Académico

Una vez comprendido cómo se mide el rendimiento académico, surge la pregunta respecto a los factores que influyen en dicho desempeño. El nivel socioeconómico (NSE) es ampliamente reconocido como uno de los determinantes más influyentes. En Chile, la fuerte correlación entre NSE y logros académicos ha sido ampliamente documentada, reflejando desigualdades estructurales y una marcada segregación escolar.

Los estudiantes provenientes de hogares con mayores ingresos tienden a obtener mejores resultados debido a múltiples factores:

- Las familias con mayor poder adquisitivo pueden invertir en recursos educativos complementarios como academias, tutores, materiales didácticos y actividades extracurriculares (Torche, 2005).
- Padres con mayor nivel educativo tienden a crear ambientes que estimulan el desarrollo cognitivo y emocional. Se ha encontrado una correlación positiva entre el nivel educativo de los padres y los puntajes en pruebas estandarizadas (Anand et al., 2009).

- Los colegios ubicados en contextos vulnerables suelen enfrentar desafíos estructurales, como menor financiamiento, alta rotación docente e infraestructura deficiente.

El NSE no solo determina el acceso a recursos, sino también la calidad de las experiencias escolares. Sus impactos incluyen:

- En colegios particulares pagados, las expectativas académicas suelen ser más elevadas, lo que se traduce en un mejor desempeño estudiantil (Anand et al., 2009).
- En contextos vulnerables, los estudiantes enfrentan desmotivación y bajas expectativas por parte de sus docentes y familias (Anand et al., 2009).
- La composición socioeconómica de los establecimientos influye en el rendimiento general, ya que estudiantes de NSE alto se benefician del entorno académico que crean sus pares (Torche, 2005).
- A largo plazo, los egresados de colegios particulares pagados tienen más probabilidades de acceder a universidades de élite y mejores oportunidades laborales (Torche, 2005).

3.4 Formación Docente y su Impacto en el Aprendizaje

La calidad docente es otro de los factores ampliamente discutidos en la literatura educativa y es considerada clave para mejorar los aprendizajes de los estudiantes (Rivkin et al., 2005; Hanushek y Rivkin, 2006). En Chile, las diferencias en la formación inicial y el desarrollo profesional de los profesores entre tipos de colegios reflejan y refuerzan las desigualdades del sistema educativo.

Con la Ley 20.903 (Ley N°20.903, 2016), se buscó mejorar la formación docente mediante:

- Establecimiento de requisitos de ingreso más exigentes a las carreras de pedagogía.
- Promoción de estándares nacionales para fortalecer la formación inicial.
- Creación del Sistema de Desarrollo Profesional Docente, con evaluaciones periódicas y una carrera profesional estructurada.

Pese a estos esfuerzos, la implementación ha sido desigual y los efectos sobre la calidad del sistema educativo siguen siendo mixtos.

Diversos estudios han evidenciado el impacto positivo de contar con docentes de alta calidad:

- Un docente eficaz puede mejorar el rendimiento académico de sus estudiantes hasta en un 20 % respecto a un docente promedio (Rivkin et al., 2005).
- La presencia de un profesor altamente efectivo puede compensar un año completo de aprendizaje adicional en comparación con uno menos eficaz (Hanushek y Rivkin, 2006).

Si bien se han introducido reformas significativas, aún persisten brechas en el acceso y la calidad de la formación docente, especialmente entre los diferentes tipos de establecimientos.

3.5 Teorías Educativas Relacionadas

Las diferencias en el rendimiento académico también pueden analizarse desde distintas perspectivas teóricas. Una de ellas es la teoría del capital humano, propuesta por Becker (Becker, 1964), que plantea que la educación incrementa las habilidades productivas de los individuos, generando mayores retornos económicos a futuro.

- Bajo esta lógica, la educación es una inversión que se traduce en mejores ingresos y condiciones de vida.
- Los países que invierten más en educación tienden a tener mayor crecimiento económico sostenido.

No obstante, esta teoría ha sido criticada por no considerar suficientemente las condiciones estructurales de origen, como el nivel socioeconómico, que pueden afectar los beneficios de la educación. Además, se centra principalmente en los aspectos económicos, dejando de lado dimensiones sociales, culturales y emocionales.

Otra perspectiva relevante es la Teoría Ecológica del Desarrollo Humano de Bronfenbrenner (Bronfenbrenner, 1979), que propone que el aprendizaje está influido por múltiples niveles de interacción: desde el entorno familiar inmediato (microsistema) hasta las políticas públicas y la cultura (macrosistema). En el contexto chileno, esta teoría resulta particularmente útil para entender cómo factores externos al aula afectan el rendimiento escolar.

Estas teorías permiten enriquecer el análisis del sistema educativo y orientan el diseño de políticas que promuevan tanto la equidad como la calidad educativa.

3.6 Valor Agregado en Educación

El concepto de valor agregado ha cobrado creciente importancia en el análisis de la calidad educativa, ya que busca aislar el impacto real que una escuela tiene sobre el aprendizaje de sus estudiantes, controlando por factores externos como el NSE, las características del hogar o el rendimiento inicial (Thieme et al., 2016; Anand et al., 2009; Gibbons y Silva, 2011).

Se define como la contribución neta de un establecimiento al progreso académico de sus estudiantes, y se estima mediante modelos estadísticos que consideran:

1. Rendimiento inicial: Situación de partida del estudiante.
2. Contexto socioeconómico: Ingreso del hogar, escolaridad de los padres, entorno social.
3. Características del colegio: Recursos, infraestructura, calidad docente.

Este enfoque permite responder preguntas clave como:

- ¿Cuánto mejora un estudiante debido a la calidad del colegio al que asiste?
- ¿Qué instituciones logran reducir las brechas entre alumnos de diferentes niveles socioeconómicos?

3.6.1. Metodologías Para el Cálculo del Valor Agregado

Entre las técnicas más utilizadas se encuentran:

1. Modelos de regresión lineal: Permiten estimar la contribución específica de un colegio al controlar por variables externas.

2. Modelos multinivel: Consideran la estructura jerárquica de los datos, separando el efecto de los estudiantes del efecto escuela.
3. Análisis longitudinal: Evalúan el progreso de los mismos estudiantes a lo largo del tiempo.

3.6.2. *Ventajas del Enfoque de Valor Agregado*

1. Permite una evaluación más justa, al reconocer diferencias de contexto.
2. Sirve como herramienta de política educativa, al identificar prácticas efectivas.
3. Promueve la mejora continua, al centrarse en el progreso y no solo en los resultados absolutos.

Aunque este enfoque no está exento de desafíos técnicos, su aplicación puede ser fundamental para avanzar hacia un sistema más equitativo y centrado en la mejora real del aprendizaje de los estudiantes.

4. Descripción de Datos

Antes de presentar la metodología econométrica empleada para estimar el valor agregado en los aprendizajes escolares, resulta fundamental detallar la fuente de los datos utilizados en el presente estudio. Este capítulo tiene por objetivo describir el origen de la base de datos, así como caracterizar las variables disponibles originalmente. Proporcionar esta información permite contextualizar el análisis posterior y garantizar la transparencia en el tratamiento de los datos.

4.1 Fuente de Datos

La investigación emplea datos provenientes del Sistema de Medición de la Calidad de la Educación (SIMCE), una evaluación estandarizada de carácter nacional que mide el rendimiento académico de los estudiantes en Chile. Esta prueba se aplica anualmente en distintos niveles escolares (4° básico, 8° básico y 2° medio) y evalúa áreas fundamentales como Lenguaje y Matemática. En algunos niveles, también incluye asignaturas como Ciencias Naturales y Ciencias Sociales.

Además de los puntajes obtenidos por los estudiantes, el SIMCE recopila información contextual mediante cuestionarios dirigidos a alumnos, padres y apoderados, docentes y equipos directivos de los establecimientos. Estos instrumentos permiten recolectar datos detallados sobre características demográficas, socioeconómicas, familiares y escolares, fundamentales para un análisis integral del rendimiento académico.

Para este estudio se consideraron los datos correspondientes a los años 2010 y 2011 para 4° básico, y 2014 y 2015 para 8° básico. Esta selección permite realizar un seguimiento longitudinal de los mismos estudiantes, lo que facilita la estimación del valor agregado atribuible al establecimiento educativo.

Asimismo, se incorporaron datos del Ministerio de Educación de Chile (MINEDUC), específicamente de las bases públicas sobre rendimiento escolar individual. Estas fuentes complementan la información del SIMCE, proporcionando variables adicionales relevantes para el desarrollo del modelo econométrico.

4.2 Base de Datos Originales

Antes de detallar las transformaciones y selecciones realizadas sobre las bases de datos para la construcción del modelo econométrico, resulta fundamental presentar la estructura original de las fuentes empleadas. Esta contextualización permite comprender el volumen y la diversidad de información disponible, así como las decisiones posteriores relacionadas con su depuración.

Las bases de datos del SIMCE se organizan en cuatro bloques principales: información del estudiante, de sus padres o apoderados, del establecimiento educacional y del cuerpo docente. Cada uno de estos bloques contiene variables específicas que permiten caracterizar tanto el rendimiento académico como el entorno educativo y familiar del estudiante.

Tabla 1

Información sobre la base de datos de los alumnos.

	2010	2011
Número de variables	26	23
Número de observaciones	255.448	247.667

Tabla 2

Información sobre la base de datos de los apoderados.

	2010	2011
Número de variables	165	182
Número de observaciones	222.187	213.163

Tabla 3*Información sobre la base de datos de los profesores.*

	2010	2011
Número de variables	198	215
Número de observaciones	10.520	9.713

Tabla 4*Información sobre la base de datos de los establecimientos educacionales.*

	2010	2011
Número de variables	48	48
Número de observaciones	7.928	7.741

Adicionalmente, se incorporaron bases de datos del MINEDUC, las cuales permiten complementar la información del SIMCE con datos transversales sobre el rendimiento y características generales de los estudiantes. A continuación, se presenta un resumen de su contenido:

Tabla 5*Información sobre la base de datos del MINEDUC.*

	2010	2011	2014	2015
Número de variables	24	25	34	31
Número de observaciones	3.335.825	3.326.746	3.238.586	3.227.534

5. Consolidación de Datos

Tras la presentación de las bases de datos originales, esta sección tiene por objetivo describir el proceso de selección, limpieza y transformación de variables que permitió consolidar la información empleada en la estimación del modelo econométrico.

En particular, se expone el tratamiento aplicado a las bases de datos del SIMCE y del MINEDUC, junto con los criterios adoptados para la generación de nuevas variables a partir de los datos originales. Esta etapa de consolidación fue esencial para asegurar la consistencia y comparabilidad de la información.

5.1 Selección de Variables de Base de Datos del SIMCE

5.1.1. *Base de Datos de los Alumnos*

Las variables seleccionadas desde la base de datos de estudiantes corresponden a aquellas necesarias para identificar de manera única a cada individuo, medir su rendimiento académico y posibilitar la vinculación con otras bases de datos complementarias, como las de padres, docentes y establecimientos educacionales. Estas variables están disponibles tanto en las bases correspondientes a los años 2010 como 2011, es decir, cuando los estudiantes cursaban 4° básico.

Para el nivel de 8° básico, se requiere contar con exactamente las mismas columnas, ya que estas permiten depurar la base de datos y recuperar los puntajes obtenidos por los estudiantes en las pruebas SIMCE. Estas variables se pueden ver a mayor detalle en el **Anexo A**.

5.1.2. *Base de Datos de los Apoderados*

Las variables seleccionadas desde la base de datos de apoderados permiten caracterizar el entorno socioeconómico del estudiante y estimar el nivel educativo alcanzado por las figuras paterna y materna. Además, se incluye una variable que permite aproximar el grado de estimulación académica en el hogar, a través de la cantidad de libros disponibles.

Estas variables son recolectadas cuando los estudiantes cursan 4° básico, ya que el objetivo es analizar su impacto sobre los resultados académicos obtenidos posteriormente, en 8° básico. Estas variables pueden ser vistas en el **Anexo B**.

5.1.3. *Base de Datos de los Profesores*

Las variables seleccionadas desde la base de datos de docentes permiten identificar el nivel educativo alcanzado por los profesores responsables del proceso de enseñanza de los estudiantes. Esta información resulta relevante para estimar el posible impacto que el grado académico del cuerpo docente tiene sobre el rendimiento de los alumnos en las pruebas SIMCE aplicadas en 8° básico.

Dado que se busca analizar el efecto de esta formación docente previa, estas variables son requeridas únicamente para el año en que los estudiantes cursan 4° básico. Para ver más detalles de estas variables seleccionadas, estas se encuentran en el **Anexo C**.

5.1.4. *Base de Datos de los Establecimientos*

Las variables seleccionadas desde la base de datos de establecimientos permiten identificar el tipo de colegio al que asistieron los estudiantes, tanto en 4° básico como en 8° básico. Esta

información es fundamental para filtrar y conservar únicamente aquellos establecimientos que se encuentran dentro del alcance de este estudio (específicamente, colegios particulares subvencionados y particulares pagados).

Asimismo, estas variables permiten determinar si los estudiantes permanecieron en el mismo tipo de establecimiento a lo largo del tiempo o si realizaron un cambio, lo cual puede tener implicancias relevantes en la interpretación de los resultados. Estas se encuentran en el **Anexo D**.

5.2 Selección de Variables de Base de Datos del MINEDUC

Las variables seleccionadas desde la base de datos del Ministerio de Educación tienen como objetivo principal estimar la relación entre el promedio general anual del estudiante (obtenido durante 4° básico) y su desempeño en las pruebas SIMCE rendidas en 8° básico. Esta información permite incorporar una medida complementaria del rendimiento académico histórico del alumno.

Adicionalmente, estas variables cumplen un rol relevante en el proceso de depuración y validación de los datos utilizados en el análisis. Dichas variables se encuentran disponibles en el **Anexo E**.

5.3 Unión de Base de Datos

Una vez identificadas las bases de datos relevantes y seleccionadas las variables pertinentes para el modelo, resulta fundamental detallar el procedimiento mediante el cual estas se integran entre sí.

El proceso de fusión se realiza de la siguiente manera:

- La base de datos de estudiantes se vincula con la base de datos de docentes utilizando las columnas `cod_curso` y `rbd`, generando así la primera base de datos provisoria.

- Posteriormente, esta base se une con la base de datos de apoderados, a través del identificador único del alumno `idalumno`, conformando una segunda base de datos provisoria.
- A continuación, se incorpora la información de los establecimientos educacionales mediante la columna `rbd`, enriqueciendo la base anterior con datos institucionales y manteniendo su condición de base provisoria.
- Finalmente, se integra la base consolidada con los datos provenientes del Ministerio de Educación, tanto para 4° como 8° básico, utilizando como clave la variable `mrun`. Este paso permite conformar la base de datos definitiva, que será utilizada en las etapas posteriores del análisis, según se describe en el siguiente capítulo.

5.4 Limpieza de las Bases de Datos

Para avanzar con el desarrollo de la investigación, fue necesario realizar un proceso de depuración de la base de datos. Este procedimiento tuvo como objetivo principal restringir el análisis únicamente a los establecimientos de interés, es decir, aquellos clasificados como particulares subvencionados y particulares pagados.

Asimismo, se exigió que todos los estudiantes incluidos en la muestra dispusieran de información completa y consistente para la construcción de las variables requeridas por el modelo econométrico. Esta condición fue esencial para asegurar la validez de los resultados obtenidos en la estimación.

5.4.1. *Bases de Datos del MINEDUC*

El proceso de limpieza comenzó con la depuración de la base de datos proporcionada por el Ministerio de Educación, debido a que algunos estudiantes podrían haber repetido el curso tanto

en 4° como en 8° básico. Esta situación resulta problemática para el modelo, ya que dificulta la identificación del promedio general real del alumno y su relación con los resultados obtenidos posteriormente.

Con el objetivo de simplificar el análisis, se optó por considerar únicamente a aquellos estudiantes que fueron promovidos en ambos niveles. Esta condición se identificó a través de la variable `sit_fin`, cuyo valor P (según la glosa de la base de datos) indica que el estudiante fue promovido, es decir, aprobó el año escolar correspondiente.

Adicionalmente, se aplicó un criterio de limpieza sobre la variable `prom_gral`, la cual representa el promedio general del alumno en 4° básico. Dado que esta variable será utilizada como insumo en el modelo econométrico, se exigió que todos los estudiantes incluidos en la muestra cuenten con un valor registrado en dicha columna. Aquellos casos sin información fueron excluidos del análisis.

5.4.2. Bases de Datos del SIMCE

Una vez completado el proceso de filtrado inicial sobre la base de datos del MINEDUC, se procedió con la limpieza de la base de datos provista por el SIMCE, utilizando los siguientes pasos y criterios:

- En primer lugar, se excluyeron aquellos estudiantes que no presentaban puntajes válidos en alguna de las dos pruebas (Lenguaje y Matemática) aplicadas tanto en 4° como en 8° básico. Esta verificación se realizó mediante las columnas `noptje_mate4b_alu`, `noptje_lect4b_alu`, `noptje_mate8b_alu` y `noptje_lect8b_alu`, en las cuales debía constar el valor “Con Valor”.

