

**UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA**  
**DEPARTAMENTO DE INDUSTRIAS**

**Efectos del Sistema de Admisión Escolar (SAE) en los Resultados  
Académicos de los Liceos Bicentenarios en Chile: Un Estudio Comparativo  
Pre y Post SAE.**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL**

**Autor**

**Cristóbal Díaz Pinto**

**Profesor Guía**

**Bernardo Pincheira**

**Profesor Correferente**

**Rodolfo Salazar**

**VALPARAÍSO DE CHILE, Abril 2026**



## 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  
Efectos del Sistema de Admisión Escolar (SAE) en los Resultados Académicos de los Liceo Bicentenario en Chile:  
**Título del trabajo:** Un Estudio Comparativo Pre y Post SAE.

**Nombre del candidato(a):** Cristóbal Leonidas Díaz Pinto

**Carrera / Grado:** Ingeniería Civil Industrial

**Campus:** Casa Central **Departamento:** Industrias

### 2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Bernardo Pincheira Sarmiento, 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 (marcar una opción)

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 (**marcar una opción**):

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: 30 de abril Firma: 

**Estudiante o Candidato(a):**

Fecha: 30 de abril 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.*

## Resumen

Esta memoria analiza el efecto del Sistema de Admisión Escolar (SAE) en los resultados académicos de los Liceos Bicentenario en Chile, en un contexto caracterizado por alta segregación escolar y por la tensión entre equidad e inclusión, por un lado, y excelencia académica, por otro. El marco teórico se apoya en la estructura del sistema escolar chileno como cuasimercado, en la tradición selectiva de los liceos de excelencia, en el concepto de efecto par y en la distinción entre rendimiento bruto y valor añadido. Bajo este enfoque, el SAE es entendido como una reforma que reemplaza la selección académica por un mecanismo de asignación aleatoria, aumentando la heterogeneidad del alumnado y modificando las condiciones bajo las cuales operaban los Liceos Bicentenario.

Metodológicamente, la investigación utiliza datos SIMCE y cuestionarios de apoderados para 2016 y 2022, integrando información de estudiantes, hogares y establecimientos. El análisis combina pruebas t descriptivas, un modelo de efectos fijos de dos vías como estrategia principal de inferencia causal, un modelo multinivel de efectos mixtos para capturar la estructura jerárquica de los datos y el Coeficiente de Correlación Intraclase (ICC) para medir segregación escolar.

Los resultados muestran que los Liceos Bicentenario registran una fuerte caída descriptiva en sus puntajes promedio entre 2016 y 2022, especialmente en Matemática y Lectura, junto con cambios en la composición de su alumnado. A nivel econométrico, no se encuentra evidencia robusta de un efecto diferencial en Matemática en el modelo agregado una vez controlada la composición, pero sí una caída significativa en Lectura; el modelo multinivel, en cambio, detecta efectos negativos en ambas asignaturas, más intensos en Lectura. Además, la disminución del ICC sugiere una reducción parcial de la segregación entre establecimientos. En conclusión, el SAE se



asocia con una matrícula Bicentenario más diversa, una pérdida de ventaja académica relativa y una menor concentración social entre escuelas, lo que indica avances en inclusión, pero también nuevas tensiones para sostener la excelencia en un sistema no selectivo.



## Tabla de Contenidos

<b>1. Problema de Investigación</b>	<b>10</b>
<b>2. Objetivos</b>	<b>13</b>
2.1. Objetivo General . . . . .	13
2.2. Objetivos Específicos . . . . .	13
<b>3. Marco Teórico</b>	<b>14</b>
3.1. Estructura del Sistema Escolar Chileno: El Modelo de Cuasimercado . . . . .	14
3.2. Establecimientos Públicos de Excelencia y Meritocracia . . . . .	15
3.2.1. Liceos Emblemáticos: La Tradición de la Selección . . . . .	15
3.2.2. Referentes Internacionales de Excelencia Pública . . . . .	16
3.2.3. Los Liceos Bicentenario de Excelencia (LBE) . . . . .	16
3.3. Medición de Calidad y Resultados Históricos: El Rol del SIMCE . . . . .	17
3.3.1. Evolución de los Resultados Nacionales y el Efecto Pandemia . . . . .	17
3.3.2. Brechas por Dependencia Administrativa y el “Peer Effect” . . . . .	18
3.3.3. El Desempeño de los Liceos Bicentenario bajo Selección (Pre-SAE) . . . . .	19
3.3.4. El Concepto de Valor Añadido vs. Rendimiento Bruto . . . . .	20
3.3.5. La Problemática del Rendimiento Bruto Post-Reforma . . . . .	20
3.4. La Reforma de 2015 y el Sistema de Admisión Escolar (SAE) . . . . .	20
3.4.1. Diseño de Mercados y el Algoritmo de Aceptación Diferida (DA) . . . . .	21
3.4.2. De la Selección a la Aleatoriedad: “La Tómbola” . . . . .	21
3.5. Estrategia de Identificación Causal: Diferencias en Diferencias . . . . .	22



3.5.1.	Supuesto de Tendencias Paralelas . . . . .	23
3.6.	Heterogeneidad Inobservable y la Estimación mediante Efectos Fijos . . . . .	23
3.7.	Estructura Jerárquica de los Datos Educativos y Modelación Multinivel . . . . .	24
3.7.1.	El Coeficiente de Correlación Intraclase (ICC) como Medida de Segregación	25
<b>4.</b>	<b>Descripción de los Datos</b>	<b>26</b>
4.1.	Fuente . . . . .	26
4.2.	Tablas . . . . .	27
4.3.	Puntajes Promedio Matemática y Lectura . . . . .	28
<b>5.</b>	<b>Integración de los Datos</b>	<b>30</b>
5.1.	Manejo de los Datos . . . . .	30
5.1.1.	Depuración de los Datos . . . . .	31
5.1.2.	Carga de los Datos . . . . .	34
5.1.3.	Repetición para el año 2022 . . . . .	34
5.2.	Creación y Transformación de Variables . . . . .	34
5.3.	Panel de Datos final . . . . .	35
5.3.1.	Panel General . . . . .	35
5.3.2.	Panel de Escuelas . . . . .	36
<b>6.</b>	<b>Metodología</b>	<b>38</b>
6.1.	Análisis Preliminar: Comparación de Medias (T-Tests) . . . . .	38
6.1.1.	Procedimiento del análisis . . . . .	39
6.1.2.	Criterios de Significancia . . . . .	41