- De forma complementaria, y específicamente para el nivel de 8° básico, se impuso la condición de que las columnas `noptje_seg_mate8b` y `noptje_seg_lect8b` presentaran un valor igual a 0. Esto permitió asegurar que los puntajes fueran válidos y diferentes de cero para ambos niveles evaluados.
- Respecto a las encuestas respondidas por los docentes, el criterio de inclusión exigió que todas las variables seleccionadas para el modelo contaran con datos válidos, es decir, que las columnas `cprof_p05`, `cprof_p10_01`, `cprof_p10_02`, `cprof_p10_03`, `cprof_p10_04`, `cprof_p10_05`, `cprof_p03`, `cprof_p06_01`, `cprof_p06_02`, `cprof_p06_03`, `cprof_p06_04` y `cprof_p06_05` no estuvieran vacías. Esto garantizó la disponibilidad de información precisa sobre el nivel educativo de los docentes.
- Finalmente, se aplicó un filtro relacionado con el tipo de establecimiento educacional, considerando únicamente a aquellos clasificados como particulares subvencionados o particulares pagados, tal como se definió previamente en los objetivos de esta investigación. Para ello, se exigió que en la columna `cod_depe2`, tanto en 4° como en 8° básico, el valor fuera igual a 2 o 3, los cuales corresponden a dichos tipos de establecimientos según la glosa oficial de la base de datos del SIMCE.

5.5 Creación de Variables

Ahora ya con las variables originales seleccionadas y las bases de datos limpiadas, se procede con dar origen a las nuevas variables que perteneceran al modelo. La clasificación de estas se pueden ver el **Anexo F** y se explicará en el siguiente capítulo los criterios utilizados y el motivo por el cual fueron incluidas al modelo econométrico.

5.5.1. *Justificación de la Categorización de las Variables*

Las variables utilizadas en el modelo fueron recodificadas en categorías binarias o dicotómicas por criterios analíticos, metodológicos o por decisiones tomadas junto al profesor guía. A continuación, se detallan las razones específicas para cada una de ellas:

- `tipo_colegio`: Esta variable distingue entre los dos tipos de establecimientos considerados en el estudio.
 - 1 si el estudiante asiste a un colegio particular pagado.
 - 0 si asiste a un colegio particular subvencionado.
- `gen_alu`: Variable que identifica el género del estudiante. Está codificada de la siguiente manera:
 - 1 si el estudiante es de género masculino.
 - 0 si el estudiante es de género femenino.
- `nivel_estudio_profesor_i`: La categorización del nivel educativo del profesor fue definida arbitrariamente en conjunto con el profesor guía, agrupando los niveles de formación en tres categorías representativas.
 - 1 si el docente pertenece a la categoría correspondiente.
 - 0 en caso contrario.

- `vive_figura_materna`: Esta variable indica si el estudiante convive con una figura materna, de acuerdo con las respuestas entregadas en la encuesta.
 - 1 si convive con una figura materna.
 - 0 si no.

- `vive_figura_paterna`: Similar al caso anterior, esta variable refleja la convivencia con una figura paterna.
 - 1 si convive con una figura paterna.
 - 0 si no.

- `nivel_estudio_figura_materna_i`: Al igual que en el caso del nivel educativo del docente, esta variable fue dividida en tres categorías determinadas arbitrariamente con apoyo del profesor guía.
 - 1 si la figura materna pertenece a la categoría.
 - 0 si no.

- `nivel_estudio_figura_paterna_i`: Se aplicó el mismo criterio anterior para la figura paterna.
 - 1 si la figura paterna pertenece a la categoría.
 - 0 si no.

- `nse_i`: La categorización del nivel socioeconómico fue determinada en función de rangos salariales aproximados al contexto nacional de la época. La primera categoría abarca ingresos cercanos al sueldo mínimo (CymaSuite, s.f.), la segunda aproximadamente el doble del mínimo, y la tercera agrupa el resto de los casos.
 - 1 si el hogar se encuentra dentro de la categoría correspondiente.
 - 0 en caso contrario.

- `numero_libros_i`: Esta variable fue categorizada luego de un análisis exploratorio de datos. Las divisiones se establecieron de manera que cada categoría contuviera un porcentaje equilibrado de observaciones.
 - 1 si el estudiante pertenece a la categoría.
 - 0 si no.

6. Metodología

En esta etapa de la investigación, se introduce el modelo econométrico que permitirá analizar los resultados obtenidos por los estudiantes en las pruebas SIMCE. Para ello, se utiliza un enfoque de valor agregado, el cual busca medir el efecto diferencial en el rendimiento académico de los alumnos que cursaron 4° básico en un determinado tipo de establecimiento y que, posteriormente, en 8° básico, se mantuvieron o cambiaron de tipo de colegio.

Este modelo tiene por objetivo identificar el impacto atribuible a las características del establecimiento educativo sobre el aprendizaje, controlando por las condiciones previas del estudiante. Asimismo, en este capítulo se detallan las variables incluidas en el modelo, junto con su justificación y codificación correspondiente.

6.1 Creación del Modelo

Tal como se ha señalado previamente, el modelo implementado corresponde a un enfoque de valor agregado, el cual permite estimar con mayor precisión las diferencias en los puntajes obtenidos por los estudiantes y evaluar si estas diferencias resultan estadísticamente significativas. Este modelo se fundamenta en los resultados obtenidos en las pruebas de Matemática y Lenguaje del SIMCE, los cuales se relacionan con las variables seleccionadas y descritas en el capítulo anterior. De esta manera, se busca aislar el efecto específico del establecimiento educativo sobre el rendimiento académico.

En total, se estimarán cuatro modelos, dos por cada cohorte de estudio: 2010–2014 y 2011–2015. Para cada cohorte se desarrollará un modelo que evalúe el desempeño en la prueba de Lenguaje y otro en la prueba de Matemática.

6.1.1. *Variables del Modelo*

6.1.1.1. Variables Dependientes. Las siguientes variables corresponden a las variables dependientes utilizadas en los modelos de valor agregado. Estas representan el desempeño académico de los estudiantes en las pruebas SIMCE aplicadas en 8° básico:

- **ptje_lect8b_alu:** Variable que indica el puntaje obtenido por el estudiante en la prueba de Lenguaje en 8° básico. Es utilizada como variable dependiente en el modelo centrado en el desempeño en esta asignatura.
- **ptje_mate8b_alu:** Variable que indica el puntaje obtenido por el estudiante en la prueba de Matemática en 8° básico. Es utilizada como variable dependiente en el modelo correspondiente al área de Matemática.

6.1.1.2. Variables Independientes. Las siguientes variables corresponden a las variables independientes consideradas en los modelos de valor agregado. Incluyen factores del ámbito socioeconómico del estudiante, del entorno familiar y educativo, así como características del establecimiento escolar y de la calidad docente. Todas estas variables fueron medidas en 4° básico, ya que el objetivo es analizar cómo influyen en los resultados obtenidos en las pruebas SIMCE de 8° básico.

6.1.1.2.1. *Variables de Control del Establecimiento.*

- **tipo_colegio:** Variable que indica el tipo de establecimiento al que asiste el estudiante. Se incluye para evaluar si existen diferencias significativas en el rendimiento atribuibles al tipo de colegio, más allá de otros factores del modelo.

- **nivel_estudio_profesor_2, nivel_estudio_profesor_3:** Indican el máximo nivel educativo alcanzado por el docente que imparte clases al estudiante. Se incluyen para analizar si existe una relación entre el nivel académico del profesor y el desempeño del alumno.

nivel_estudio_profesor_1 corresponde a la categoría base.

6.1.1.2.2. Variables de Control del Estudiante.

- **ptje_lect4b_alu:** Puntaje obtenido en la prueba SIMCE de Lenguaje en 4° básico. Permite establecer un punto de referencia del desempeño académico del estudiante antes de 8° básico.
- **ptje_mate4b_alu:** Variable similar a la anterior, correspondiente al puntaje en la prueba SIMCE de Matemática en 4° básico.
- **prom_gral:** Promedio general del año académico del estudiante en 4° básico. Al igual que los puntajes SIMCE de ese nivel, esta variable entrega una medida del rendimiento académico previo del alumno.
- **gen_alu:** Variable que identifica el género del estudiante. Toma el valor 0 si es de género femenino, el cual se considera como categoría base en los modelos estimados.

6.1.1.2.3. Variables de Control del Contexto Familiar y Socioeconómico.

- **vive_con_figura_materna:** Indica si el estudiante convive con su figura materna (madre o madrastra). Se incorpora al análisis para evaluar el posible efecto de esta convivencia en el rendimiento académico. Toma el valor 1 si el estudiante convive con la figura materna y 0 en caso contrario, siendo esta última la categoría base en los modelos estimados.
- **vive_con_figura_paterna:** Variable análoga a la anterior, pero referida a la figura paterna.

- **nivel_estudio_figura_materna_2, nivel_estudio_figura_materna_3:** Representan el máximo nivel educativo alcanzado por la figura materna. Se incluyen para examinar si el nivel de instrucción de la madre o madrastra influye en los logros académicos del estudiante. La variable `nivel_estudio_figura_materna_1` se consideró como la categoría base.
- **nivel_estudio_figura_paterna_2, nivel_estudio_figura_paterna_3:** Variables análogas a las anteriores, pero referidas a la figura paterna. De igual manera que la variable anterior `nivel_estudio_figura_paterna_1` se consideró como la categoría base.
- **nse_2, nse_3:** Representan el nivel socioeconómico del estudiante. Se incluyen para estudiar cómo influye la disponibilidad de recursos materiales en el desempeño académico, dado que estudios previos han demostrado una relación entre mayores ingresos y mejor rendimiento escolar. Para esta variable, `nse_1` se omitió ya que fue considerada como categoría base.
- **numero_libros_2, numero_libros_3:** Indican la cantidad de libros, revistas y materiales similares disponibles en el hogar. Esta variable se considera un proxy del estímulo al aprendizaje en el entorno familiar. `numero_libros_1` se omitió ya que fue considerada como categoría base.

6.1.2. *Fórmula del Modelo*

$$\begin{aligned} \text{ptje_SIMCE8b_alu} = & \alpha + \beta_1 \cdot \text{tipo_colegio} \\ & + \beta_2 \cdot \text{nivel_estudio_profesor_2} + \beta_3 \cdot \text{nivel_estudio_profesor_3} \\ & + \beta_4 \cdot \text{ptje_SIMCE4b_alu} + \beta_5 \cdot \text{prom_gral} + \beta_6 \cdot \text{gen_alu} \\ & + \beta_7 \cdot \text{vive_con_figura_materna} + \beta_8 \cdot \text{vive_con_figura_paterna} \\ & + \beta_9 \cdot \text{nivel_estudio_figura_materna_2} + \beta_{10} \cdot \text{nivel_estudio_figura_materna_3} \\ & + \beta_{11} \cdot \text{nivel_estudio_figura_paterna_2} + \beta_{12} \cdot \text{nivel_estudio_figura_paterna_3} \\ & + \beta_{13} \cdot \text{nse_2} + \beta_{14} \cdot \text{nse_3} \\ & + \beta_{15} \cdot \text{numero_libros_2} + \beta_{16} \cdot \text{numero_libros_3} \\ & + \epsilon \end{aligned} \tag{1}$$

A continuación, se describen los elementos que componen la especificación del modelo de regresión utilizado para estimar el valor agregado en el rendimiento académico de los estudiantes:

- **ptje_SIMCE8b_alu:** Representa la variable dependiente del modelo, correspondiente al puntaje obtenido por el estudiante en la prueba SIMCE de Lenguaje o Matemática en 8° básico, según el modelo considerado.
- α : Corresponde al intercepto del modelo, es decir, el valor esperado del puntaje SIMCE en ausencia de los efectos de las variables explicativas.
- **ptje_SIMCE4b_alu:** Variable independiente que representa el puntaje obtenido por el estudiante en la prueba SIMCE de Lenguaje o Matemática en 4° básico, de acuerdo con el área correspondiente. Permite controlar por el rendimiento académico previo del alumno.

- $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9, \beta_{10}, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{15}, \beta_{16}$: Coeficientes asociados a las variables explicativas del modelo. Cada coeficiente mide el efecto marginal de la variable correspondiente sobre el puntaje SIMCE en 8° básico, manteniendo constantes las demás variables.
- ϵ : Término de error aleatorio que captura los factores no observados que pueden influir en el desempeño del estudiante y que no están incluidos explícitamente en el modelo.

6.1.3. Descripción de los Modelos

Una vez definido el modelo final (que integra todas las dimensiones relevantes para explicar el rendimiento académico), se procedió a construir una serie de modelos previos con el propósito de analizar de forma aislada distintos grupos de variables. Esta estrategia secuencial permite identificar de manera gradual y diferenciada el impacto que tienen las variables institucionales, individuales y familiares sobre los puntajes obtenidos por los estudiantes.

Cada uno de estos modelos incorpora progresivamente bloques de variables explicativas, lo que facilita no solo la estimación del efecto específico de cada dimensión, sino también la observación de cómo varía los coeficientes estimados al integrar nuevas variables de análisis. Cabe destacar que, en todos los modelos considerados, se mantiene constante la inclusión de la variable correspondiente al tipo de colegio, dado que esta representa el eje central del presente estudio.

A continuación, se describen los cuatro modelos estimados:

- **1° Modelo: Variables del Establecimiento.** Este modelo considera únicamente variables asociadas al contexto institucional, específicamente `tipo_colegio`, `nivel_estudio_profesor_2` y `nivel_estudio_profesor_3`. Su propósito es

identificar el efecto aislado de las características del establecimiento educacional sobre el desempeño académico de los estudiantes.

- **2° Modelo: Variables del Establecimiento + Variables del Estudiante.** Además de las variables del modelo anterior, se incorporan variables que capturan las características individuales del alumno, tales como `ptje_SIMCE4b_alu`, `prom_gral` y `gen_alu`. Este modelo permite examinar en qué medida el rendimiento previo y las condiciones individuales explican el desempeño en octavo básico.

- **3° Modelo: Variables del Establecimiento + Variables del Contexto Familiar y Socioeconómico.** Partiendo nuevamente de las variables del establecimiento, este modelo incorpora factores asociados al entorno del estudiante, se incluye `vive_con_figura_materna`, `vive_con_figura_paterna`, `nivel_estudio_figura_materna_2`, `nivel_estudio_figura_materna_3`, `nivel_estudio_figura_paterna_2`, `nivel_estudio_figura_paterna_3`, `nse_2`, `nse_3`, `numero_libros_2` y `numero_libros_3`. Este modelo busca aislar el impacto de las condiciones estructurales y culturales del entorno familiar sobre el desempeño académico.

- **4° Modelo: Modelo Final.** Este modelo combina la totalidad de las variables consideradas en los modelos anteriores, integrando de forma conjunta los factores institucionales, individuales y familiares. Su objetivo es evaluar de manera comprehensiva la interacción entre estas dimensiones y determinar cuál es su peso relativo en la explicación del rendimiento académico de los estudiantes.

6.1.4. *Creación de Grupos de Estudio*

Considerando el objetivo de esta investigación, centrado en analizar el rendimiento académico de los estudiantes en función del tipo de establecimiento educacional al que pertenecen, fue necesario clasificar a los alumnos en grupos según su trayectoria escolar entre 4° y 8° básico. Esta categorización permite evaluar si el cambio o permanencia en un tipo específico de colegio influye en los resultados obtenidos en las pruebas SIMCE. Los grupos definidos son los siguientes:

- **Colegio Particular Subvencionado a Particular Subvencionado (G1):** Incluye a los estudiantes que cursaron tanto 4° como 8° básico en establecimientos particulares subvencionados. Representa el grupo que se mantuvo en este tipo de institución durante todo el período de análisis.
- **Colegio Particular Pagado a Particular Pagado (G2):** Corresponde a los estudiantes que permanecieron en colegios particulares pagados tanto en 4° como en 8° básico, sin registrar cambio de dependencia.
- **Colegio Particular Subvencionado a Particular Pagado (G3):** Este grupo contempla a los estudiantes que en 4° básico asistían a un colegio particular subvencionado y que, en algún momento posterior, realizaron un cambio hacia un establecimiento particular pagado.
- **Colegio Particular Pagado a Particular Subvencionado (G4):** Incluye a los estudiantes cuya trayectoria escolar comienza en un colegio particular pagado en 4° básico, y que posteriormente migraron a un establecimiento particular subvencionado.



Esta segmentación en grupos permite analizar de manera más clara e integral a los estudiantes que compartieron condiciones escolares similares en 4° básico, facilitando así la comparación de sus resultados en 8° básico. Además, esta agrupación resulta fundamental para aplicar posteriormente pruebas estadísticas que evalúan si las diferencias observadas en los puntajes de las pruebas SIMCE son estadísticamente significativas. De este modo, se fortalece la validez del análisis realizado mediante el modelo de valor agregado, permitiendo arribar a conclusiones más precisas.