6.1.3.	Alcance interpretativo del análisis preliminar . . . . .	42
6.2.	Modelo de Inferencia Causal: Efectos Fijos. . . . .	43
6.2.1.	Lógica del Tratamiento y Adopción Escalonada . . . . .	44
6.2.2.	Especificación Econométrica del Modelo . . . . .	45
6.2.3.	Control de Composición y el Desafío del Valor Añadido . . . . .	46
6.2.4.	Limitaciones del Modelo de Efectos Fijos de Dos Vías . . . . .	47
6.3.	Modelo Multinivel de Efectos Mixtos para el Análisis de Composición y Segrega- ción Escolar . . . . .	49
6.3.1.	Fundamentación Metodológica del Modelo . . . . .	49
6.3.2.	Estructura Jerárquica de los Datos . . . . .	50
6.3.3.	Especificación Econométrica . . . . .	50
6.3.4.	Interpretación del Coeficiente de Interés . . . . .	52
6.3.5.	Aplicación del Modelo a los Objetivos Específicos . . . . .	53
6.3.6.	Medición de Segregación mediante el Coeficiente de Correlación Intraclase	53
6.3.7.	Ventajas del Modelo Multinivel . . . . .	54
6.3.8.	Alcances y Limitaciones del Modelo . . . . .	55
6.4.	Limitaciones de Identificación y Pruebas de Robustez . . . . .	56
<b>7.</b>	<b>Resultados</b>	<b>58</b>
7.1.	Resultados preliminares de las Pruebas-T de tipo Longitudinal . . . . .	58
7.1.1.	Bicentenarios 2016 vs. 2022 . . . . .	58
7.1.2.	Municipales 2016 vs. 2022 . . . . .	59
7.1.3.	Subvencionado 2016 vs. 2022 . . . . .	60



7.1.4.	Privado 2016 vs. 2022 . . . . .	61
7.2.	Resultados preliminares de las Pruebas-T de tipo Transversal . . . . .	62
7.2.1.	Bicentenarios vs. Municipales . . . . .	62
7.2.2.	Bicentenarios vs. Subvencionados . . . . .	63
7.2.3.	Bicentenarios vs. Privados . . . . .	65
7.3.	Resultados del Modelo de Efectos Fijos de Dos Vías . . . . .	66
7.3.1.	Resultados para Matemática . . . . .	66
7.3.2.	Resultados para Lectura . . . . .	67
7.3.3.	Controles de composición y asociaciones relevantes . . . . .	68
7.4.	Resultados del Modelo Multinivel de Efectos Mixtos . . . . .	70
7.4.1.	Resultados académicos del modelo multinivel . . . . .	70
7.4.2.	Asociaciones relevantes del modelo multinivel académico . . . . .	72
7.4.3.	Resultados sobre composición del alumnado . . . . .	73
7.5.	Resultados de segregación escolar: Coeficiente de Correlación Intraclase . . . . .	75
<b>8.</b>	<b>Conclusiones</b>	<b>77</b>
8.1.	Conclusiones sobre el cambio en la composición del alumnado . . . . .	77
8.2.	Conclusiones sobre los resultados descriptivos preliminares . . . . .	78
8.3.	Conclusiones del modelo de efectos fijos de dos vías . . . . .	80
8.4.	Conclusiones del modelo multinivel de efectos mixtos . . . . .	81
8.5.	Conclusiones sobre segregación escolar . . . . .	83
8.6.	Conclusiones respecto de las preguntas de investigación . . . . .	85
8.7.	Comparación de resultados con las expectativas previas . . . . .	86



8.8. Conclusiones generales . . . . .	87
<b>9. Limitaciones</b>	<b>89</b>
9.1. Limitaciones del Diseño Temporal . . . . .	89
9.2. Limitaciones de Identificación Causal . . . . .	90
9.3. Limitaciones de la Base de Datos y de la Construcción Muestral . . . . .	91
9.4. Limitaciones asociadas a la Transformación de Variables . . . . .	91
9.5. Limitaciones de Operacionalización Conceptual . . . . .	93
9.6. Limitaciones del Modelo de Efectos Fijos de Dos Vías . . . . .	93
9.7. Limitaciones del Modelo Multinivel . . . . .	94
9.8. Cierre del Capítulo . . . . .	95
<b>Anexos</b>	<b>97</b>



## Lista de Tablas

1.	Tabla con los promedios y número de observaciones de las pruebas para los años 2016 y 2022. . . . .	17
2.	Tabla descriptiva del archivo _rbd. . . . .	27
3.	Tabla descriptiva del archivo _cpad. . . . .	27
4.	Tabla descriptiva del archivo _alu. . . . .	28
5.	Tabla con las estadísticas descriptivas de las pruebas para 2016 y 2022. . . . .	29
6.	Análisis longitudinal descriptivo: Bicentenario 2016 vs. 2022. . . . .	58
7.	Análisis longitudinal descriptivo: Municipal 2016 vs. 2022. . . . .	59
8.	Análisis longitudinal descriptivo: Subvencionado 2016 vs 2022. . . . .	60
9.	Análisis longitudinal descriptivo: Privado 2016 vs. 2022. . . . .	61
10.	Análisis descriptivo comparativo: Municipal (2016). . . . .	62
11.	Análisis descriptivo comparativo: Municipal (2022). . . . .	62
12.	Análisis descriptivo comparativo: Subvencionado (2016). . . . .	63
13.	Análisis descriptivo comparativo: Subvencionado (2022). . . . .	64
14.	Análisis descriptivo comparativo: Privado (2016). . . . .	65
15.	Análisis descriptivo comparativo: Privado (2022). . . . .	65
16.	Resultados del modelo de efectos fijos de dos vías: Matemática. . . . .	66
17.	Resultados del modelo de efectos fijos de dos vías: Lectura. . . . .	67
18.	Coeficientes relevantes de los controles en el modelo de efectos fijos de dos vías. . . . .	68
19.	Resultados académicos del modelo multinivel de efectos mixtos. . . . .	70
20.	Coeficientes relevantes del modelo multinivel para resultados académicos. . . . .	72



21.	Resultados del modelo multinivel para composición del alumnado. . . . .	73
22.	Coefficiente de Correlación Intraclase (ICC) por variable y año. . . . .	75
23.	Criterios de imputación estocástica para la variable Ingreso Familiar. . . . .	97
24.	Criterios de conversión e imputación estocástica para los años de escolaridad de los padres. . . . .	97
25.	Matriz de Correlaciones completa para el Modelo de Efectos Fijos de Dos Vías (Modelo 6.2). . . . .	98
26.	Matriz de Correlaciones completa para el Modelo Multinivel de Efectos Mixtos (Modelo 6.3). . . . .	99

## Lista de Figuras

1.	Mapa de Correlación para el Modelo de Efectos Fijos . . . . .	69
2.	Mapa de Correlación para el Modelo Multinivel de Efectos Mixtos . . . . .	74

## 1. Problema de Investigación

El sistema educativo chileno ha experimentado transformaciones profundas durante las últimas décadas, impulsadas por la necesidad de reducir la segmentación estructural que históricamente ha condicionado el acceso a una educación de calidad. Diversos estudios han demostrado que Chile presenta uno de los niveles más altos de segregación escolar socioeconómica dentro de la Organización para la Cooperación y el Desarrollo Económico (OCDE (2017); Bellei (2013)), generando desigualdades sustanciales en las trayectorias educativas y limitando la movilidad social ascendente.

En este escenario, dos políticas públicas han sido clave y a su vez, han marcado la discusión educativa reciente en Chile: la creación de los Liceos Bicentenarios en 2010, orientados a fortalecer la educación pública a través del fomento del alto rendimiento académico y la disciplina escolar; y la implementación del Sistema de Admisión Escolar (SAE) a partir del año 2016, cuyo objetivo es democratizar el acceso a establecimientos educacionales mediante un sistema de asignación aleatoria ponderada. Ambas responden a problemáticas distintas, pero convergen en un punto crítico: el acceso a establecimientos de calidad en el sistema escolar público.

Los Liceos Bicentenarios fueron creados con el objetivo de ser instituciones de excelencia dentro del sistema escolar público, siendo seleccionados y financiados estratégicamente para ofrecer una educación de alta calidad. En su etapa inicial, estos establecimientos mostraron mejoras significativas en pruebas como SIMCE y PSU, y lograron tasas de egreso superiores al promedio nacional (Villar (2020)). No obstante, sus buenos resultados estuvieron estrechamente asociados a procesos selectivos de admisión, que favorecían la homogeneidad académica y socioeconómica de sus estudiantes.

La implementación del SAE, regulada por la Ley de Inclusión Escolar (Ley N.º 20.845), prohibió la selección por mérito académico en todos los establecimientos públicos y subvencionados, obligando a los Liceos Bicentenarios a modificar sus mecanismos de ingreso. Este cambio estructural eliminó filtros previos como entrevistas, pruebas y promedios, reemplazándolos por un algoritmo aleatorizado que prioriza criterios como la cercanía geográfica y la existencia de hermanos matriculados. Este sistema, conocido popularmente como “la tómbola”, busca garantizar que todos los estudiantes tengan la misma posibilidad de acceder a un establecimiento educacional, sin discriminación por desempeño previo ni nivel socioeconómico.

La implementación del SAE generó un cambio estructural en la forma en que los estudiantes acceden a los Liceos Bicentenarios. Antes del SAE, estos liceos aplicaban mecanismos de selección —como pruebas, entrevistas o exigencias de promedio— para admitir a estudiantes, lo que les permitía conformar comunidades académicamente homogéneas y con altos niveles de rendimiento.

Un número significativo de Liceos Bicentenarios reportan dificultades —no cuantificadas— para sostener su rendimiento tras el SAE (Carrasco et al. (2024)), debido al aumento en la heterogeneidad del alumnado. Además, algunos informes técnicos y estudios regionales (Sillard et al. (2018)) han mostrado que, si bien el SAE ha promovido una mayor inclusión, también ha generado desafíos en la gestión académica y en la percepción de calidad por parte de las comunidades educativas. Esto ha dado paso a preguntas fundamentales que todavía no han sido suficientemente respondidas por la evidencia empírica:

¿Cómo ha cambiado el perfil de los estudiantes que ingresan a los Liceos Bicentenarios tras la implementación del SAE?

¿Existe una relación entre el ingreso aleatorio de estudiantes y una eventual disminución

—o mantenimiento— de los resultados académicos institucionales?

¿Ha logrado el SAE disminuir la segregación escolar dentro de los Liceos Bicentenarios o solo ha redistribuido el problema?

¿Es posible mantener estándares de excelencia académica en un sistema de ingreso no selectivo?

Estas preguntas reflejan el tenso equilibrio entre equidad y calidad que atraviesa la educación pública en Chile. Por una parte, el SAE representa un avance en términos de justicia social y democratización del acceso. Por otra, pone a prueba la capacidad de los establecimientos de alto rendimiento para mantener sus resultados sin seleccionar a su alumnado.

En este contexto, resulta fundamental realizar una evaluación empírica del impacto del Sistema de Admisión Escolar en el desempeño académico de los Liceos Bicentenarios, considerando tanto indicadores cuantitativos como las percepciones de los actores involucrados. Comprender esta dinámica permitirá aportar evidencia al debate educativo, y contribuir a un modelo que combine inclusión con altos estándares pedagógicos.

Además, esta investigación se vincula con una discusión más amplia sobre el derecho a una educación de calidad en contextos de creciente diversidad. La literatura internacional ha demostrado que las políticas de asignación escolar no son neutras: configuran oportunidades educativas y condiciones de aprendizaje (Lubienski et al. (2009); Hoxby (2000)). Por tanto, es clave estudiar no solo los resultados, sino también los mecanismos intermedios a través de los cuales el SAE afecta la organización escolar, las prácticas pedagógicas y la relación entre comunidades escolares.

## 2. Objetivos

### 2.1. Objetivo General

Evaluar la influencia del Sistema de Admisión Escolar (SAE) en los resultados académicos de los Liceos Bicentenarios en Chile, analizando el desempeño institucional antes y después de su implementación, con el propósito de comprender cómo esta política de acceso influye en la calidad educativa de dichos establecimientos.

### 2.2. Objetivos Específicos

Comparar los resultados SIMCE de segundo medio de los Liceos Bicentenarios antes y después de la implementación del SAE, con el fin de identificar posibles variaciones en el rendimiento institucional atribuibles al cambio en el sistema de admisión, y evaluar su impacto en la calidad educativa.

Analizar los cambios en el perfil socioeconómico y académico de los estudiantes admitidos en los Liceos Bicentenarios tras la implementación del SAE, para comprender cómo ha influido esta política en la composición del estudiantado.

Estimar el efecto causal de la introducción del Sistema de Admisión Escolar (SAE) en la variación de los resultados académicos de los Liceos Bicentenarios, controlando por la nueva composición del estudiantado, esto con el objetivo de determinar si el método de ingreso afecta directamente los indicadores de desempeño académico y la gestión escolar.