7. Estadísticas Descriptivas

En este capítulo, se presentan las estadísticas descriptivas de las variables analizadas. Este análisis descriptivo proporciona una visión general de los datos, permitiendo identificar tendencias y distribuciones clave antes de aplicar el modelo econométrico.

7.1 Variables Dependientes

Tabla 6

Estadísticas descriptivas de los puntajes de Lectura y Matemática (2014 y 2015).

Prueba	Año	Promedio	Desv. Estándar	Mínimo	Máximo	Número Obs.
Lenguaje	2014	254,73	51,51	111,57	373,24	84.347
Matemática	2014	278,97	45,75	145,06	402,69	84.347
Lenguaje	2015	257,33	48	116,65	373,16	76.790
Matemática	2015	281,81	46,79	139,74	394,39	76.790

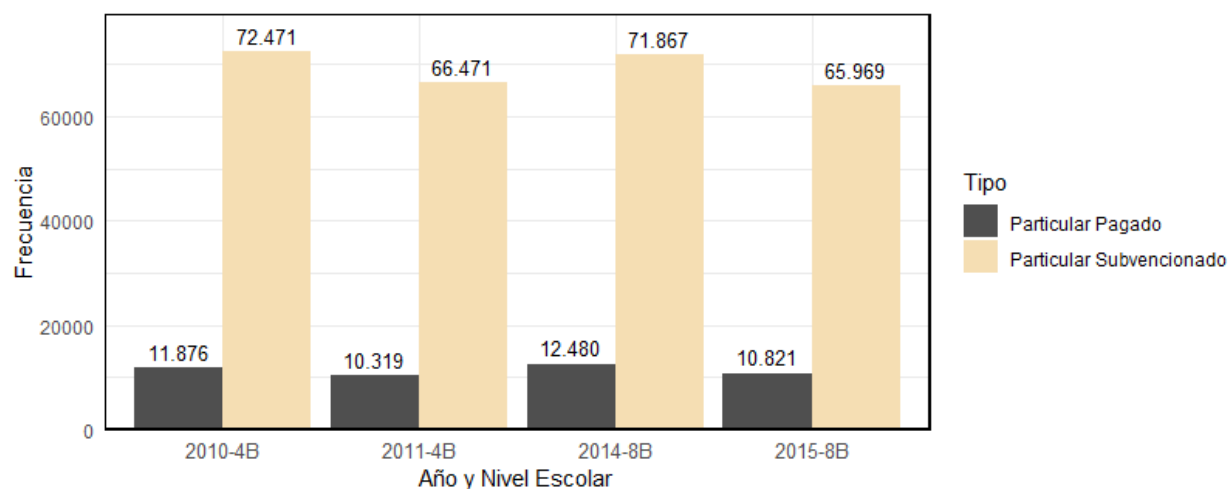
En términos generales, los estudiantes presentaron puntajes promedio superiores en la prueba de Matemática en comparación con la de Lenguaje, en ambos años analizados. Asimismo, se evidencia una leve mejora en los resultados de ambas asignaturas entre los años 2014 y 2015. La mayor desviación estándar observada en Lenguaje indica una mayor dispersión en los puntajes, lo que sugiere una mayor heterogeneidad en el rendimiento de los estudiantes en esta área. Finalmente, la presencia de puntajes extremos refleja una distribución amplia en los resultados obtenidos.

7.2 Variables Independientes

7.2.1. Tipo de Colegio

Figura 1

Frecuencia por Tipo de Colegio y Año Escolar:

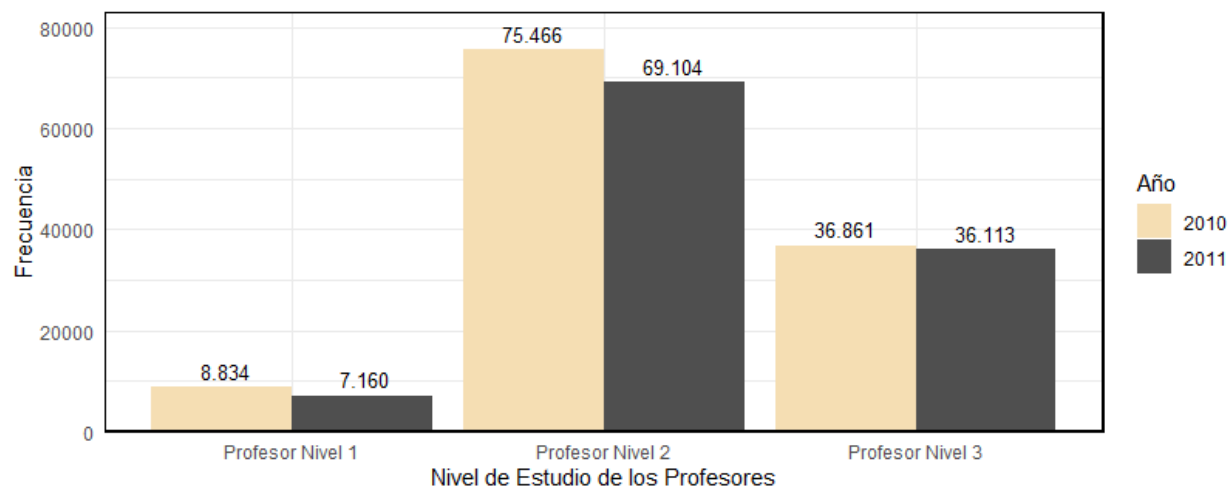


Se puede observar que, en todos los años analizados, la mayor proporción de estudiantes pertenece a establecimientos particulares subvencionados, con una frecuencia considerablemente superior en comparación con los colegios particulares pagados. En promedio, los estudiantes de colegios particulares pagados representan aproximadamente un 15 % del total, mientras que el 85 % restante corresponde a alumnos de colegios particulares subvencionados. Esta distribución se mantiene de forma consistente a lo largo de los años analizados.

7.2.2. Nivel de Estudio Profesor

Figura 2

Frecuencia por categoría para la variable Nivel de Estudio Profesor.



El grupo con formación correspondiente al Nivel 2 continúa siendo el más predominante, lo que indica que la mayoría de los docentes posee estudios universitarios y/o diplomados. La leve disminución observada en el Nivel 3 sugiere estabilidad en el número de profesores con formación de posgrado, manteniéndose porcentajes similares entre ambos años. En contraste, los docentes con formación de Nivel 1 (escuela normal y/o centros de formación técnica) representan una proporción menor del total en ambos periodos analizados.

7.2.3. Puntajes SIMCE

Tabla 7

Estadísticas descriptivas de los puntajes de Lectura y Matemática (2010 y 2011).

Prueba	Año	Promedio	Desv. Estándar	Mínimo	Máximo	Número Obs.
Lenguaje	2010	286,77	44,45	105,96	356,54	84.347
Matemática	2010	271,88	49,06	105,23	387,17	84.347
Lenguaje	2011	280,55	45,81	114,75	357,41	76.790
Matemática	2011	273,91	46,69	107,21	382,25	76.790

Durante los años 2010 y 2011, los estudiantes obtuvieron puntajes promedio más altos en la prueba de Lenguaje que en Matemática. Sin embargo, mientras que el rendimiento en Lenguaje experimentó una leve disminución, en Matemática se observó una ligera mejora. En ambos casos, la dispersión de los resultados fue considerable, siendo más pronunciada en Matemática, lo que sugiere una mayor heterogeneidad en el desempeño estudiantil.

7.2.4. Promedio General

Tabla 8

Estadísticas descriptivas para la variable Promedio General.

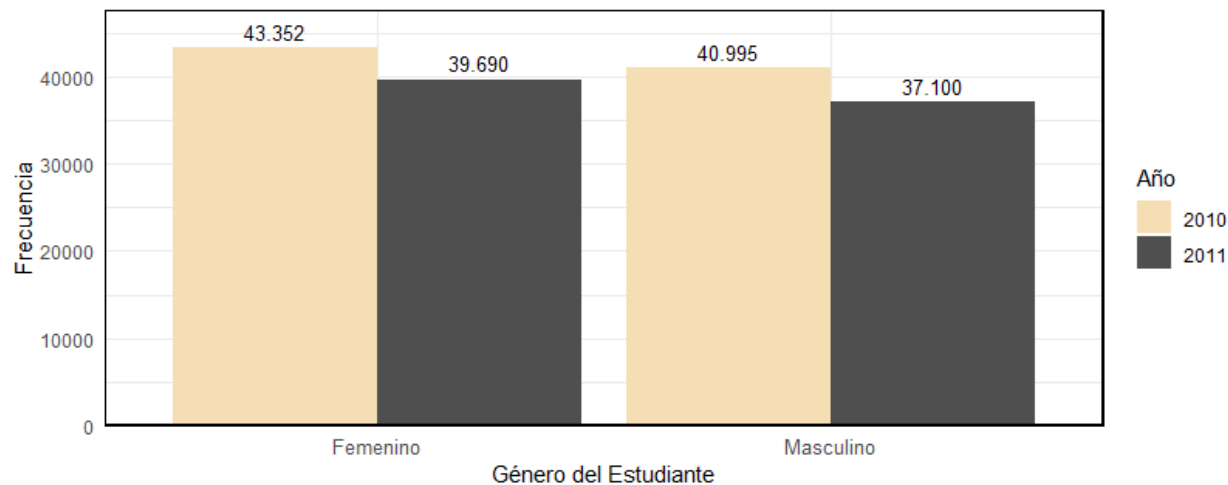
	Año	Promedio	Desv. Estándar	Mínimo	Máximo	Número Obs.
Promedio General	2010	6,01	0,47	4,4	7,0	84.347
	2011	5,99	0,47	4,42	7,0	76.790

El promedio general de los estudiantes en los años 2010 y 2011 muestran una tendencia muy similar en ambos periodos. Si bien el promedio y el puntaje mínimo son levemente superiores en 2010, la diferencia es marginal. La desviación estándar se mantiene constante en 0,47 en ambos años, lo que sugiere una dispersión de puntajes equivalente.

7.2.5. Género del Estudiante

Figura 3

Frecuencia por categoría para la variable Género del Estudiante.

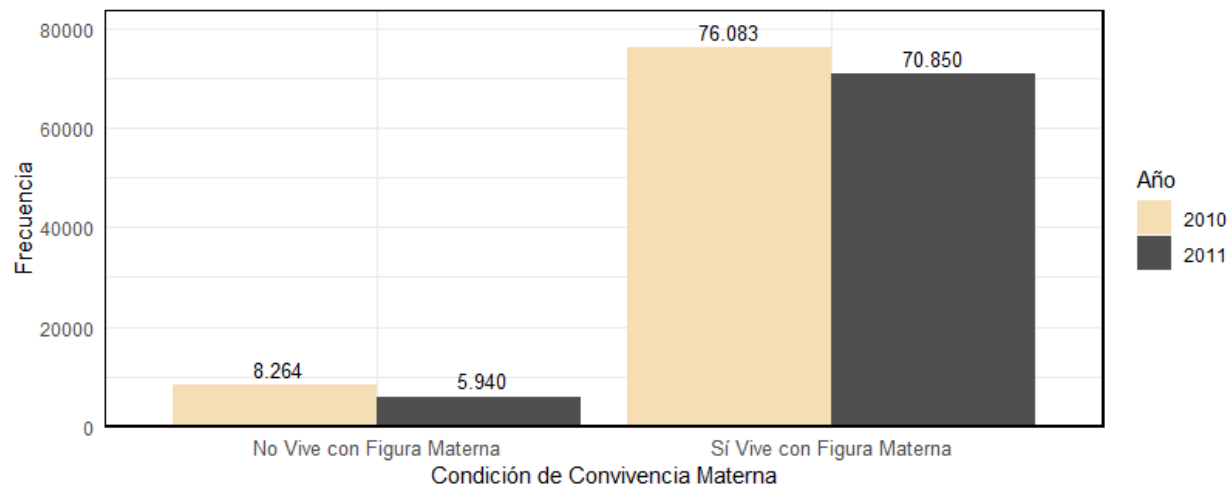


Durante los años 2010 y 2011, la distribución por género muestra una mayor proporción de estudiantes de sexo femenino en comparación con el masculino. Esta tendencia se mantiene constante en ambos años, con diferencias relativamente estables en las frecuencias.

7.2.6. *Vive con Figura Materna*

Figura 4

Frecuencia por categoría para la variable Vive con Figura Materna.

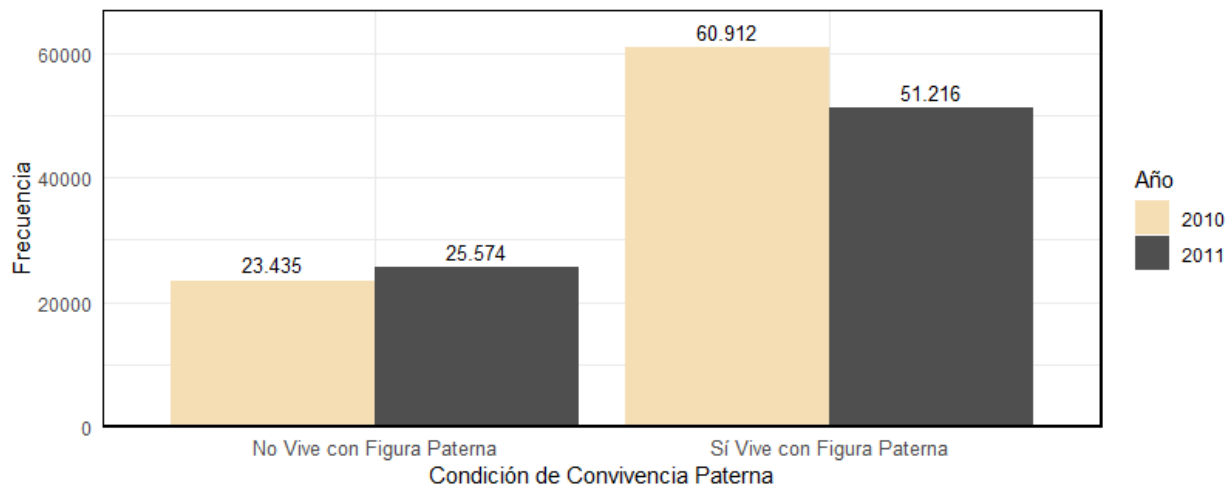


Durante los años 2010 y 2011, la mayoría de los estudiantes residía con una figura materna. Si bien se evidencia una leve disminución en ambas categorías, el patrón general se mantiene constante: más del 90 % de los alumnos convivía con una figura materna en el hogar.

7.2.7. *Vive con Figura Paterna*

Figura 5

Frecuencia por categoría para la variable Vive con Figura Paterna.

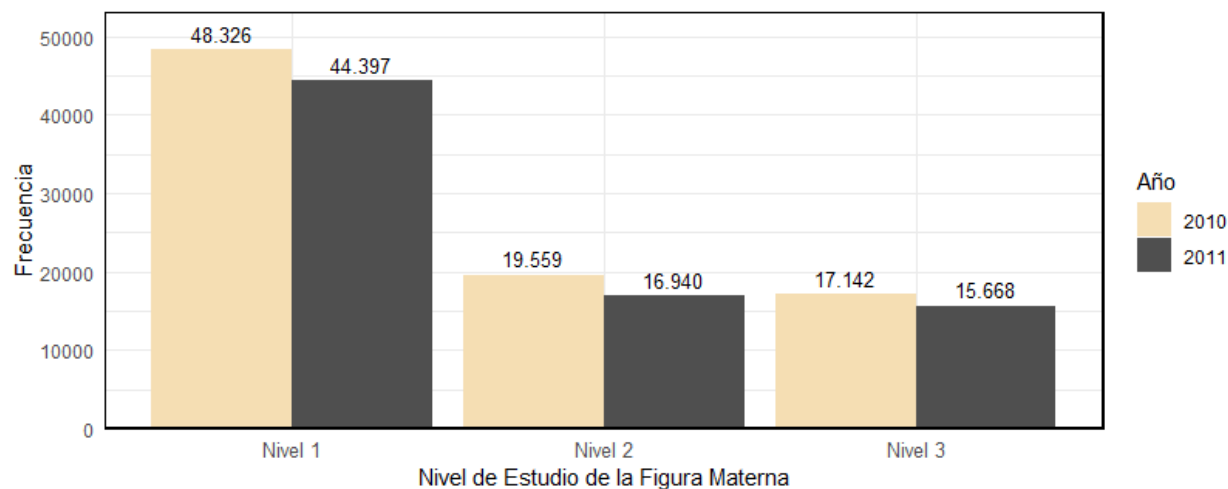


En ambos años, la mayoría de los estudiantes convivía con una figura paterna; sin embargo, en 2011 esta proporción se redujo levemente. Paralelamente, se observó un ligero aumento en el número de estudiantes que no residían con dicha figura. Al comparar esta situación con la de la figura materna, se evidencia que una mayor proporción de alumnos no convivía con una figura paterna, lo que indica una presencia relativamente menor de esta en el hogar. En términos generales, la figura paterna resulta más ausente que la materna.

7.2.8. Nivel de Estudio Figura Materna

Figura 6

Frecuencia por categoría para la variable Nivel de Estudio Figura Materna.

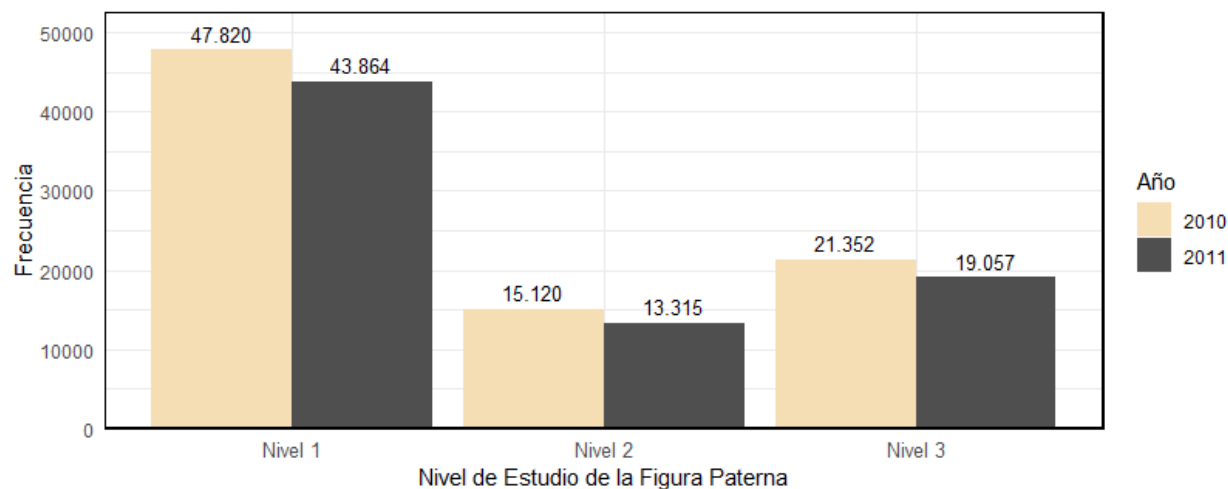


Entre 2010 y 2011 se observa una ligera disminución general en la proporción de figuras maternas en cada uno de los niveles educativos. La mayor concentración se presenta en el Nivel 1, lo que indica que una parte significativa de las madres posee un bajo nivel de escolaridad. Este patrón podría tener implicancias relevantes al momento de analizar el contexto familiar y su posible influencia en el rendimiento académico del estudiante.

7.2.9. Nivel de Estudio Figura Paterna

Figura 7

Frecuencia por categoría para la variable Nivel de Estudio Figura Paterna.

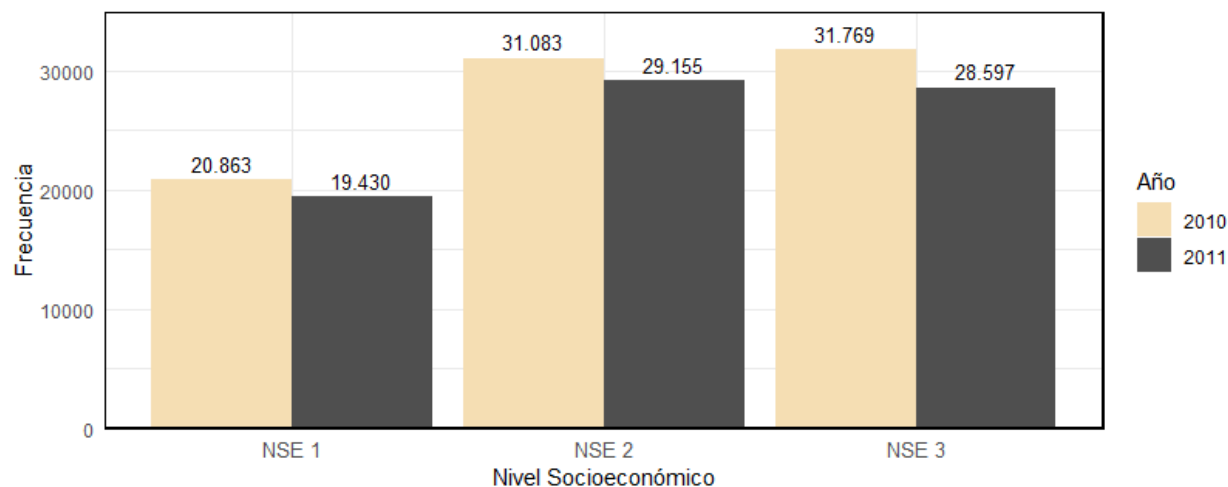


Entre los años 2010 y 2011 se evidencia una disminución en la proporción de figuras paternas en todos los niveles educativos. Al igual que en el caso de las figuras maternas, la mayor concentración se encuentra en el Nivel 1, lo que refleja un entorno familiar con bajo nivel de escolaridad. No obstante, a diferencia de lo observado en las madres, un porcentaje relevante de figuras paternas alcanza el Nivel 3, lo que sugiere una mayor presencia de estudios de posgrado en esta categoría.

7.2.10. Nivel Socioeconómico

Figura 8

Frecuencia por categoría para la variable NSE.

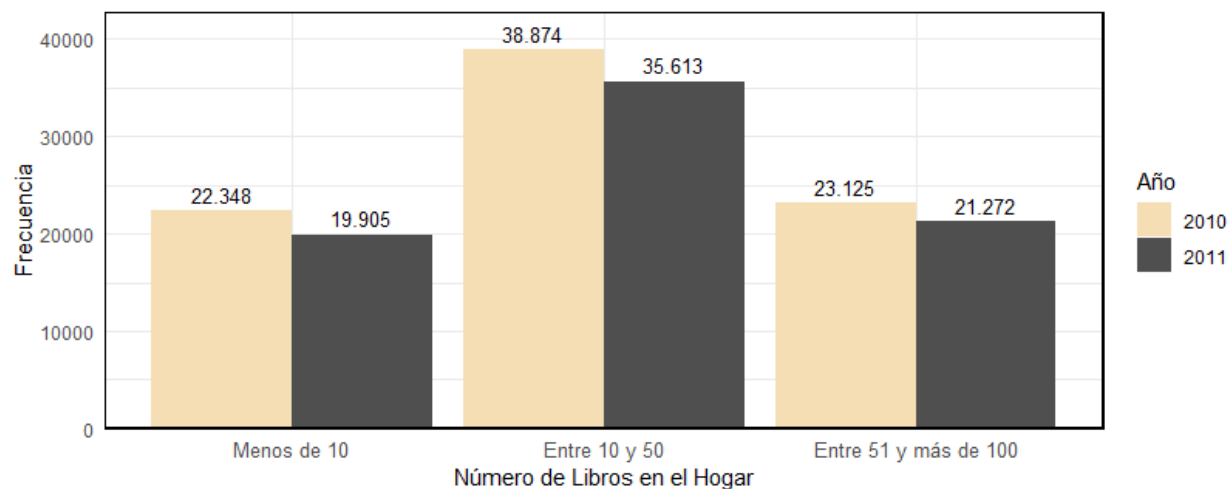


En ambos años, la mayoría de los estudiantes proviene de niveles socioeconómicos medios y altos. Este resultado es coherente con estudios previos, los cuales indican que las familias que acceden a establecimientos particulares subvencionados y particulares pagados (objeto de estudio en esta investigación) tienden a presentar mayores niveles de ingreso económico.

7.2.11. *Número de Libros en el Hogar*

Figura 9

Frecuencia por categoría para la variable Nivel de Número de Libros en el Hogar.



En ambos años, la mayoría de los estudiantes reportó tener entre 10 y 50 libros en el hogar, lo cual podría reflejar un nivel básico de acceso a material de lectura. Asimismo, la categoría con menos de 10 libros presenta la menor frecuencia, lo que sugiere que, en general, existía algún grado de estímulo hacia la lectura en los hogares de los estudiantes.

8. Resultados

En este capítulo se presentan los resultados obtenidos a partir de las regresiones realizadas y de la aplicación de los T-Test para la comparación entre los distintos grupos de estudio definidos previamente. El análisis se enfoca en la evaluación de la significancia estadística de las variables incluidas en el modelo, así como en la interpretación de los coeficientes estimados. Asimismo, se pone especial énfasis en la comparación de los resultados entre los grupos, con el objetivo de identificar posibles diferencias significativas en el rendimiento académico en función del tipo de establecimiento educacional y de los cambios ocurridos entre 4° y 8° básico.

8.1 Modelo de Lenguaje

8.1.1. SIMCE 2010-2014

Tabla 9
Tabla resumen regresiones prueba de Lenguaje, año 2010-2014.

	1° Modelo	2° Modelo	3° Modelo	4° Modelo
Intercepto	246,6726***	-42,7677***	230,6963***	-37,3704***
	(246,6726)***	(-42,7677)***	(230,6962)***	(-37,3704)***
Colegio Part. Pagado	29,4552***	10,6802***	9,1885***	3,5514***
	(29,4551)***	(10,6802)***	(9,1884)***	(3,5514)***
Nivel 2 de estudio profesor	3,4026***	1,3413**	2,8892***	1,2222**
	(3,4025)***	(1,3413)**	(2,8892)***	(1,222)***
Nivel 3 de estudio profesor	1,9907***	1,0747***	1,9045***	1,0839***
	(1,9906)***	(1,0747)***	(1,9045)***	(1,0839)***
Ptje. Lectura SIMCE 4B	-	0,5385***	-	0,5253***
	-	(0,5385)***	-	(0,5253)***
Promedio general	-	23,5865***	-	22,5778***
	-	(23,5865)***	-	(22,5778)***
Género Masculino	-	-4,0184***	-	-4,3226***
	-	(-4,0184)***	-	(-4,3226)***
Sí convive con la figura materna	-	-	2,1686***	0,2671
	-	-	(2,1686)***	(0,2671)
Sí convive con la figura paterna	-	-	0,5683	-0,5803
	-	-	(0,5682)	(-0,5803)
Nivel 2 de estudio figura materna	-	-	4,2276***	0,9756*
	-	-	(4,2275)***	(0,9756)*
Nivel 3 de estudio figura materna	-	-	11,1741***	4,5336***
	-	-	(11,1740)***	(4,5336)***
Nivel 2 de estudio figura paterna	-	-	4,2258***	1,4829***
	-	-	(4,2258)***	(1,4829)***
Nivel 3 de estudio figura paterna	-	-	11,2367***	5,0414***
	-	-	(11,2366)***	(5,0414)***
NSE nivel 2	-	-	4,8905***	0,9897**
	-	-	(4,8904)***	(0,9897)**
NSE nivel 3	-	-	8,5239***	1,4461**
	-	-	(8,5238)***	(1,4461)**
Nivel 2 de número de libros	-	-	4,7929***	2,0126***
	-	-	(4,7928)***	(2,0126)***
Nivel 3 de número de libros	-	-	10,5241***	4,9529***
	-	-	(10,5240)***	(4,9529)***
R^2 Ajustado	0,0407	0,3959	0,0835	0,4017
Nº Observaciones	84.347			

Notas. *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$. La primera fila de valores para cada una de las variables corresponde a la primera regresión, mientras que la segunda fila de valores (los que se encuentran entre paréntesis) corresponden a los resultados obtenidos mediante la técnica de errores robustos.

Una vez aplicados ambos métodos de estimación en los cuatro modelos propuestos, se observa que los niveles de significancia de las variables explicativas se mantienen estables, sin variaciones relevantes entre los resultados obtenidos con errores estándar tradicionales y aquellos ajustados con errores robustos. Al comparar el modelo final, que incorpora la totalidad de las variables, con los modelos parciales, se advierte una reducción general en la magnitud de los coeficientes. Esto sugiere que parte del efecto atribuido inicialmente a ciertas variables podría estar capturando influencias compartidas con otras variables incluidas en el modelo completo.

En el modelo final, los estudiantes que asisten a colegios particulares pagados presentan un puntaje promedio 3,6 puntos superior en la prueba de Lenguaje en comparación con aquellos que asisten a colegios particulares subvencionados, controlando por el resto de las variables. Este resultado indica un efecto positivo pero moderado del tipo de establecimiento.

Respecto al nivel académico del profesorado, tanto el segundo como el tercer nivel de estudios muestran coeficientes positivos y estadísticamente significativos. No obstante, no se evidencia una progresión clara del efecto conforme se incrementa el nivel educativo, lo cual sugiere que otros factores podrían estar modulando esta relación.

En cuanto al género del estudiante, los resultados revelan que los alumnos del género masculino obtienen, en promedio, 4,32 puntos menos que sus pares femeninas, manteniéndose constante el resto de las variables.

Las variables de convivencia con las figuras parentales (materna y paterna) no muestran efectos estadísticamente significativos en el modelo final. Por el contrario, el capital educativo familiar sí presenta una influencia positiva: tanto para la figura paterna como para la materna, el nivel 3 de escolaridad muestra efectos relevantes, siendo más fuerte el asociado al padre (5,04 puntos en promedio) frente al de la madre (4,53 puntos).



El nivel socioeconómico del hogar también tiene un impacto positivo en el rendimiento en Lenguaje. Sin embargo, los incrementos de puntaje asociados al segundo y tercer nivel del NSE son más bien acotados, con efectos de aproximadamente 1 punto.

Finalmente, el acceso a material de lectura en el hogar, medido por la cantidad de libros, se asocia de forma positiva y significativa al rendimiento académico. Los estudiantes pertenecientes al segundo y tercer nivel de esta variable obtienen, en promedio, 2 y 5 puntos más, respectivamente, que aquellos con menor acceso a libros.

8.1.2. SIMCE 2011-2015
Tabla 10
Tabla resumen regresiones prueba de Lenguaje, año 2011-2015.

	1° Modelo	2° Modelo	3° Modelo	4° Modelo
Intercepto	249,6934***	1,9284	236,7238***	6,2779***
	(249,6933)***	(1,9284)	(236,7238)***	(6,2779)***
Colegio Part. Pagado	23,872***	6,3769***	7,4109***	1,2293*
	(23,8719)***	(6,3769)***	(7,4109)***	(1,2293)*
Nivel 2 de estudio profesor	3,8856***	1,4247**	3,7002***	1,4385**
	(3,8856)***	(1,4247)**	(3,7002)***	(1,4385)**
Nivel 3 de estudio profesor	1,9847***	0,7569**	1,7814***	0,7061**
	(1,9847)***	(0,7569)**	(1,7814)***	(0,7061)**
Ptje. Lectura SIMCE 4B	-	0,4926***	-	0,4848***
	-	(0,4926)***	-	(0,4848)***
Promedio general	-	19,5425***	-	18,7793***
	-	(19,5425)***	-	(18,7793)***
Género Masculino	-	-4,9180***	-	-5,0256***
	-	(-4,9180)***	-	(-5,0255)***
Sí convive con la figura materna	-	-	2,4040***	-0,4150
	-	-	(2,4040)***	(-0,4150)
Sí convive con la figura paterna	-	-	-0,0665	-0,6141*
	-	-	(-0,0665)	(-0,6141)*
Nivel 2 de estudio figura materna	-	-	2,9844***	0,4632
	-	-	(2,9844)***	(0,4632)
Nivel 3 de estudio figura materna	-	-	8,9987***	2,8860***
	-	-	(8,9987)***	(2,8860)***
Nivel 2 de estudio figura paterna	-	-	3,3388***	2,8860**
	-	-	(3,3388)***	(1,1510)**
Nivel 3 de estudio figura paterna	-	-	8,5115***	3,5465***
	-	-	(8,5115)***	(3,5465)***
NSE nivel 2	-	-	4,0243***	1,1575**
	-	-	(4,0243)***	(1,1575)**
NSE nivel 3	-	-	6,7793***	1,2696**
	-	-	(6,7793)***	(1,2696)**
Nivel 2 de número de libros	-	-	3,1193***	0,4633
	-	-	(3,1192)***	(0,4633)
Nivel 3 de número de libros	-	-	9,3331***	3,9053***
	-	-	(9,3331)***	(3,9053)***
R^2 Ajustado	0,0300	0,3748	0,0627	0,3785
Nº Observaciones	76.790			

Notas. ***p<0,001; **p<0,01; *p<0,05. La primera fila de valores para cada una de las variables corresponde a la primera regresión, mientras que la segunda fila de valores (los que se encuentran entre paréntesis) corresponden a los resultados obtenidos mediante la técnica de errores robustos.

Tras estimar los cuatro modelos de regresión correspondientes, se observa que los niveles de significancia de algunas variables cambian levemente en el modelo final respecto de los modelos parciales. En particular, ciertas variables que eran estadísticamente significativas al estar aisladas pierden significancia al ser incorporadas junto con el resto de los factores. Además, se aprecia una disminución general en la magnitud de los coeficientes estimados en el modelo completo, lo que sugiere la existencia de relaciones compartidas entre variables que atenúan su efecto individual.