Determinar si la implementación del SAE ha influido en los niveles de segregación escolar en los Liceos Bicentenarios, a fin de establecer si esta política ha contribuido a una mayor diversidad social y académica en estos establecimientos, promoviendo entornos más inclusivos.

### 3. Marco Teórico

El presente capítulo constituye la base conceptual y el soporte empírico que enmarcan la problemática central de esta investigación: el equilibrio entre calidad y equidad en la educación pública chilena. Para comprender el impacto del Sistema de Admisión Escolar (SAE) en los Liceos Bicentenario de Excelencia (LBE), es imperativo realizar un recorrido desde la arquitectura macro del sistema escolar hasta los mecanismos algorítmicos específicos que determinan la asignación de estudiantes.

#### 3.1. Estructura del Sistema Escolar Chileno: El Modelo de Cuasimercado

El sistema educativo chileno ha experimentado transformaciones profundas, impulsadas históricamente por la necesidad de gestionar el acceso a una educación de calidad en un contexto de alta segmentación. Desde la reforma de la década de 1980, Chile consolidó un modelo de “cuasimercado” educativo, caracterizado por el financiamiento a la demanda mediante vouchers y la libertad de elección de los padres (Hsieh and Urquiola (2006)).

Esta arquitectura se sostiene sobre tres pilares de administración:

1. **Educación Municipal / Servicios Locales de Educación Pública (SLEP):** Representan la oferta estatal tradicional. Históricamente, han atendido a la población más vulnerable y han sido los más expuestos a la segregación estructural del sistema.
2. **Educación Particular Subvencionada (Subvencionados):** Instituciones de propiedad privada pero financiadas mayoritariamente con fondos públicos por alumno. Este sector permitió una expansión masiva de la matrícula, pero también facilitó la selección arbitraria de estudiantes hasta la reforma de 2015.

**3. Educación Particular Pagada (Particulares):** Establecimientos financiados exclusivamente por las familias de los estudiantes, que operan fuera del sistema de subvención estatal y concentran a la élite socioeconómica, cuentan con la libertad absoluta para la selección de estudiantes pero igual se rigen por el Ministerio de Educación.

Diversos estudios demuestran que Chile presenta uno de los niveles más altos de segregación escolar socioeconómica dentro de la OCDE, lo que limita la movilidad social ascendente y genera trayectorias educativas desiguales basadas en el origen del estudiante (Valenzuela et al. (2010); González et al. (2017)).

### **3.2. Establecimientos Públicos de Excelencia y Meritocracia**

La noción de “excelencia” en la educación pública chilena no es un concepto estático, sino que ha evolucionado desde la tradición republicana hacia políticas de intervención focalizada.

#### **3.2.1. Liceos Emblemáticos: La Tradición de la Selección**

Antes de la creación de los LBE, el estandarte de la máxima calidad pública residía en los “Liceos Emblemáticos” como el Instituto Nacional o el Liceo 1, por mencionar algunos. Estas instituciones basaban su prestigio en procesos de admisión altamente competitivos y meritocráticos. Su rol histórico fue servir como motores de movilidad social para estudiantes talentosos de estratos medios y bajos, aunque a costa de una alta segregación académica respecto al resto del sistema municipal.

### 3.2.2. Referentes Internacionales de Excelencia Pública

El debate sobre escuelas públicas selectivas es global. Dos modelos relevantes para contrastar con la realidad chilena son:

- **Magnet Schools (Estados Unidos):** Escuelas públicas con currículos especializados (STEM, Artes) que atraen a estudiantes de diversos distritos. Aunque buscan la desegregación, a menudo utilizan criterios de admisión basados en el talento.
- **Grammar Schools (Reino Unido):** Escuelas estatales selectivas que admiten alumnos basados en el examen “11-plus”. Al igual que los liceos emblemáticos chilenos, son defendidas por su capacidad de potenciar el talento académico, pero criticadas por profundizar la brecha de clases.

### 3.2.3. Los Liceos Bicentenario de Excelencia (LBE)

La política de los LBE se inicia en 2010 como una estrategia del Ministerio de Educación (MINEDUC) para fortalecer la educación pública mediante el fomento del alto rendimiento y la disciplina. Su objetivo fundamental fue la promoción de la movilidad social concentrando recursos en establecimientos bajo un modelo de gestión basado en la exigencia académica y un liderazgo directivo fuerte. Hasta antes de 2015, los LBE utilizaban pruebas de selección y entrevistas para conformar comunidades académicamente homogéneas, lo que facilitaba la gestión pedagógica y minimizaba los desafíos de la diversidad en el aula (Gómez, 2014).

### 3.3. Medición de Calidad y Resultados Históricos: El Rol del SIMCE

Para comprender el impacto del Sistema de Admisión Escolar (SAE) sobre los Liceos Bicentenario de Excelencia (LBE), es fundamental analizar la trayectoria de los resultados de aprendizaje en Chile. El Sistema de Medición de la Calidad de la Educación (SIMCE) actúa como el principal censo de aprendizaje a nivel nacional, evaluando áreas críticas como Lectura y Matemáticas. No obstante, estos resultados han estado históricamente condicionados por la segmentación del sistema (Eyzaguirre and Fontaine (1999); Himmel (1997)).

#### 3.3.1. Evolución de los Resultados Nacionales y el Efecto Pandemia

Los resultados históricos en Chile muestran una estabilidad relativa con brechas persistentes, pero el periodo comprendido entre 2016 y 2022 revela una inflexión crítica. Mientras que el año 2016 representa un periodo de consolidación de políticas de mejora educativa pre-reforma total, el año 2022 captura un escenario complejo: el retorno a la presencialidad post-pandemia y la plena implementación del SAE en la Región Metropolitana y el resto del país. +4

De acuerdo con las estadísticas descriptivas de este estudio, se observa una variación significativa en los puntajes brutos nacionales entre ambos periodos:

Prueba	Año	Promedio	N° de Obs.
Matemática	2016	267.72	162766
Matemática	2022	253.24	159248
Lectura	2016	249.35	161448
Lectura	2022	244.58	157446

Tabla 1: Tabla con los promedios y número de observaciones de las pruebas para los años 2016 y 2022.

Esta caída, especialmente pronunciada en Matemáticas de aproximadamente 14 puntos, no puede atribuirse a un solo factor, pero establece una fuerte base sobre la cual realizar otro tipo de

análisis.

### 3.3.2. Brechas por Dependencia Administrativa y el “Peer Effect”

La literatura económica ha documentado que los resultados de aprendizaje están fuertemente condicionados por la composición del entorno educativo. En este contexto, surge el concepto del *Peer Effect* (Efecto Par), el cual postula que el rendimiento de un estudiante se ve influenciado por las características sociodemográficas y académicas de sus compañeros. Formalmente, este fenómeno se puede modelar mediante una función de producción educativa de la siguiente forma:

$$Y_{it} = \alpha + \beta X_{it} + \gamma \bar{X}_{-i,t} + \epsilon_{it} \quad (1)$$

Donde:

- $Y_{it}$  representa el resultado académico (puntaje SIMCE) del estudiante  $i$  en el periodo  $t$ .
- $X_{it}$  es un vector de características individuales y familiares del estudiante (ej. nivel socioeconómico, escolaridad de los padres).
- $\bar{X}_{-i,t}$  representa el promedio de las características de sus compañeros de nivel en el establecimiento, excluyendo al estudiante  $i$ .
- $\gamma$  es el coeficiente que captura la magnitud del efecto par; un valor positivo indica que interactuar con pares de mayor capital cultural potencia el rendimiento individual.
- $\epsilon_{it}$  es el término de error aleatorio.

Históricamente, los Liceos Bicentenario de Excelencia (LBE) maximizaban el componente  $\gamma \bar{X}_{-i,t}$  mediante procesos de selección que favorecerían la homogeneidad académica y socioeconómica.

mica. Al concentrar estudiantes con altas habilidades previas, se minimizaba la dispersión en el aula, facilitando la gestión pedagógica.

Con la implementación del SAE, la variable  $\bar{X}_{-i,t}$  se vuelve exógena al liceo, aumentando la heterogeneidad del alumnado. El desafío de esta investigación consiste en determinar si la gestión institucional de los LBE es capaz de compensar una posible disminución en el valor de los atributos del grupo de pares tras la eliminación de la selección.

### 3.3.3. El Desempeño de los Liceos Bicentenario bajo Selección (Pre-SAE)

Desde su creación en 2010, los LBE mostraron una trayectoria de éxito innegable en términos de resultados de salida. En sus primeros años, lograron puntajes promedio en SIMCE y pruebas de acceso a la universidad significativamente superiores al promedio nacional y a liceos de similares condiciones socioeconómicas. Por lo que desde el momento de su creación hasta antes de la implementación del SAE, se podía desatacar lo siguiente:

- **Resultados Académicos:** Tasas de egreso superiores al promedio nacional y mejoras sustanciales en pruebas estandarizadas
- **Gestión y Disciplina:** Un modelo basado en la exigencia y liderazgos directivos fuertes.
- **Sesgo de Selección:** Es imperativo notar que estos resultados se daban en un contexto donde los liceos podían filtrar a su alumnado, conformando comunidades académicamente homogéneas.

### **3.3.4. El Concepto de Valor Añadido vs. Rendimiento Bruto**

En la economía de la educación, es crucial distinguir entre el rendimiento bruto (el puntaje final) y el valor añadido (cuánto aprendió el alumno gracias a la escuela). Las métricas de rendimiento bruto suelen estar sesgadas por el Efecto Par, es decir, si un liceo selecciona solo a los mejores estudiantes, sus altos puntajes pueden deberse más al capital cultural previo de los alumnos que a la calidad de la enseñanza.

### **3.3.5. La Problemática del Rendimiento Bruto Post-Reforma**

Con la implementación de la Ley de Inclusión y el SAE, el mecanismo de entrada a los LBE cambió de un criterio de mérito a uno de aleatoriedad ponderada. Esto genera un aumento en la heterogeneidad del alumnado.

Desde una perspectiva de ingeniería industrial aplicada a políticas públicas, esto altera los inputs del proceso educativo. Si un LBE mantiene sus resultados a pesar de recibir una matrícula más diversa y con menores niveles de preparación inicial, su valor añadido es, teóricamente, mayor que en el periodo de selección. Por el contrario, si los resultados caen, se abre el debate sobre si el modelo pedagógico de "excelencia" era realmente efectivo por su gestión o si solo era un subproducto de la selección arbitraria de talentos.

## **3.4. La Reforma de 2015 y el Sistema de Admisión Escolar (SAE)**

La Ley N.º 20.845 de Inclusión Escolar marcó el fin de la selección por mérito académico en todos los establecimientos que reciben fondos del Estado. El objetivo fue reducir la segmentación y garantizar que el acceso a escuelas de calidad no dependiera de la capacidad de pago o del

desempeño previo (Villaseca (2022); Carrasco and Honey (2019)).

### 3.4.1. Diseño de Mercados y el Algoritmo de Aceptación Diferida (DA)

El SAE es una plataforma centralizada que utiliza un algoritmo de Diseño de Mercados para asignar cupos. Este algoritmo (Roth and Sönmez (2005); Angrist and Pischke (2009)), una variante del modelo de Gale and Shapley (1962), posee propiedades matemáticas fundamentales para la equidad:

1. **Estabilidad:** Asegura que no existan incentivos para que un alumno y una escuela prefieran emparejarse fuera del sistema.
2. **No Manipulabilidad (Strategy-proofness):** Esta es la propiedad técnica más relevante. Garantiza que la mejor estrategia para los padres sea declarar sus preferencias reales. Esto elimina la ventaja que tenían antes las familias con más información o capital cultural (Abdulkadiroğlu et al. (2009); Correa et al. (2022)).

### 3.4.2. De la Selección a la Aleatoriedad: “La Tómbola”

Cuando un establecimiento como un LBE tiene más demanda que cupos, el SAE aplica criterios de prioridad (hermanos, alumnos prioritarios) y, como último recurso, la aleatoriedad. Este cambio estructural obligó a los LBE a modificar sus mecanismos de ingreso, reemplazando las pruebas por un algoritmo que prioriza la inclusión, lo que genera un aumento en la heterogeneidad socio-académica de la matrícula (Sillard et al. (2018); Padilla et al. (2023)).

### 3.5. Estrategia de Identificación Causal: Diferencias en Diferencias

Para aislar el efecto causal de la implementación del SAE sobre el desempeño de los Liceos Bicentenario, se utiliza un modelo de Diferencias en Diferencias (DiD) con datos de panel a nivel de estudiante (Angrist and Pischke, 2009). Esta técnica permite controlar por factores no observados que son constantes en el tiempo y por tendencias temporales comunes a todos los establecimientos (Angrist and Pischke, 2009).