En cuanto al tipo de establecimiento, los estudiantes que asisten a colegios particulares pagados obtienen, en promedio, 1,22 puntos más en la prueba de Lenguaje que sus pares de colegios particulares subvencionados, manteniéndose constantes las demás variables. Aunque este efecto es menor que el observado en el periodo 2010–2014, sigue siendo estadísticamente significativo.

La formación académica del profesorado sigue mostrando una influencia positiva y significativa sobre el rendimiento de los estudiantes. Sin embargo, no se evidencia una relación claramente creciente entre el nivel de estudios del docente y el puntaje obtenido, indicando que un mayor nivel académico no se traduce necesariamente en un impacto proporcionalmente mayor en los resultados.

Respecto al género, los estudiantes del género masculino obtienen, en promedio, 5 puntos menos que las estudiantes mujeres.

En el ámbito familiar, la convivencia con la figura materna no presenta un efecto estadísticamente significativo en el modelo final, mientras que la convivencia con la figura paterna muestra una relación negativa, aunque de escasa magnitud (aproximadamente -0,61 puntos). En cuanto al nivel educativo de las figuras parentales, tanto para la figura materna como paterna, el nivel 3 de escolaridad se asocia a un mejor desempeño académico. El efecto más destacado corresponde a la figura paterna, con un aumento promedio de 3,55 puntos, seguido por la figura materna con un

incremento de 2,89 puntos.

El NSE del hogar mantiene un efecto positivo moderado sobre el rendimiento. Los estudiantes pertenecientes al nivel 2 presentan un aumento estimado de 1,16 puntos, y aquellos del nivel 3 un incremento de 1,27 puntos, en comparación con quienes se ubican en el nivel más bajo.

Por último, el capital cultural del hogar, medido a través del número de libros disponibles, muestra un impacto positivo únicamente en su nivel más alto. Específicamente, los estudiantes del nivel 3 de esta variable obtienen en promedio 3,91 puntos más en la prueba de Lenguaje, mientras que el nivel intermedio no presenta efectos significativos.

8.1.3. *Aplicación de T-Test Pruebas de Lenguaje*

Dado que uno de los objetivos centrales de esta investigación es evaluar si existen diferencias significativas en el rendimiento académico según el tipo de establecimiento educacional (tanto en aquellos estudiantes que se mantuvieron en el mismo tipo de colegio como en aquellos que realizaron un cambio), se aplicó el T-Test.

Esta herramienta permite contrastar si las diferencias de puntaje observadas entre los distintos grupos definidos previamente (G1, G2, G3 y G4) son estadísticamente significativas, es decir, si existen fundamentos suficientes para afirmar que el tipo de trayectoria escolar seguida por los estudiantes tiene un efecto relevante sobre los resultados obtenidos en la prueba SIMCE de Lenguaje y Matemática.

8.1.3.1. Grupo 1: Part. Subvencionados a Part. Subvencionados vs. Grupo 3: Part. Subvencionados a Part. Pagados. Para el periodo 2010-2014 y 2011-2015, se aplicó un T-test de diferencia de medias entre los grupos G1 (estudiantes que se mantuvieron en colegios particu-

lares subvencionados) y G3 (estudiantes que comenzaron en colegios particulares subvencionados y luego migraron a particulares pagados). Los resultados se presentan a continuación:

Tabla 11

Resultados T-test prueba de Lenguaje G1 vs G3, año 2010–2014.

	Valor
Estadístico t	-4,76
Valor p	2,083·e-06

Tabla 12

Resultados T-test prueba de Lenguaje G1 vs G3, año 2011–2015.

	Valor
Estadístico t	1,25
Valor p	0,211

El valor p mostrado en la **Tabla 11** es considerablemente menor al umbral convencional de significancia (0,05), lo que indica que la diferencia entre ambos grupos es estadísticamente significativa. Específicamente, los estudiantes del grupo G3 presentan, en promedio, un puntaje superior al del grupo G1 en la prueba de Lenguaje. Esto sugiere que el cambio desde un colegio particular subvencionado a uno particular pagado se asocia con un mejor desempeño académico, al menos en esta área evaluada.

Mientras que para los alumnos de los años 2010-2015 se tiene que no se rechaza la hipótesis nula, ya que el valor p obtenido (0,211) supera el umbral habitual de significancia. Esto implica que no se encuentra evidencia estadísticamente significativa que respalde una diferencia de puntaje en la prueba de Lenguaje entre los grupos G1 y G3.

De forma interesante, en este caso el promedio de puntajes fue levemente superior para los estudiantes del grupo G1, con una diferencia de 1,25 puntos a favor de quienes se mantuvieron en colegios particulares subvencionados. Sin embargo, esta diferencia no es estadísticamente concluyente, por lo que no puede atribuirse con certeza a un efecto del tipo de establecimiento educativo.

8.1.3.2. Grupo 2: Part. Pagados a Part. Pagados vs. Grupo 4: Part. Pagados a Part.

Subvencionados. Los resultados para estos grupos fueron los siguientes:

Tabla 13

Resultados T-test prueba de Lenguaje G2 vs G4, año 2010–2014.

	Valor
Estadístico t	3,54
Valor p	0,0004247

Tabla 14

Resultados T-test prueba de Lenguaje G2 vs G4, año 2011–2015.

	Valor
Estadístico t	0,85
Valor p	0,3975

Los resultados para el año 2010-2014 permiten concluir que existe una diferencia estadísticamente significativa entre ambos grupos. Específicamente, los estudiantes del grupo G2 obtuvieron, en promedio, 3,54 puntos más que los del grupo G4 en la prueba de Lenguaje. Esto sugiere que aquellos alumnos que permanecieron en colegios particulares pagados lograron mejores resultados académicos en comparación con quienes se cambiaron a un colegio particular subvencionado.

Mientras que para el año 2011-2015 de acuerdo con los resultados obtenidos, no se encuentra evidencia estadísticamente significativa que permita afirmar la existencia de una diferencia en los puntajes de la prueba de Lenguaje entre ambos grupos. Esto indica que, para este periodo, los estudiantes que permanecieron en colegios particulares pagados (G2) no obtuvieron un desempeño significativamente diferente al de aquellos que se trasladaron a colegios particulares subvencionados (G4).

8.1.4. Conclusiones Modelo de Lenguaje

En relación con los valores estimados de los coeficientes obtenidos en los modelos de regresión para la prueba de Lenguaje, se observa que la variable correspondiente al tipo de colegio no presenta un efecto particularmente relevante. Si bien en algunos modelos los estudiantes que asistieron a establecimientos particulares pagados muestran una leve mejora respecto a aquellos de colegios particulares subvencionados, esta diferencia no es lo suficientemente amplia como para considerarse concluyente de manera robusta. En promedio, las diferencias observadas fluctúan entre 1 y 3,6 puntos, lo cual representa una magnitud moderada, y su significancia estadística varía según el modelo considerado.

De manera similar, los resultados asociados al nivel académico del profesorado reflejan un efecto positivo al aumentar el nivel educativo del docente. No obstante, estos aumentos tienden a ser acotados y no siempre consistentes. En los modelos estimados, las mejoras en el rendimiento no superan los 1,5 puntos en promedio al pasar de un nivel de formación docente a otro más alto, lo que sugiere que este factor, si bien influye, no representa un determinante central del rendimiento en la prueba de Lenguaje.

Respecto a las características individuales del estudiante, el rendimiento previo (medido por el puntaje SIMCE de cuarto básico y el promedio general) destacan como los predictores más fuertes y estables del desempeño posterior en Lenguaje. Ambas variables presentan efectos positivos significativos y de alta magnitud, lo que reafirma la importancia del desempeño escolar acumulado como predictor del rendimiento futuro.

En cuanto al género, los resultados indican que los estudiantes de género masculino obtienen sistemáticamente un menor puntaje que sus pares femeninas, con diferencias que alcanzan

hasta 5 puntos. Esta brecha se mantiene constante en los distintos modelos, lo que sugiere una diferencia de rendimiento persistente entre ambos géneros en el área de Lenguaje.

En relación con las variables familiares, la convivencia con la figura materna no resultó significativa en ninguno de los modelos finales, lo que sugiere que su presencia no se asocia directamente a una mejora en el rendimiento. En contraste, la convivencia con la figura paterna presentó un efecto negativo en algunos modelos, aunque su magnitud fue menor a un punto, por lo que si bien estadísticamente significativa en ciertos casos, su relevancia práctica es limitada.

Por otra parte, el nivel educativo de las figuras parentales muestra una influencia más clara sobre los resultados. En particular, el nivel 3 de escolaridad tanto en la figura materna como paterna se asocia con incrementos positivos en el rendimiento de los estudiantes, con efectos que oscilan entre 2,8 y 5 puntos. Este patrón sugiere que el capital educativo del hogar desempeña un rol importante, especialmente cuando alcanza niveles altos de formación.

Asimismo, el nivel socioeconómico muestra una relación positiva con el rendimiento, aunque de menor magnitud. En los modelos donde esta variable fue significativa, el incremento al pasar del nivel bajo al alto fue inferior a 1,5 puntos, lo que indica una influencia moderada, pero consistente, del contexto económico en el desempeño en Lenguaje.

Un efecto más notorio se observa en la variable relacionada con el número de libros en el hogar. En particular, los estudiantes que pertenecen al nivel más alto de esta categoría, obtienen hasta 5 puntos más en la prueba lo que refleja una fuerte asociación entre el acceso a recursos lectores en el hogar y el rendimiento académico.

En complemento a los modelos de regresión, se realizaron comparaciones de medias mediante pruebas T-test entre grupos de estudiantes que cambiaron o permanecieron en distintos tipos de colegios a lo largo del tiempo. En términos generales, se identificaron diferencias estadística-

mente significativas a favor de quienes migraron desde colegios particulares subvencionados a pagados, así como de quienes se mantuvieron en establecimientos particulares pagados frente a aquellos que se trasladaron a subvencionados. No obstante, estas diferencias no se mantienen de manera sistemática en todos los casos, y en algunos periodos no se observaron diferencias estadísticamente significativas, lo que impide atribuir un efecto directo y sostenido al tipo de trayectoria escolar seguida por los estudiantes.

En definitiva, los resultados del modelo de valor agregado para la prueba de Lenguaje no permiten establecer con certeza una relación directa y consistente entre el tipo de establecimiento educacional y una mejora sustantiva en el rendimiento académico. Aunque en ciertos casos se observaron diferencias significativas, la evidencia apunta a que otros factores (como el rendimiento previo, el capital educativo y cultural del hogar, y el género del estudiante) juegan un rol más determinante en explicar las variaciones en los puntajes obtenidos.

8.2 Modelo de Matemática

8.2.1. SIMCE 2010-2014

Tabla 15
Tabla resumen regresiones prueba de Matemática, año 2010-2014.

	1° Modelo	2° Modelo	3° Modelo	4° Modelo
Intercepto	269,8908***	45,9404***	249,4352***	48,4739***
	(269,8907)***	(45,9404)***	(249,4352)***	(48,4739)***
Colegio Part. Pagado	38,9398***	16,8115***	16,8115***	8,3400***
	(38,9398)***	(15,4951)***	(16,8114)***	(8,3400)***
Nivel 2 de estudio profesor	3,1467***	0,7810*	2,5967***	0,6749
	(3,1467)***	(0,7810)*	(2,5966)***	(0,6749)
Nivel 3 de estudio profesor	1,7918***	0,3503	1,6803***	0,3649
	(1,7917)***	(0,3503)	(1,6803)***	(0,3649)
Ptje. Matemática SIMCE 4B	-	0,5427***	-	0,5232***
	-	(0,5427)***	-	(0,5232)***
Promedio general	-	13,5084***	-	12,7261***
	-	(13,5084)***	-	(12,7261)***
Género Masculino	-	2,4629***	-	2,4192***
	-	(2,4629)***	-	(2,4192)***
Sí convive con la figura materna	-	-	3,3547***	1,6065***
	-	-	(3,3547)***	(1,6066)***
Sí convive con la figura paterna	-	-	1,6367***	0,0026
	-	-	(1,6367)***	(0,0026)
Nivel 2 de estudio figura materna	-	-	4,6694***	1,4737***
	-	-	(4,6693)***	(1,4737)***
Nivel 3 de estudio figura materna	-	-	10,4876***	3,1729***
	-	-	(10,4875)***	(3,172)***
Nivel 2 de estudio figura paterna	-	-	4,5622***	1,5561***
	-	-	(4,5621)***	(1,5561)***
Nivel 3 de estudio figura paterna	-	-	11,5692***	4,7468***
	-	-	(11,5691)***	(4,7468)***
NSE nivel 2	-	-	7,4404***	2,7223**
	-	-	(7,4404)***	(2,7223)**
NSE nivel 3	-	-	13,4839***	5,3026**
	-	-	(13,4838)***	(5,3026)**
Nivel 2 de número de libros	-	-	5,3338***	1,8012***
	-	-	(5,3337)***	(1,8013)***
Nivel 3 de número de libros	-	-	9,7188***	3,0600***
	-	-	(9,7187)***	(3,0600)***
R^2 Ajustado	0,0890	0,5078	0,1573	0,516
Nº Observaciones	84.347			

Notas. *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$. La primera fila de valores para cada una de las variables corresponde a la primera regresión, mientras que la segunda fila de valores (los que se encuentran entre paréntesis) corresponden a los resultados obtenidos mediante la técnica de errores robustos.

En los modelos estimados para la prueba de Matemática, se observa que la variable correspondiente al tipo de colegio presenta un efecto más consistente y de mayor magnitud en comparación con lo observado en la prueba de Lenguaje. En particular, en el modelo final, los estudiantes que asistieron a establecimientos particulares pagados obtuvieron, en promedio, 8,34 puntos más que sus pares de colegios particulares subvencionados, manteniendo constantes el resto de las variables. Esta diferencia es estadísticamente significativa y representa una de las asociaciones más robustas del modelo, lo que sugiere que el tipo de colegio tiene un impacto más claro sobre el rendimiento en Matemática.

En contraste, el nivel académico de los docentes, aunque significativo en modelos aislados, pierde significancia en el modelo completo. Esto indica que su efecto sobre el rendimiento se diluye al incorporar otras variables explicativas, lo que sugiere que no es un determinante estructural del rendimiento en esta prueba cuando se considera el conjunto de factores familiares, escolares e individuales.

Respecto al género del estudiante, el efecto observado es opuesto al de la prueba de Lenguaje: los alumnos del género masculino presentan mejores resultados, con un promedio de 2,41 puntos por encima de las estudiantes femeninas. Esta diferencia, que se mantiene significativa en los modelos estimados, sugiere una ventaja de rendimiento masculino en el área de Matemática, lo que contrasta con la desventaja observada en Lenguaje.

En el ámbito familiar, se observa que la convivencia con la figura materna se asocia con un incremento leve pero significativo en el rendimiento, cercano a 1,6 puntos, mientras que la convivencia con la figura paterna no muestra un efecto relevante en el modelo final. Esto sugiere una posible influencia positiva del acompañamiento materno, aunque de magnitud moderada.

Por otro lado, el nivel educativo de las figuras parentales sigue desempeñando un papel

importante. Cuando estas alcanzan el nivel más alto de escolaridad (nivel 3), el efecto sobre el puntaje es claramente positivo: la figura materna se asocia con un aumento de 3,17 puntos, y la figura paterna con un incremento aún mayor, cercano a los 4,74 puntos. Estos resultados refuerzan la hipótesis de que el capital educativo en el hogar tiene un impacto relevante, particularmente en áreas de razonamiento lógico-matemático.

El nivel socioeconómico muestra nuevamente un efecto positivo y de mayor magnitud que en Lenguaje. Los estudiantes pertenecientes al nivel medio (NSE 2) presentan un aumento de 2,72 puntos, mientras que aquellos en el nivel alto (NSE 3) alcanzan un incremento de 5,30 puntos. Esta asociación confirma que el contexto económico es un determinante importante del rendimiento académico, especialmente en Matemática.

Finalmente, la variable asociada al número de libros en el hogar evidencia una influencia significativa: los estudiantes que se ubican en el nivel más alto de esta categoría presentan un rendimiento promedio 3,06 puntos superior al del grupo base.

En conjunto, los resultados del modelo de valor agregado para Matemática muestran una relación más consistente y significativa entre el tipo de establecimiento educativo y el rendimiento académico. A diferencia de lo observado en Lenguaje, en este caso el tipo de colegio, el nivel socioeconómico, el capital cultural y el género del estudiante emergen como factores con impactos claros y significativos.

8.2.2. SIMCE 2011-2015

Tabla 16
Tabla resumen regresiones prueba de Matemática, año 2011-2015.