La especificación econométrica principal se define como:

$$Y_{ist} = \alpha + \beta_1(LBE_s) + \beta_2(Post_t) + \delta(LBE_s \times Post_t) + \lambda X_{ist} + \epsilon_{ist} \quad (2)$$

Donde:

- $Y_{ist}$  es el puntaje SIMCE del estudiante  $i$ , en el establecimiento  $s$ , en el año  $t$ .
- $LBE_s$  es una variable dicotómica que toma el valor de 1 si el establecimiento es un Liceo Bicentenario de Excelencia y 0 en caso contrario (grupo de control).
- $Post_t$  es una variable dicotómica para el periodo post-reforma, que toma el valor de 1 para el año 2022 y 0 para el año base 2016.
- $LBE_s \times Post_t$  es el término de interacción que representa el efecto del tratamiento.
- El coeficiente  $\delta$  es el estimador de Diferencias en Diferencias, que captura el impacto neto del SAE sobre los resultados de los LBE.
- $X_{ist}$  es un vector de controles a nivel de estudiante y establecimiento (ej. índice socioeconómico, dependencia administrativa, región) que busca reducir la varianza del error y corregir

posibles sesgos de selección.

- $\epsilon_{ist}$  es el término de error estocástico.

### 3.5.1. Supuesto de Tendencias Paralelas

La validez del estimador  $\delta$  descansa sobre el supuesto de que, en ausencia de la reforma, la brecha de rendimiento entre los LBE y el grupo de control se habría mantenido constante en el tiempo (Angrist and Pischke, 2009). Dado que los LBE fueron seleccionados precisamente por su potencial de mejora, el análisis debe ser cauteloso en controlar por variables preexistentes que aseguren que el grupo de control es un contrafactual válido para la trayectoria de los Liceos Bicentenario.

## 3.6. Heterogeneidad Inobservable y la Estimación mediante Efectos Fijos

La estrategia de Diferencias en Diferencias (DiD) proporciona un marco lógico para la identificación causal, pero su estimación empírica en el ámbito educativo enfrenta el desafío de la heterogeneidad inobservable. Los establecimientos escolares no son unidades idénticas que solo difieren en su condición de Bicentenario; poseen características estructurales, culturales y territoriales que son constantes en el tiempo y que inciden directamente en su desempeño histórico.

En la literatura de economía de la educación, se argumenta que ignorar estas características invariables (como la reputación acumulada de un liceo, su infraestructura o el capital social de su ubicación geográfica) genera sesgos de variable omitida. Para corregir esto, la estimación del DiD suele operacionalizarse a través de un Modelo de Efectos Fijos de Dos Vías (Two-Way Fixed Effects, TWFE) (Angrist and Pischke, 2009). Teóricamente, la inclusión de efectos fijos a nivel

de establecimiento permite absorber toda la heterogeneidad inobservable constante en el tiempo, mientras que los efectos fijos temporales capturan shocks sistémicos que afectan a todas las escuelas por igual, como el impacto generalizado de la pandemia de COVID-19 y el posterior retorno a la presencialidad. De esta forma, el parámetro causal se identifica aislando exclusivamente la variación intra-establecimiento a través del tiempo.

### **3.7. Estructura Jerárquica de los Datos Educativos y Modelación Multinivel**

A pesar de la robustez de los modelos de efectos fijos agregados, la investigación educativa se caracteriza por trabajar con datos que poseen una naturaleza intrínsecamente jerárquica o anidada. Los estudiantes no son unidades de observación aisladas; se encuentran agrupados dentro de aulas, que a su vez pertenecen a establecimientos educacionales.

El supuesto clásico de independencia de las observaciones, fundamental en los modelos de regresión lineal estándar (Mínimos Cuadrados Ordinarios), se viola sistemáticamente en contextos escolares. Dos estudiantes matriculados en un mismo Liceo Bicentenario comparten el mismo equipo directivo, infraestructura, clima escolar y un entorno socioeconómico de pares. Por consiguiente, sus resultados académicos y trayectorias están correlacionados (Raudenbush and Bryk, 2002). Ignorar esta dependencia conduce a una subestimación de los errores estándar y a inferencias incorrectas.

Para abordar esta complejidad teórica, surge la necesidad de emplear Modelos Multinivel (o de Efectos Mixtos). Estos modelos permiten modelar simultáneamente las características del nivel 1 (estudiante y hogar) y del nivel 2 (establecimiento), reconociendo explícitamente que la varianza en el rendimiento escolar es el resultado de la interacción entre los atributos individuales

y el contexto institucional. En el marco de la evaluación del SAE, esto resulta fundamental para separar el rendimiento bruto del valor añadido y para analizar cómo la nueva composición de pares interactúa con la capacidad institucional de los liceos.

### **3.7.1. El Coeficiente de Correlación Intraclase (ICC) como Medida de Segregación**

La estructura jerárquica de los modelos multinivel ofrece, además, una herramienta teórica y metodológica de alto valor para medir la segregación escolar. Desde la perspectiva de la política pública, la segregación se entiende como la distribución desigual de estudiantes con distintas características (socioeconómicas o académicas) entre las escuelas de un sistema (Bellei, 2013).

Estadísticamente, el grado en que el sistema educativo segrega a sus estudiantes puede cuantificarse a través del Coeficiente de Correlación Intraclase (ICC). El ICC mide la proporción de la varianza total de una variable (por ejemplo, el ingreso familiar o el puntaje SIMCE) que es explicada exclusivamente por las diferencias entre establecimientos, en lugar de las diferencias al interior de los mismos.

Un ICC cercano a 0 indicaría un sistema perfectamente integrado, donde la diversidad al interior de cualquier escuela es idéntica a la diversidad total del país. Por el contrario, un ICC elevado (como históricamente ha presentado Chile) refleja un sistema altamente segregado, donde los estudiantes de características similares se concentran en las mismas escuelas. La evaluación de políticas inclusivas, como el Sistema de Admisión Escolar, requiere teóricamente observar si existe una contracción en el ICC a través del tiempo, lo que empíricamente evidenciaría una mayor mixtura social y académica al interior de las aulas institucionales.

## 4. Descripción de los Datos

Antes de continuar con la explicación de los métodos y modelos utilizados para evaluar la implementación del SAE en el sistema educacional chileno, es de suma importancia explicar el origen de estos mismos. Es por esto que en el presente capítulo se presentarán a detalle las características de las muestras, junto con estadísticas descriptivas generales que permitan contextualizar sobre los datos a utilizar, los cuales han sido recogidos del Sistema de Medición de la Calidad de la Educación (SIMCE) y de el Ministerio de Educación, ambos para los años 2016 y 2022.