	1° Modelo	2° Modelo	3° Modelo	4° Modelo
Intercepto	272,3338***	23,9542***	252,0612***	28,4170***
	(272,3338)***	(23,9542)***	(252,0612)***	(28,4170)***
Colegio Part. Pagado	38,9487***	14,2893***	16,0704***	6,2598***
	(38,9487)***	(14,2893)***	(16,0703)***	(6,2598)***
Nivel 2 de estudio profesor	3,5378***	0,2723	3,2251***	0,2743
	(3,5378)***	(0,2723)	(3,2251)***	(0,2743)
Nivel 3 de estudio profesor	2,2608***	0,1368	1,9875***	0,0946
	(2,2607)***	(0,1368)	(1,9875)***	(0,0946)
Ptje. Matemática SIMCE 4B	-	0,5898***	-	0,5703***
	-	(0,5898)***	-	(0,5703)***
Promedio general	-	15,3229***	-	14,3074***
	-	(15,3229)***	-	(14,3074)***
Género Masculino	-	4,6976***	-	4,7351***
	-	(4,6976)***	-	(4,7351)***
Sí convive con la figura materna	-	-	3,7001***	0,4304
	-	-	(3,7001)***	(0,4304)
Sí convive con la figura paterna	-	-	0,9692**	-0,3899
	-	-	(0,9691)**	(-0,3899)
Nivel 2 de estudio figura materna	-	-	5,2870***	2,2859***
	-	-	(5,2869)***	(2,2859)***
Nivel 3 de estudio figura materna	-	-	11,6443***	3,8944***
	-	-	(11,6443)***	(3,8944)***
Nivel 2 de estudio figura paterna	-	-	4,7811***	2,4626***
	-	-	(4,7810)***	(2,4626)***
Nivel 3 de estudio figura paterna	-	-	11,9655***	5,8140***
	-	-	(11,9655)***	(5,8140)***
NSE nivel 2	-	-	6,6120***	2,8714***
	-	-	(6,6120)***	(2,8714)***
NSE nivel 3	-	-	12,0504***	4,9575***
	-	-	(12,0503)***	(4,9575)***
Nivel 2 de número de libros	-	-	5,2731***	1,8149***
	-	-	(5,2731)***	(1,8149)***
Nivel 3 de número de libros	-	-	10,5133***	3,4820***
	-	-	(10,5132)***	(3,4820)***
R^2 Ajustado	0,0820	0,5261	0,1505	0,5362
Nº Observaciones	76.790			

Notas. *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$. La primera fila de valores para cada una de las variables corresponde a la primera regresión, mientras que la segunda fila de valores (los que se encuentran entre paréntesis) corresponden a los resultados obtenidos mediante la técnica de errores robustos..

Los resultados obtenidos a partir del modelo de regresión para la prueba de Matemática en el periodo 2011–2015 reflejan un patrón consistente con el observado en el periodo anterior. En este caso, la variable correspondiente al tipo de establecimiento educacional mantiene su relevancia estadística. Específicamente, se estima que los estudiantes que asistieron a colegios particulares pagados obtuvieron, en promedio, 6,26 puntos más que aquellos que asistieron a colegios particulares subvencionados, manteniéndose constante el resto de los factores considerados. Este efecto, si bien levemente menor al estimado para el periodo 2010–2014, sigue siendo significativo y confirma la existencia de una brecha asociada al tipo de colegio en el rendimiento en Matemática.

En lo referente al nivel educativo de los docentes, al igual que en el modelo anterior, los efectos positivos que inicialmente se observan en los modelos aislados desaparecen en el modelo completo. Esto indica que el nivel de estudios del profesor no presenta un impacto significativo cuando se consideran simultáneamente las demás variables explicativas, y sugiere que este no constituye un determinante estructural del desempeño en Matemática.

Por su parte, el género del estudiante presenta nuevamente un efecto positivo a favor del alumnado masculino. Los estudiantes de género masculino muestran, en promedio, un puntaje superior en cerca de 4,73 puntos respecto a sus pares femeninas. Este efecto es considerablemente más alto que el observado en el periodo anterior y refuerza la tendencia de mayor rendimiento masculino en esta área específica.

En cuanto a las variables del entorno familiar, la convivencia con la figura materna no tiene un efecto significativo en el modelo final, mientras que la convivencia con la figura paterna presenta un efecto negativo, aunque de magnitud muy baja (inferior a medio punto). Esta relación sugiere que, al controlar por otros factores, la convivencia con estas figuras parentales no incide de forma concluyente en el rendimiento matemático.

No obstante, el nivel educativo de las figuras parentales se mantiene como un predictor relevante. La figura paterna con escolaridad de nivel 3 se asocia con un incremento de 5,81 puntos en promedio, mientras que la figura materna con el mismo nivel educativo aporta un aumento de aproximadamente 3,89 puntos. Estos resultados refuerzan el valor del capital educativo del hogar en la explicación del desempeño académico, especialmente cuando ambos padres alcanzan niveles más altos de escolaridad.

Respecto al nivel socioeconómico, el modelo muestra un patrón claro y positivo: los estudiantes del NSE nivel 2 obtienen en promedio 2,87 puntos más, mientras que aquellos del nivel 3 presentan un incremento de cerca de 4,96 puntos, en comparación con los del nivel más bajo. Estos efectos sugieren que el entorno económico del estudiante es un componente importante en el rendimiento matemático, posiblemente debido a mejores condiciones materiales y de acceso a apoyo educativo.

Por último, la disponibilidad de libros en el hogar sigue mostrando una influencia positiva. Aquellos estudiantes que reportan el nivel más alto de acceso a libros (nivel 3) logran un aumento promedio de 3,48 puntos en la prueba.

En resumen, el modelo estimado para el periodo 2011–2015 reafirma las conclusiones del análisis previo: el tipo de colegio, el nivel educativo de los padres, el nivel socioeconómico y el entorno cultural del hogar continúan siendo determinantes significativos del rendimiento en la prueba de Matemática.

8.2.3. *Aplicación de T-Test Pruebas de Matemática*

Se aplicó la misma metodología bajo los mismos fundamentos aplicados para la prueba de Lenguaje, pero ahora para la prueba de Matemática.

8.2.3.1. Grupo 1: Part. Subvencionados a Part. Subvencionados vs. Grupo 3: Part.

Subvencionados a Part. Pagados. Para estos grupos, los resultados son los siguientes:

Tabla 17

Resultados T-test prueba de Matemática G1 vs G3, año 2010-2014.

	Valor
Estadístico t	-5,04
Valor p	5,209·e-07

Tabla 18

Resultados T-test prueba de Matemática G1 vs G3, año 2011-2015.

	Valor
Estadístico t	-5,95
Valor p	3,591·e-09

Los resultados del T-test para el periodo 2010–2014 muestran una diferencia estadísticamente significativa entre los grupos G1 y G3. En particular, los alumnos que se mantuvieron en colegios particulares subvencionados (G1) obtuvieron, en promedio, un puntaje 5,04 puntos inferior en la prueba de Matemática en comparación con aquellos que migraron a colegios particulares pagados (G3).

Por lo tanto, se concluye que el cambio desde un establecimiento particular subvencionado a uno particular pagado se asocia con una mejora significativa en el rendimiento en Matemática para este período.

Los resultados obtenidos para el periodo 2011–2015 indican una diferencia estadísticamente significativa entre los grupos G1 y G3. Específicamente, los estudiantes que comenzaron en un colegio particular subvencionado y luego se trasladaron a uno particular pagado (G3) obtuvieron, en promedio, 5,95 puntos más en la prueba de Matemática que aquellos que permanecieron en el mismo tipo de establecimiento (G1).

En consecuencia, al igual que en el análisis del periodo anterior, se evidencia que el cambio hacia un establecimiento particular pagado está asociado a una mejora significativa en el rendimiento en Matemática.

8.2.3.2. Grupo 2: Part. Pagados a Part. Pagados vs. Grupo 4: Part. Pagados a Part.

Subvencionados. Para estos grupos, los valores son los siguientes:

Tabla 19

Resultados T-test prueba de Matemática G2 vs G4, año 2010-2014.

	Valor
Estadístico t	0,27
Valor p	0,7839

Tabla 20

Resultados T-test prueba de Matemática G2 vs G4, año 2011-2015.

	Valor
Estadístico t	0,7
Valor p	0,4867

En la comparación entre los grupos G1 y G3 para el periodo 2010–2014, los resultados del T-test indican que no existe evidencia estadística suficiente para afirmar que hay una diferencia significativa entre los puntajes obtenidos por ambos grupos.

Esto sugiere que los estudiantes que se trasladaron desde un colegio particular pagado a uno particular subvencionado no presentan un rendimiento significativamente distinto en la prueba de Matemática en comparación con aquellos que permanecieron en colegios particular pagados.

Al igual que en el año anterior, los resultados del T-test para el periodo 2011–2015 indican que no existe una diferencia significativa entre los grupos G1 y G3.

En consecuencia, no se observa una mejora ni un empeoramiento en el rendimiento en la prueba de Matemática entre los estudiantes que permanecieron en colegios particular pagados y aquellos que se trasladaron a colegios particular subvencionados.

8.2.4. Conclusiones Modelo de Matemática

En relación con los coeficientes estimados en los modelos para la prueba de Matemática, se observa que asistir a un establecimiento particular pagado se asocia, en promedio, con un incremento de aproximadamente 6 puntos en el rendimiento respecto a aquellos estudiantes que

cursaron sus estudios en colegios particulares subvencionados. Esta diferencia, contrasta a lo observado en la prueba de Lenguaje, se presenta de forma más consistente y con mayor magnitud, lo que sugiere un efecto más robusto del tipo de colegio sobre el desempeño en esta área.

En cuanto al nivel educativo del profesorado, los resultados no evidencian un patrón claro ni uniforme. En general, no se detecta una relación sistemática entre un mayor grado académico de los docentes y un aumento significativo en los puntajes de Matemática. Esto indica que, si bien la formación del profesorado podría influir en el rendimiento, no constituye un factor determinante dentro de los modelos estimados.

Respecto a las variables familiares, la mera presencia de las figuras parentales en el hogar no mostró efectos significativos sobre el desempeño académico. No obstante, el nivel educativo alcanzado por las figuras materna y paterna sí evidenció una influencia positiva. En particular, el máximo nivel considerado para la figura paterna (nivel 3) se asocia con un incremento promedio cercano a 6 puntos en la prueba de Matemática, lo que destaca el rol del capital educativo familiar en el logro académico.

En relación con las condiciones del hogar, el nivel socioeconómico presentó un efecto positivo moderado. Los estudiantes pertenecientes al nivel más alto de esta categoría obtuvieron, en promedio, hasta 5 puntos adicionales en comparación con aquellos de menor NSE. Asimismo, el número de libros en el hogar mostró una influencia positiva, aunque más acotada: los alumnos en el nivel superior de esta variable lograron incrementos de hasta 3 puntos en su rendimiento, reflejando la importancia del acceso a recursos culturales en el hogar.

Complementando el análisis con pruebas de comparación de medias (T-Test), se observaron diferencias significativas entre los grupos G1 (alumnos que permanecieron en colegios particulares subvencionados) y G3 (alumnos que migraron a particulares pagados). En ambos periodos analiza-

dos, el cambio hacia un establecimiento particular pagado se asoció con mejoras estadísticamente significativas en el rendimiento, con diferencias cercanas a los 6 puntos. Por el contrario, la comparación entre los grupos G2 (que permanecieron en colegios pagados) y G4 (que migraron a subvencionados) no arrojó diferencias significativas en ninguno de los periodos, lo que sugiere que el traslado hacia un colegio subvencionado no conllevó un deterioro sustantivo en el rendimiento.

En suma, los resultados para la prueba de Matemática indican una relación más clara y consistente entre el tipo de establecimiento y el desempeño académico. Sin embargo, al igual que en Lenguaje, los factores familiares y contextuales (en especial el nivel educativo de los padres y el capital cultural del hogar) juegan un papel relevante y complementario en la explicación del rendimiento de los estudiantes.

8.3 Matriz de Correlaciones

Finalmente, para complementar los análisis anteriores, se construyeron matrices de correlación para cada año y para ambas pruebas estandarizadas. Estas se encuentran representadas en: **Anexo G, Anexo H, Anexo I y Anexo J.**

Los resultados muestran que las correlaciones de mayor magnitud se mantienen consistentes en ambos periodos analizados y en ambas pruebas (Lenguaje y Matemática). En particular, se destaca la fuerte correlación entre los promedios obtenidos por los estudiantes en cuarto básico y tanto su promedio general como el puntaje alcanzado en octavo básico. Esto sugiere que un buen desempeño académico en cuarto básico tiende a mantenerse en el tiempo, reflejándose en mejores resultados en niveles posteriores.

En cuanto a la variable tipo de colegio, se observa una alta correlación con el nivel educativo máximo alcanzado por las figuras materna y paterna, así como con el nivel socioeconómico del estudiante. Este hallazgo es consistente con la literatura previa, que señala una estrecha relación entre el tipo de establecimiento y las condiciones familiares y sociales del entorno del alumno.

9. Conclusiones

El presente estudio tuvo como objetivo evaluar el impacto del tipo de establecimiento educacional en el desempeño académico de los estudiantes, utilizando modelos de valor agregado aplicados a las pruebas SIMCE de Lenguaje y Matemática, en los periodos 2010–2014 y 2011–2015. Se incorporaron diversas variables de control relacionadas con el entorno familiar, el nivel socioeconómico y el capital cultural del hogar, con el fin de aislar el efecto del establecimiento y obtener estimaciones más robustas.

Los resultados obtenidos para la prueba de Lenguaje no permiten establecer una relación clara ni consistente entre el tipo de colegio y el rendimiento académico. Aunque en el periodo 2010–2014 se observó una diferencia significativa a favor de los estudiantes que migraron desde colegios particulares subvencionados a particulares pagados, esta diferencia no se replicó en el periodo 2011–2015. Además, la magnitud del efecto asociado al tipo de colegio fue moderada, y en varios modelos no resultó estadísticamente significativa. Esto sugiere que la relación entre el tipo de establecimiento y el desempeño en Lenguaje podría depender de factores contextuales específicos de cada cohorte o año.

En contraste, para la prueba de Matemática, los resultados fueron más consistentes y robustos. En ambos periodos, los estudiantes que asistieron a establecimientos particulares pagados obtuvieron mejores resultados en comparación con sus pares de colegios particulares subvencionados, con diferencias cercanas a los 6 puntos. Además, el análisis de T-Test mostró diferencias estadísticamente significativas entre los grupos G1 y G3, confirmando que el cambio hacia un colegio particular pagado se asoció con un mejor rendimiento en Matemática. Por el contrario, los estudiantes que se trasladaron en sentido inverso (G2 a G4) no mostraron un deterioro en su de-

sempañ, lo que sugiere que el tipo de establecimiento puede tener un efecto positivo, pero su ausencia no necesariamente implica una pérdida significativa.

En términos de los factores asociados al entorno familiar, destaca el efecto positivo del nivel educativo de las figuras parentales. Tanto en Lenguaje como en Matemática, se observó una relación directa entre un mayor nivel académico de los padres y mejores resultados en las pruebas, siendo este efecto particularmente fuerte para la figura paterna. No obstante, la presencia o ausencia de estas figuras en el hogar no mostró efectos significativos, salvo en el caso de Lenguaje, donde la convivencia con la figura paterna se asoció a una leve disminución en el puntaje.

En cuanto a las variables de capital cultural y socioeconómico, como el nivel socioeconómico y el número de libros en el hogar, ambas mostraron efectos positivos en el rendimiento, con impactos más notorios en la prueba de Matemática. Estos resultados refuerzan la importancia del contexto familiar en el desarrollo académico de los estudiantes, especialmente en lo relativo a los recursos culturales y materiales disponibles.

En conjunto, los hallazgos de esta investigación indican que, si bien el tipo de establecimiento puede influir en el desempeño académico (particularmente en Matemática), este no es el único ni el principal factor determinante. La influencia del entorno familiar, el capital educativo de los padres y el acceso a recursos culturales también juegan un rol fundamental. Asimismo, la falta de consistencia en los resultados para la prueba de Lenguaje sugiere que los efectos del tipo de colegio pueden variar en el tiempo y deben ser interpretados con cautela.

10. Limitaciones

Este estudio presenta ciertas limitaciones que deben ser consideradas para interpretar adecuadamente los resultados y su alcance. Reconocer estas restricciones permite enmarcar los hallazgos dentro de su contexto metodológico y analítico, evitando así generalizaciones inapropiadas o conclusiones que excedan lo observado.

Para comenzar, si bien el estudio utiliza información proveniente del SIMCE, una fuente oficial y estandarizada de evaluación a nivel nacional, el análisis se restringe únicamente a dos cohortes de estudiantes: aquellos evaluados en los periodos 2010–2014 y 2011–2015. Esto implica que los resultados obtenidos son representativos solo para dichos grupos, y no pueden ser extrapolados directamente a la totalidad del sistema educativo chileno ni a otras generaciones de estudiantes. Además, el seguimiento longitudinal se basa en la disponibilidad de *run* coincidentes entre años, lo que excluye a estudiantes con datos incompletos.

Una segunda limitación se relaciona con el marco temporal de la investigación. Al enfocarse en periodos específicos, el estudio no permite observar efectos de largo plazo ni variaciones estructurales del sistema educativo que pudieran haber influido en el rendimiento estudiantil con posterioridad. Aunque los periodos analizados fueron seleccionados de forma estratégica para abarcar trayectorias escolares completas y evaluar el impacto de ciertos cambios en la composición de los establecimientos, la falta de datos más recientes impide evaluar si los patrones observados se mantienen, intensifican o revierten en el tiempo.