### 4.1. Fuente

El Sistema de Medición de la Calidad de la Educación (SIMCE) es una evaluación estandarizada la cual se rinde a nivel nacional que mide el rendimiento y desempeño académico de los estudiantes en Chile en distintas áreas de aprendizaje, siendo las más importantes Lenguaje y Matemáticas. Además de los puntajes de dichas pruebas, el SIMCE incorpora una gran cantidad de cuestionarios, los cuales se encuentran dirigidos a estudiantes, padres y apoderados, profesores y establecimientos.

Estos cuestionarios recopilan información sobre el rendimiento académico, las características socioeconómicas y demográficas, así como las condiciones del entorno educativo, lo cual es de vital importancia para comprender las características demográficas y socioeconómicas de cada estudiante, lo cual nos permite analizar de mejor manera el rendimiento académico de cada uno de los estudiantes y los factores externos que influyen en este.

Para este estudio se utilizaron los datos del SIMCE de los años 2016 y 2022. La elección de estos años radica en la implementación del SAE en el sistema de admisión, esto para cumplir con el mandato de la Ley de Inclusión, promulgada en el año 2015, cabe destacar que la última región

en aplicar el SAE fue la Región Metropolitana en el año 2019. El año 2016 se utiliza como periodo pre-reforma, mientras que el año 2022 corresponde al periodo post reforma, esto nos permitirá tener una visión clara sobre la variación de los puntajes, todo el procedimiento será detallado en los capítulos siguientes.

## 4.2. Tablas

Los resultados del SIMCE se encuentran organizados en diferentes tablas, las cuales facilitan la visualización de la información. Cada cuestionario de los mencionados anteriormente se encuentra ubicado en un único archivo Excel, y la glosa para cada variable se encuentra en otro documento ordenado por hojas. Es importante destacar que las variables son distintas para cada año, por lo que revisar la glosa es fundamental para desarrollar la investigación.

Para esta investigación se trabajará con el cuestionario de padres y apoderados, la tabla de puntajes promedio por establecimiento educacional y la tabla con puntajes de cada alumno, y a continuación se presentan las observaciones y variables para cada uno de los archivos a utilizar.

- `_rbd`: Puntajes promedio de cada establecimiento participe del SIMCE de segundo medio.

	<b>SIMCE 2016</b>	<b>SIMCE 2022</b>
Observaciones	2902	2976
Variables	49	40

Tabla 2: Tabla descriptiva del archivo `_rbd`.

- `_cpad`: Cuestionario de padres y apoderados.

	<b>SIMCE 2016</b>	<b>SIMCE 2022</b>
Observaciones	167169	164984
Variables	118	99

Tabla 3: Tabla descriptiva del archivo `_cpad`.

- `_alu`: Resultados de la prueba SIMCE para cada alumno.

	<b>SIMCE 2016</b>	<b>SIMCE 2022</b>
Observaciones	248158	246754
Variables	38	28

Tabla 4: Tabla descriptiva del archivo `_alu`.

### 4.3. Puntajes Promedio Matemática y Lectura

A continuación se presentan las estadísticas descriptivas de los puntajes correspondientes a las pruebas de matemática y lectura, las cuales fueron aplicadas a estudiantes de segundo medio de todo Chile en los años 2016 y 2022. Estos resultados son valores iniciales de las distintas bases de datos, es decir, que aún no se les realiza un proceso de limpieza.

Este primer acercamiento es fundamental para la investigación, ya que permite observar la magnitud de los puntajes en bruto y establecer tendencias de manera preliminar, y posteriormente, permite usarlos de manera comparativa con los resultados ya procesados.

Cabe destacar que ambos puntajes constituyen variables dependientes del modelo económico que se utilizará en este estudio, por lo que tener una medición en bruto de estas permitirá evaluar en que medida influyen otros factores en los puntajes para cada grupo de este estudio.

De manera preliminar, es evidente una variación de los puntajes en ambas evaluaciones a nivel nacional, esto será el eje central de la investigación y se estudiará más a detalle en los siguientes capítulos.

En el siguiente capítulo se realizará la integración de datos, proceso en el cual se seleccionarán las variables necesarias para realizar este estudio, en donde, además, se contempla la transformación de variables cualitativas para facilitar el estudio y además se mencionará la creación de variables nuevas para la realización del modelo. Se detallará también el flujo de integración

y depuración seguido hasta consolidar la base analítica y, con ello, seleccionar el modelo más adecuado para realizar el estudio.

A modo de resumir, se unificarán todas las tablas mencionadas en este capítulo para cada uno de los años en un solo panel de datos utilizando las variables `idalumno` y `rbd` para realizar dicha unificación.

A continuación se aprecia un cuadro resumen con las estadísticas descriptivas de ambas pruebas:

<b>Prueba</b>	<b>Año</b>	<b>Promedio</b>	<b>Mediana</b>	<b>Desviación Estándar</b>	<b>N° de Obs.</b>
Matemática	2016	267.72	269.45	64.96	162766
Matemática	2022	253.24	246.75	61.69	159248
Lectura	2016	249.35	248.72	52.18	161448
Lectura	2022	244.58	247.15	53.25	157446

Tabla 5: Tabla con las estadísticas descriptivas de las pruebas para 2016 y 2022.

## 5. Integración de los Datos

En esta sección se describe el proceso paso a paso para integrar las tablas de manera correcta y obtener un panel de datos que nos permita realizar el análisis adecuado para este estudio. El primer paso es la depuración de los datos, en donde se eliminan columnas innecesarias y/o vacías de cada una de las tablas, esto con el fin de garantizar la consistencia del conjunto de datos. Acto seguido, se detallan las transformaciones realizadas para integrar de manera correcta la información, la creación de variables nuevas según se requiera y también se menciona la transformación de variables cualitativas en formatos adecuados (por ejemplo, variables dicotómicas o categóricas) que permitan su uso en los modelos econométricos.

Este proceso fue fundamental para poder realizar un análisis robusto sobre el impacto del sistema de admisión escolar en los resultados académicos de los estudiantes en Chile durante los años 2016 y 2022.

### 5.1. Manejo de los Datos

Para realizar la integración de los datos de manera correcta se comenzó con la edición de cada archivo en su formato Excel, para optimizar de esa forma la carga de datos y que esta fuera más rápida, en este proceso se eliminaron las columnas innecesarias de cada base de datos para cargar a RStudio solo lo necesario, la carga de las distintas bases de datos disponibles fue el paso siguiente, tanto para el año 2016 como para el 2022. Cada base de datos contiene información específica sobre los estudiantes, sus puntajes en las pruebas SIMCE, las características de los establecimientos educativos y el entorno familiar. Realizar un correcto manejo de los diferentes conjuntos de datos es crucial para obtener un resultado coherente en la investigación.