Respecto a las variables utilizadas, si bien se incluyeron factores individuales, familiares y escolares, algunas dimensiones relevantes del proceso educativo no pudieron ser consideradas por limitaciones propias del set de datos. Esto podría dejar fuera factores relevantes que, al no estar

controlados, podrían incidir en los resultados.

Asimismo, las variables socioeconómicas y familiares utilizadas, como el nivel de estudios de los padres o la cantidad de libros en el hogar, corresponden a autodeclaraciones hechas en el contexto de la aplicación del SIMCE. Aunque estas variables son informativas, pueden presentar ciertos grados de error o imprecisión, lo cual podría afectar la estimación de sus efectos reales. Adicionalmente, estas medidas capturan condiciones en un momento específico y no necesariamente reflejan cambios que pueden haber ocurrido durante el periodo de estudio.

Otro aspecto a considerar es la naturaleza del análisis cuantitativo desarrollado. Si bien los modelos de regresión y las pruebas T-test permiten detectar asociaciones estadísticas significativas, estos métodos no permiten establecer relaciones de causalidad directa entre las variables. Por tanto, los efectos observados deben interpretarse como correlaciones controladas por un conjunto de covariables, y no como consecuencias atribuibles exclusivamente a un factor aislado, como el tipo de establecimiento.

Finalmente, es importante destacar que el enfoque del modelo de valor agregado, si bien ofrece una mejora metodológica al considerar el rendimiento previo de los estudiantes, sigue estando expuesto a posibles sesgos por variables omitidas no observables. Factores como la motivación del estudiante, apoyo familiar no capturado por las variables disponibles, o decisiones estratégicas de cambio de establecimiento pueden influir en los resultados sin estar adecuadamente reflejados en los modelos.

En conjunto, estas limitaciones no invalidan los hallazgos del estudio, pero sí deben ser tenidas en cuenta al momento de interpretar sus implicancias. Futuros trabajos podrían ampliar el análisis incluyendo más cohortes, variables complementarias o métodos cualitativos que permitan captar dimensiones menos observables del rendimiento académico.

11. Anexos

Anexo A: Variables utilizadas desde la base de datos de alumnos provista por el SIMCE.

Tabla 21

Variables utilizadas desde la base de datos de alumnos provista por el SIMCE.

	2010 y 2014	2011 y 2015	Uso
Variables	mrun	mrun	Identificador único asignado por el MINEDUC a cada estudiante. Permite el cruce transversal de información entre distintos años y con las bases provenientes del MINE-DUC.
	idalumno	idalumno	Identificador interno del SIMCE para cada alumno dentro de un año determinado. Es esencial para vincular las encuestas respondidas por padres, apoderados y profesores.
	gen_alu	gen_alu	Permite identificar el género del alumno, es decir, si este es masculino o femenino.
	rbd	rbd	Corresponde al código único asignado a cada establecimiento educacional. Esta variable permite vincular la base de datos de los estudiantes con la de los establecimientos, facilitando así la identificación del tipo de colegio al que asiste cada alumno.
	cod_curso	cod_curso	Identificador del curso al que pertenece el estudiante dentro del establecimiento. Permite relacionar al alumno con los docentes correspondientes a su curso.
	ptje_lect4b_alu	ptje_lect4b_alu	Puntaje obtenido por el estudiante en la prueba de Lenguaje en 4° básico.
	ptje_mate4b_alu	ptje_mate4b_alu	Puntaje obtenido por el estudiante en la prueba de Matemática en 4° básico.

Tabla 22Continuación de **Tabla 21**.

	2010 y 2014	2011 y 2015	Uso
Variables	noptje_lect4b_alu noptje_lect8b_alu	noptje_lect4b_alu noptje_lect8b_alu	Permite identificar si el estudiante posee un puntaje válido en la prueba de Lenguaje, tanto en 4° básico como en 8° básico. Esta variable resulta fundamental para los procesos de depuración de las bases de datos y la selección de observaciones válidas para el análisis.
	noptje_mate4b_alu noptje_mate8b_alu	noptje_mate4b_alu noptje_mate8b_alu	Permite identificar si el estudiante posee un puntaje válido en la prueba de Matemática, tanto en 4° básico como en 8° básico. Esta variable resulta fundamental para los procesos de depuración de las bases de datos y la selección de observaciones válidas para el análisis.
		noptje_seg_lect8b	Variable que indica si, por razones de seguridad, el estudiante dispone o no de un puntaje válido en la prueba de Lenguaje. Esta información es utilizada en el proceso de depuración de la base de datos, permitiendo excluir observaciones no aptas para el análisis.
		noptje_seg_mate8b	Variable que indica si, por razones de seguridad, el estudiante dispone o no de un puntaje válido en la prueba de Matemática. Esta información es utilizada en el proceso de depuración de la base de datos, permitiendo excluir observaciones no aptas para el análisis.

Anexo B: Variables utilizadas de la bases de datos de apoderados provista por SIMCE.**Tabla 23***Variables utilizadas de la bases de datos de apoderados provista por SIMCE.*

	2010 y 2014	2011 y 2015	Uso
Variables	idalumno	idalumno	Identificador interno del SIMCE para cada alumno dentro de un año determinado. Es esencial para vincular las encuestas respondidas por los apoderados y estudiantes.
	rbd	rbd	Corresponde al código único asignado a cada establecimiento educacional. Esta variable permite vincular la base de datos de los estudiantes con la de los establecimientos, facilitando así la identificación del tipo de colegio al que asiste cada alumno.
	cpad_p05_01 cpad_p05_02 cpad_p05_03 cpad_p05_04	cpad_p05_01 cpad_p05_02 cpad_p05_03 cpad_p05_04	VARIABLES para identificar si el estudiante reside con una figura paterna y/o materna, lo cual permite contextualizar su entorno familiar y de apoyo en el hogar.
	cpad_p08	cpad_p07	Variable que permite identificar la cantidad de libros presentes en el hogar del estudiante. Esta variable, en conjunto con el nivel educativo de los padres, permite dimensionar el capital cultural disponible en el hogar y su posible influencia en el rendimiento académico del estudiante.

Tabla 24

Continuación de Tabla 23.

	2010 y 2014	2011 y 2015	Uso
Variables	cpad_p09_01	cpad_p08_01	<p>Permite identificar el nivel educativo alcanzado por la figura paterna del estudiante. Esta información permite evaluar si existe una relación significativa entre el rendimiento académico del estudiante y el grado académico de su figura paterna.</p>
	cpad_p09_02	cpad_p08_02	
	cpad_p09_03	cpad_p08_03	
	cpad_p09_04	cpad_p08_04	
	cpad_p09_05	cpad_p08_05	
	cpad_p09_06	cpad_p08_06	
	cpad_p09_07	cpad_p08_07	
	cpad_p09_08	cpad_p08_08	
	cpad_p09_09	cpad_p08_09	
	cpad_p09_10	cpad_p08_10	
	cpad_p09_11	cpad_p08_11	
	cpad_p09_12	cpad_p08_12	
	cpad_p09_13	cpad_p08_13	
	cpad_p09_14	cpad_p08_14	
	cpad_p09_15	cpad_p08_15	
	cpad_p09_16	cpad_p08_16	
	cpad_p09_17	cpad_p08_17	
	cpad_p09_18	cpad_p08_18	
	cpad_p09_19	cpad_p08_29	
	cpad_p09_20	cpad_p08_20	
	cpad_p09_21	cpad_p08_21	
cpad_p10_01	cpad_p09_01	<p>Permite identificar el nivel educativo alcanzado por la figura materna del estudiante. Esta variable permite analizar la relación entre el desempeño académico del estudiante y el capital educativo de su figura materna.</p>	
cpad_p10_02	cpad_p09_02		
cpad_p10_03	cpad_p09_03		
cpad_p10_04	cpad_p09_04		
cpad_p10_05	cpad_p09_05		
cpad_p10_06	cpad_p09_06		
cpad_p10_07	cpad_p09_07		
cpad_p10_08	cpad_p09_08		
cpad_p10_09	cpad_p09_09		
cpad_p10_10	cpad_p09_10		
cpad_p10_11	cpad_p09_11		
cpad_p10_12	cpad_p09_12		
cpad_p10_13	cpad_p09_13		
cpad_p10_14	cpad_p09_14		
cpad_p10_15	cpad_p09_15		
cpad_p10_16	cpad_p09_16		
cpad_p10_17	cpad_p09_17		
cpad_p10_18	cpad_p09_18		
cpad_p10_19	cpad_p09_29		
cpad_p10_20	cpad_p09_20		
cpad_p10_21	cpad_p09_21		

Tabla 25*Continuación Tabla 24 .*

	2010 y 2014	2011 y 2015	Uso
Variables	cpad_p11_01	cpad_p10_01	Para identificar la cantidad de ingresos mensuales percibidos en el hogar del estudiante. Esta variable sirve para estimar su nivel socioeconómico y su posible influencia en el acceso a recursos o actividades que podrían afectar su rendimiento académico.
	cpad_p11_02	cpad_p10_02	
	cpad_p11_03	cpad_p10_03	
	cpad_p11_04	cpad_p10_04	
	cpad_p11_05	cpad_p10_05	
	cpad_p11_06	cpad_p10_06	
	cpad_p11_07	cpad_p10_07	
	cpad_p11_08	cpad_p10_08	
	cpad_p11_09	cpad_p10_09	
	cpad_p11_10	cpad_p10_10	
	cpad_p11_11	cpad_p10_11	
	cpad_p11_12	cpad_p10_12	
	cpad_p11_13	cpad_p10_13	
	cpad_p11_14	cpad_p10_14	
	cpad_p11_15	cpad_p10_15	

Anexo C: Variables utilizadas de la bases de datos de los profesores provista por SIMCE.**Tabla 26***Variables utilizadas de la bases de datos de los profesores provista por SIMCE.*

	2010 y 2014	2011 y 2015	Uso
Variables	rbd	rbd	Corresponde al código único asignado a cada establecimiento educacional. Esta variable permite vincular la base de datos de los estudiantes con la de los establecimientos, facilitando así la identificación del tipo de colegio al que asiste cada alumno.
	cod_curso	cod_curso	Identificador del curso al que pertenece el estudiante dentro del establecimiento. Permite relacionar al alumno con los docentes correspondientes a su curso.
	cprof_p05 cprof_p10_01 cprof_p10_02 cprof_p10_03 cprof_p10_04 cprof_p10_05	cprof_p03 cprof_p06_01 cprof_p06_02 cprof_p06_03 cprof_p06_04 cprof_p06_05	Para identificar el grado académico del profesor del estudiante. Esta información permite evaluar si existe una relación entre el rendimiento académico del estudiante y el nivel de formación alcanzado por sus docentes.

Anexo D: Variables utilizadas de la bases de datos de los establecimientos educacionales provista por SIMCE.**Tabla 27***Variables utilizadas de la bases de datos de los establecimientos educacionales provista por SIMCE.*

	2010 y 2014	2011 y 2015	Uso
Variables	rbd	rbd	Corresponde al código único asignado a cada establecimiento educacional. Esta variable permite vincular la base de datos de los estudiantes con la de los establecimientos, facilitando así la identificación del tipo de colegio al que asiste cada alumno.
	cod_depe2	cod_depe2	Código que permite identificar el tipo de establecimiento al que asiste el estudiante. Para los fines de esta investigación, solo se consideran aquellos pertenecientes a los tipos particular subvencionado y particular pagado.

Anexo E: Variables utilizadas de la bases de datos del MINEDUC.**Tabla 28***Variables utilizadas de la bases de datos del MINEDUC.*

	2010 y 2014	2011 y 2015	Uso
Variables	mr _{run}	mr _{run}	Identificador único asignado por el MINEDUC a cada estudiante. Permite el cruce transversal de información entre distintos años y con las bases provenientes del MINEDUC.
	prom _{gral}	prom _{gral}	Esta variable corresponde a la calificación promedio final obtenida por el estudiante en una escala de 1 a 7. Su inclusión en el modelo permite evaluar si existe una relación entre el rendimiento académico general del alumno y los puntajes obtenidos en la prueba SIMCE. Dado que el objetivo es analizar el efecto de este desempeño previo en los resultados obtenidos en 8° básico, esta variable se considera únicamente para los años 2010 y 2011, es decir, cuando los estudiantes cursaban 4° básico.
	sit _{fin}	sit _{fin}	Esta variable indica la condición académica del estudiante al finalizar el año escolar, especificando si fue promovido al curso siguiente o si repitió el año. Su uso en el análisis permite realizar una depuración adecuada de la base de datos, excluyendo aquellos casos que podrían introducir sesgos en la estimación del modelo, como alumnos que no progresaron de curso.

Anexo F: Variables creadas para el modelo econométrico**Tabla 29***Variables creadas para el modelo econométrico.*

Variable Creada	Categorías Pertencientes	Tipo de Variable
tipo_colegio	<ul style="list-style-type: none">• Particular Subvencionado• Particular Pagado	Binaria
gen_alu	<ul style="list-style-type: none">• Femenino• Masculino	Binaria
nivel_estudio_profesor_1	<ul style="list-style-type: none">• Educación escuela normal• Educación Instituto o centro de formación técnica	Binaria
nivel_estudio_profesor_2	<ul style="list-style-type: none">• Educación universidad• Educación grado Diplomado	Binaria
nivel_estudio_profesor_3	<ul style="list-style-type: none">• Educación grado Postítulo• Educación grado Magíster• Educación grado Doctorado	Binaria
vive_figura_materna	<ul style="list-style-type: none">• Vive con figura materna• No vive con figura materna	Binaria
vive_figura_paterna	<ul style="list-style-type: none">• Vive con figura paterna• No vive con figura paterna	Binaria
nivel_estudio_figura_materna_1	<ul style="list-style-type: none">• No estudió• Educación básica completa/ incompleta• Educación media completa/ incompleta	Binaria
nivel_estudio_figura_materna_2	<ul style="list-style-type: none">• Educación centro de formación técnica completa/ incompleta	Binaria
nivel_estudio_figura_materna_3	<ul style="list-style-type: none">• Educación universidad completa/ incompleta• Educación grado magíster• Educación grado doctorado	Binaria
nivel_estudio_figura_paterna_1	<ul style="list-style-type: none">• No estudió• Educación básica completa/ incompleta• Educación media completa/ incompleta	Binaria
nivel_estudio_figura_paterna_2	<ul style="list-style-type: none">• Educación centro de formación técnica completa/ incompleta	Binaria

Tabla 30Continuación *Tabla 29*.

Variable Creada	Categorías Pertencientes	Tipo de Variable
nivel_estudio_figura_paterna_3	<ul style="list-style-type: none">• Educación universidad completa/incompleta• Educación grado magíster• Educación grado doctorado	Binaria
nse_1	<ul style="list-style-type: none">• Menos de \$100.000• Entre \$100.001 y \$200.00	Binaria
nse_2	<ul style="list-style-type: none">• Entre \$200.001 y \$300.000• Entre \$300.001 y \$400.000• Entre \$400.001 y \$500.000	Binaria
nse_3	<ul style="list-style-type: none">• Entre \$500.001 y \$600.000• Entre \$600.001 y \$800.000• Entre \$800.001 y \$1.000.000• Entre \$1.000.001 y \$1.200.000• Entre \$1.200.002 y \$1.400.000• Entre \$1.400.001 y \$1.600.000• Entre \$1.600.001 y \$1.800.000• Entre \$1.800.001 y \$2.000.000• Entre \$2.000.001 y \$2.200.000• Sobre \$2.200.000	Binaria
numero_libros_1	<ul style="list-style-type: none">• Vacío• Doble marca• Ninguno• Menos de 10	Binaria
numero_libros_2	<ul style="list-style-type: none">• Entre 10 y 50	Binaria
numero_libros_3	<ul style="list-style-type: none">• Entre 51 y 100• Más de 100	Binaria

Anexo G: Resultado prueba de heterocedasticidad.**Tabla 31***Resultado prueba de heterocedasticidad mediante el test de Breusch-Pagan.*

Año	Prueba		Estadístico	Valor P	Conclusión
2010	Lenguaje	1 Modelo	16,5315	0,0009	Existe Heterocedasticidad
		2 Modelo	187,5248	0	Existe Heterocedasticidad
		3 Modelo	16,9964	0,1995	No existe Heterocedasticidad
		3 Modelo	210,5466	0	Existe Heterocedasticidad
	Matemática	1 Modelo	227,7846	0	Existe Heterocedasticidad
		2 Modelo	198,5774	0	Existe Heterocedasticidad
		3 Modelo	193,8299	0	Existe Heterocedasticidad
		4 Modelo	209,4643	0	Existe Heterocedasticidad
2011	Lenguaje	1 Modelo	28,1214	0	Existe Heterocedasticidad
		2 Modelo	262,5093	0	Existe Heterocedasticidad
		3 Modelo	40,3394	0,0001	Existe Heterocedasticidad
		3 Modelo	279,0638	0	Existe Heterocedasticidad
	Matemática	1 Modelo	358,3921	0	Existe Heterocedasticidad
		2 Modelo	391,5582	0	Existe Heterocedasticidad
		3 Modelo	351,7767	0	Existe Heterocedasticidad
		4 Modelo	415,6656	0	Existe Heterocedasticidad

Anexo H: Matriz correlación representada por números, prueba Lenguaje año 2010-2014.

Tabla 32
Matriz correlación representada por números, prueba Lenguaje año 2010-2014.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.00	0.20	0.03	0.02	0.59	0.47	-0.10	0.02	0.04	0.05	0.21	0.03	0.22	-0.07	0.21	-0.03	0.17
2	0.20	1.00	0.05	-0.01	0.18	0.20	0.01	0.02	0.07	0.06	0.45	-0.02	0.50	-0.29	0.48	-0.12	0.32
3	0.03	0.05	1.00	0.05	0.02	0.04	0.00	0.00	0.01	0.01	0.04	0.01	0.04	-0.03	0.04	-0.02	0.01
4	0.02	-0.01	0.05	1.00	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.59	0.18	0.02	0.02	1.00	0.51	-0.09	0.02	0.03	0.07	0.19	0.04	0.20	-0.06	0.21	-0.02	0.15
6	0.47	0.20	0.04	0.01	0.51	1.00	-0.10	0.05	0.10	0.06	0.20	0.03	0.21	-0.07	0.22	-0.02	0.15
7	-0.10	0.01	0.00	0.01	-0.09	-0.10	1.00	0.00	0.02	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01
8	0.02	0.02	0.00	0.00	0.02	0.05	0.00	1.00	0.19	0.01	0.00	0.01	0.02	0.01	0.02	0.02	0.02
9	0.04	0.07	0.01	0.00	0.03	0.10	0.02	0.19	1.00	0.01	0.03	0.03	0.07	0.03	0.15	0.01	0.05
10	0.05	0.06	0.01	0.00	0.07	0.06	0.00	0.01	0.01	1.00	-0.25	0.26	0.09	-0.01	0.18	0.05	0.06
11	0.21	0.45	0.04	0.00	0.19	0.20	0.00	0.01	0.03	-0.25	1.00	0.00	0.50	-0.24	0.46	-0.11	0.34
12	0.03	-0.02	0.01	0.00	0.02	0.05	0.00	0.00	0.03	0.26	0.00	1.00	-0.24	0.02	0.13	0.05	0.02
13	0.22	0.50	0.04	0.00	0.02	0.06	0.00	0.01	0.03	0.09	0.50	-0.24	1.00	0.02	0.52	-0.09	0.33
14	-0.07	-0.29	-0.03	0.00	-0.06	-0.07	0.00	0.02	0.07	0.09	-0.27	0.02	1.00	-0.27	0.52	-0.09	0.33
15	0.21	0.48	0.04	0.00	0.21	0.22	0.00	0.01	0.03	0.26	0.00	0.02	-0.27	1.00	-0.58	0.10	-0.13
16	-0.03	-0.12	-0.02	0.00	-0.02	-0.02	0.00	0.01	0.15	0.18	0.46	0.13	0.52	-0.58	1.00	-0.05	0.33
17	0.17	0.32	0.01	0.00	0.15	0.15	0.01	0.02	0.05	0.06	0.34	0.02	0.33	-0.13	0.33	1.00	-0.57
17																	1.00

Nota. La tabla muestra las correlaciones entre variables. El valor de la fila indica la correlación con cada columna correspondiente.

Anexo I: Matriz correlación representada por números, pueba Matemática año 2010-2014.

Tabla 33
Matriz correlacion representada por números, pueba Matemática año 2010-2014.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.00	0.30	0.04	0.02	0.50	0.49	0.06	0.04	0.07	0.08	0.27	0.05	0.30	-0.10	0.30	-0.03	0.21
2	0.30	1.00	0.05	-0.01	0.18	0.20	0.01	0.02	0.07	0.06	0.45	-0.02	0.50	-0.29	0.48	-0.12	0.32
3	0.04	0.05	1.00	0.05	0.02	0.04	0.00	0.00	0.01	0.01	0.04	0.01	0.04	-0.03	0.04	-0.02	0.01
4	0.02	-0.01	0.05	1.00	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.50	0.18	0.02	0.02	1.00	0.51	-0.09	0.02	0.03	0.07	0.19	0.04	0.20	-0.06	0.21	-0.02	0.15
6	0.49	0.20	0.04	0.01	0.51	1.00	-0.10	0.05	0.10	0.06	0.20	0.03	0.21	-0.07	0.22	-0.02	0.15
7	0.06	0.01	0.00	0.01	-0.09	-0.10	1.00	0.00	0.02	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01
8	0.04	0.02	0.00	0.00	0.02	0.05	0.00	1.00	0.19	0.01	0.00	0.01	0.02	0.01	0.02	0.02	0.02
9	0.07	0.07	0.01	0.00	0.03	0.10	0.02	0.19	1.00	0.01	0.03	0.03	0.07	0.03	0.15	0.01	0.05
10	0.08	0.06	0.01	0.00	0.07	0.06	0.00	0.01	0.01	1.00	-0.25	0.26	0.09	-0.01	0.18	0.05	0.06
11	0.27	0.45	0.04	0.00	0.19	0.20	0.01	0.00	0.03	-0.25	1.00	0.00	0.50	-0.24	0.46	-0.11	0.34
12	0.05	-0.02	0.01	0.00	0.06	0.06	0.00	0.00	0.03	0.26	1.00	1.00	-0.24	0.02	0.13	0.05	0.02
13	0.30	0.50	0.04	0.00	0.03	0.03	0.00	0.01	0.03	0.26	0.00	-0.24	1.00	0.02	0.52	-0.09	0.33
14	-0.10	-0.29	-0.03	0.00	0.07	0.07	0.01	0.02	0.07	0.09	0.50	-0.24	1.00	-0.27	0.52	-0.09	0.33
15	0.30	0.48	0.04	0.00	-0.06	-0.07	0.00	0.01	0.03	-0.01	-0.24	0.02	-0.27	1.00	-0.58	0.10	-0.13
16	-0.03	-0.12	-0.02	0.00	0.21	0.22	0.01	0.02	0.15	0.18	0.46	0.13	0.52	-0.58	1.00	-0.05	0.33
17	0.21	0.32	0.01	0.00	-0.02	-0.02	0.00	0.02	0.01	0.05	-0.11	0.05	-0.09	0.10	-0.05	1.00	-0.57
17	0.21	0.32	0.01	0.00	0.15	0.15	0.01	0.02	0.05	0.06	0.34	0.02	0.33	-0.13	0.33	-0.57	1.00

Nota. La tabla muestra las correlaciones entre variables. El valor de la fila indica la correlación con cada columna correspondiente.

Anexo J: Matriz correlación representada por números, pueba Lenguaje año 2011-2015.

Tabla 34
Matriz correlación representada por números, pueba Lenguaje año 2011-2015

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.00	0.17	0.03	0.02	0.58	0.45	-0.11	0.03	0.04	0.04	0.18	0.03	0.19	-0.06	0.18	-0.04	0.16
2	0.17	1.00	0.04	-0.01	0.18	0.21	0.01	0.04	0.10	0.05	0.44	-0.02	0.49	-0.28	0.47	-0.13	0.32
3	0.03	0.04	1.00	0.03	0.03	0.03	0.00	0.01	0.01	0.00	0.02	0.00	0.03	-0.02	0.03	-0.01	0.00
4	0.02	-0.01	0.03	1.00	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
5	0.58	0.18	0.03	0.03	1.00	0.52	-0.09	0.04	0.04	0.05	0.18	0.04	0.19	-0.06	0.19	-0.02	0.15
6	0.45	0.21	0.03	0.00	0.52	1.00	-0.10	0.07	0.10	0.05	0.20	0.03	0.21	-0.07	0.22	-0.03	0.16
7	-0.11	0.01	0.00	0.01	-0.09	-0.10	1.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
8	0.03	0.04	0.01	0.00	0.04	0.07	0.01	1.00	0.19	0.03	0.04	0.02	0.05	0.00	0.06	0.01	0.04
9	0.04	0.10	0.01	0.00	0.04	0.10	0.02	0.19	1.00	0.02	0.06	0.05	0.11	0.01	0.19	0.01	0.08
10	0.04	0.05	0.00	0.00	0.05	0.05	0.00	0.03	0.02	1.00	-0.25	0.26	0.09	0.00	0.17	0.04	0.06
11	0.18	0.44	0.00	0.00	0.18	0.20	0.00	0.03	0.06	-0.25	1.00	0.00	0.50	-0.23	0.47	-0.11	0.34
12	0.03	-0.02	0.00	0.00	0.03	0.03	1.00	0.01	0.02	0.00	0.50	1.00	-0.23	0.02	0.13	0.05	0.03
13	0.19	0.49	0.03	0.00	0.19	0.21	0.00	0.04	0.06	0.00	-0.23	-0.23	1.00	0.02	0.52	-0.10	0.34
14	-0.06	-0.28	-0.02	0.00	-0.06	-0.07	0.00	0.02	0.05	0.00	0.50	0.00	0.02	1.00	0.26	0.11	-0.13
15	0.18	0.47	0.03	0.00	0.19	0.22	0.00	0.05	0.11	0.09	0.50	-0.23	1.00	-0.26	0.52	0.11	-0.13
16	-0.04	-0.13	-0.01	0.00	-0.02	-0.03	0.00	0.06	0.19	0.17	-0.23	0.13	0.52	0.02	-0.57	1.00	-0.58
17	0.16	0.32	0.00	0.00	0.15	0.16	0.00	0.04	0.08	0.06	0.34	0.03	0.34	-0.13	0.34	-0.58	1.00

Nota. La tabla muestra las correlaciones entre variables. El valor de la fila indica la correlación con cada columna correspondiente.

Anexo K: Matriz correlación representada por números, pueba Matemática año 2011-2015.

Tabla 35
Matriz correlación representada por números, pueba Matemática año 2011-2015.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.00	0.28	0.03	0.02	0.50	0.50	0.08	0.05	0.08	0.08	0.27	0.05	0.29	-0.09	0.29	-0.04	0.22
2	0.28	1.00	0.04	-0.01	0.18	0.21	0.01	0.04	0.10	0.05	0.44	-0.02	0.49	-0.28	0.47	-0.13	0.32
3	0.03	0.04	1.00	0.03	0.03	0.03	0.00	0.01	0.01	0.00	0.02	0.00	0.03	-0.02	0.03	-0.01	0.00
4	0.02	-0.01	0.03	1.00	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
5	0.50	0.18	0.03	0.03	1.00	0.52	-0.09	0.04	0.04	0.05	0.18	0.04	0.19	-0.06	0.19	-0.02	0.15
6	0.50	0.21	0.03	0.00	0.52	1.00	-0.10	0.07	0.10	0.05	0.20	0.03	0.21	-0.07	0.22	-0.03	0.16
7	0.08	0.01	0.00	0.01	-0.09	-0.10	1.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
8	0.05	0.04	0.01	0.00	0.04	0.07	0.01	1.00	0.19	0.03	0.04	0.02	0.05	0.00	0.06	0.01	0.04
9	0.08	0.10	0.01	0.00	0.04	0.10	0.02	0.19	1.00	0.02	0.06	0.05	0.11	0.01	0.19	0.01	0.08
10	0.08	0.05	0.00	0.00	0.05	0.05	0.00	0.03	0.02	1.00	-0.25	0.26	0.09	0.00	0.17	0.04	0.06
11	0.27	0.44	0.02	0.00	0.18	0.20	0.00	0.04	0.06	-0.25	1.00	0.00	0.50	-0.23	0.47	-0.11	0.34
12	0.05	-0.02	0.00	0.01	0.04	0.03	0.00	0.02	0.05	0.26	0.00	1.00	-0.23	0.02	0.13	0.05	0.03
13	0.29	0.49	0.03	0.00	0.19	0.21	0.00	0.05	0.11	0.09	0.50	-0.23	1.00	-0.26	0.52	-0.10	0.34
14	-0.09	-0.28	-0.02	0.00	-0.06	-0.07	0.00	0.00	0.01	0.00	-0.23	0.02	-0.26	1.00	-0.57	0.11	-0.13
15	0.29	0.47	0.03	0.00	0.19	0.22	0.01	0.06	0.19	0.17	0.47	0.13	0.52	-0.57	1.00	-0.07	0.34
16	-0.04	-0.13	-0.01	0.00	-0.02	-0.03	0.00	0.01	0.01	0.04	-0.11	0.05	-0.10	0.11	-0.07	1.00	-0.58
17	0.22	0.32	0.00	0.00	0.15	0.16	0.00	0.04	0.08	0.06	0.34	0.03	0.34	-0.13	0.34	-0.58	1.00

Nota. La tabla muestra las correlaciones entre variables. El valor de la fila indica la correlación con cada columna correspondiente.

Referencias

- Anand P., Mizala A., y Repetto A. (2009). Using school scholarships to estimate the effect of private education on the academic achievement of low-income students in Chile. *Economics of Education Review*, 28(3):370–381. <https://doi.org/10.1016/j.econedurev.2008.03.005>.
- Becker G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226041223.001.0001>.
- Bellei C. (2013). Study of socioeconomic and academic segregation in Chilean education. *Comparative Education Review*, 57(2):181–215. <http://dx.doi.org/10.4067/S0718-07052013000100019>.
- Bronfenbrenner U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press. <https://doi.org/10.2307/j.ctv26071r6>.
- CymaSuite (s.f.). Sueldo mínimo en Chile. <https://www.cymasuite.com/sueldo-minimo-chile>. Consultado el 22 de junio de 2025.
- DEMRE (2024). *Sobre el DEMRE*. DEMRE. <https://demre.cl/el-demre/sobre-el-demre>.
- Elacqua G. (2012). The impact of school choice and public policy on segregation: Evidence from Chile. *International Journal of Educational Development*, 32(3):444–453. <https://doi.org/10.1016/j.ijedudev.2011.08.003>.

Gibbons S. y Silva O. (2011). Faith primary schools: Better schools or better pupils? *Journal of Labor Economics*, 29(3):589–638.

Gutiérrez G. y Carrasco A. (2021). Chile's enduring educational segregation: A trend unchanged by different cycles of reform. *British Educational Research Journal*, 47(6):1510–1532. <https://doi.org/10.1002/berj.3746>.

Hanushek E. A. y Rivkin S. G. (2006). Teacher Quality and Student Achievement. *Handbook of the Economics of Education*, 2:1051–1078. <https://hanushek.stanford.edu/sites/default/files/publications/Hanushek%2BRivkin%202006%20HbEEdu%202.pdf#:~:text=References%201075%20Handbook%20of%20the,6>.

Ladd H. F. y Walsh R. P. (2002). Implementing Value-Added Measures of School Effectiveness: Getting the Incentives Right. *Economics of Education Review*, 21(1):1–17. <http://randallwalsh.com/wp-content/uploads/2017/08/implementing-value-added-measures-of-school-effectiveness.pdf>.

Ley N°18956 (1990). Diario Oficial de la República de Chile. Santiago de Chile, 11 de febrero de 1990.

Ley N°20.248 (2008). Diario Oficial de la República de Chile. Santiago de Chile, 27 de agosto de 2008.

Ley N°20845 (2015). Diario Oficial de la República de Chile. Santiago de Chile, 8 de junio de 2015.

Ley N°20.903 (2016). Diario Oficial de la República de Chile. Santiago de Chile, 1 de abril de 2016.

Ministerio de Educación (s.f.). SIMCE. <https://www.ayudamineduc.cl/ficha/simce>. Consultado el 14 de octubre de 2024.

OECD (2021). Education Policy Outlook Chile. *OECD Publishing*. <https://doi.org/10.1787/d625ec2b-en>.

Quaresma M. L. (2017). Excellence in high-performing public schools in Chile: Students' perceptions and experiences. *University of Chicago Press, Chicago, IL*, 14(1):28–53. <https://doi.org/10.1086/691250>.

Reardon S. F. (2011). The Widening Academic Achievement Gap Between the Rich and the Poor. *Educational Policy Analysis Archives*, 19(1):1–28. <https://doi.org/10.14507/epaa.v19n1.2011>.

Rivkin S. G., Hanushek E. A., y Kain J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>.

Santiago P. y Others (2017). Education in Chile: Reviews of National Policies for Education. *OECD Publishing*. <https://doi.org/10.1787/9789264284420-en>.

Thieme C., Prior D., Tortosa-Ausina E., y Gempp R. (2016). Value added, educational accountability approaches and their effects on schools' rankings: Evidence from Chile. *European Journal of Operational Research*, 253(2):456–471. <https://doi.org/10.1016/j.ejor.2016.01.023>.

Torche F. (2005). Privatization Reform and Inequality of Educational Opportunity: The Ca-



se of Chile. *Sociology of Education*, 78(4):316–343. <https://doi.org/10.1177/003804070507800403>.