### 5.1.1. Depuración de los Datos

La Depuración de los Datos hace referencia al proceso de limpieza de las bases de datos antes de realizar cualquier tipo de carga al software a utilizar, en este caso, RStudio. El primer paso de este proceso es seleccionar solo las variables necesarias para el estudio, de esta forma se evita tener que seleccionar las variables directamente en RStudio ya que solo estarán cargadas únicamente las variables necesarias para el estudio.

A continuación se detallarán las variables seleccionadas para el estudio, esto para cada una de las bases de datos disponibles.

- `_rbd`: Para la base de datos que contiene información sobre los puntajes de cada establecimiento se seleccionaron las siguientes variables para el estudio.
  1. `agno`: Variable que contiene el año en el que se rindió el SIMCE.
  2. `rbd`: Es el identificador único para cada establecimiento educacional (Rol Base de Datos).
  3. `nombre_establecimiento`: Es el nombre completo que posee el establecimiento.
  4. `codigo_region`: Número que nos permite identificar la región en la cual se ubica el establecimiento.
  5. `nombre_region`: Nombre de la región en la cual se encuentra el establecimiento.
  6. `codigo_comuna`: Número que nos permite identificar la comuna en la cual se ubica el establecimiento.
  7. `nombre_comuna`: Nombre de la comuna en la cual se encuentra el establecimiento.

8. `dependencia`: Valor numérico que nos indica el tipo de administración que se encarga del establecimiento educacional.
  9. `gse`: Grupo socioeconómico promedio de los alumnos del establecimiento.
  10. `urbano_rural`: Código que nos indica si el establecimiento se encuentra en una zona urbana o rural.
  11. `alumnos_lect`: Número de alumnos que rindieron el SIMCE de Lectura.
  12. `alumnos_mat`: Número de alumnos que rindieron el SIMCE de Matemática.
  13. `promedio_lect`: Promedio del establecimiento en la prueba de Lectura.
  14. `promedio_mat`: Promedio del establecimiento en la prueba de Matemática.
  15. `es_bicentenario`: Variable dicotómica añadida que toma valor 1 cuando el establecimiento en cuestión es un Liceo Bicentenario.
- `_cpad`: Para la base de datos que contiene información sobre el cuestionario de padres y apoderados de cada uno de los estudiantes se seleccionaron las siguientes variables para el estudio.
1. `idalumno`: Identificador único asignado a cada estudiante para distinguirlo dentro de la base de datos.
  2. `ingreso`: Variable que indica el ingreso económico mensual total del grupo familiar del estudiante.
  3. `educacion_madre`: Nivel educacional más alto alcanzado o años de escolaridad de la madre del estudiante.

4. `educacion_padre`: Nivel educacional más alto alcanzado o años de escolaridad del padre del estudiante.
  5. `nacionalidad_padre`: Variable categórica que indica el país de nacimiento del padre del estudiante.
  6. `nacionalidad_madre`: Variable categórica que indica el país de nacimiento de la madre del estudiante.
  7. `nacionalidad_estudiante`: Variable categórica que indica el país de nacimiento del estudiante.
- `_alu`: Para la base de datos que contiene información sobre los resultados académicos de cada uno de los estudiantes se seleccionaron las siguientes variables para el estudio.
    1. `idalumno`: Identificador único asignado a cada estudiante para distinguirlo dentro de la base de datos.
    2. `rbd`: Identificador numérico único del establecimiento educacional (Rol Base de Datos).
    3. `puntaje_mat`: Puntaje obtenido por el estudiante en la prueba SIMCE de Matemática.
    4. `puntaje_lect`: Puntaje obtenido por el estudiante en la prueba SIMCE de Lectura.

Cabe destacar que estos no son los nombres originales de las variables, ya que estos fueron cambiados para facilitar su reconocimiento y su posterior análisis en RStudio, de esta forma se unificaron los nombres para las bases de datos de ambos años.

### 5.1.2. Carga de los Datos

Para cargar los datos en RStudio se utilizó la opción `Import Dataset from Excel`, a la cual se accede desde la opción `File` en la parte superior izquierda del software, y luego de seleccionar el archivo a cargar, se repitió la operación con las otras bases.

### 5.1.3. Repetición para el año 2022

Para los registros correspondientes al año 2022 se aplicó el mismo procedimiento descrito para los datos de 2016, con el objetivo de consolidar una base de datos sólida y comparable. De esta manera, se seleccionaron y depuraron las variables clave necesarias para el análisis econométrico, asegurando de esta forma la consistencia entre ambos años y así realizar la integración de buena manera

## 5.2. Creación y Transformación de Variables

Es importante mencionar la creación de variables que se realizó directamente en RStudio, ya que estas son fundamentales para poder realizar el análisis de manera significativa.

Las variables creadas se describen a continuación, y son las mismas para ambas bases de datos, por lo que cuando se describa una, hace referencia a la de 2016 y 2022.

- `_rbd`: En esta base de datos se crearon las siguientes variables, las cuales son todas dicotómicas:
  - `municipal_no_bic`: Variable que toma el valor 1 cuando es establecimiento es de administración municipal y no es bicentenario y el valor 0 en otro caso.

- `subvencionado_no_bic`: Variable que toma el valor 1 cuando el establecimiento es particular subvencionado y 0 en otro caso.
- `privados`: Variable que toma el valor 1 cuando el establecimiento es privado y 0 en otro caso.
- `_cpad`: En las bases de datos del cuestionario de padres y apoderados se crearon las siguientes variables:
  - `ingresos_miles`: Columna que representa un valor aleatorio del intervalo de ingresos según la respuesta marcada en el cuestionario, está expresada en miles de pesos.
  - `agnos_educacion_padre`: Columna que representa un valor aleatorio de los años de estudio según la opción marcada en el cuestionario, está medida en años.
  - `agnos_educacion_madre`: Columna que representa un valor aleatorio de los años de estudio según la opción marcada en el cuestionario, está medida en años.
  - `es_extranjero`: Variable dicotómica que toma el valor 1 cuando la nacionalidad del estudiante, y por lo menos uno de sus dos padres, es extranjera y 0 en otro caso
- `_alu`: Para estas bases de datos no se crearon variables extras.

## 5.3. Panel de Datos final

### 5.3.1. Panel General

Para este paso se fusionaron las 3 bases de datos para cada año y luego estas fueron combinadas para generar un panel único, lo cual se realizó de la siguiente forma `inner_join` para unir la base de datos de alumnos con padres, usando `idalumno` para realizar la combinación, y

`left_join` para unir el resultado anterior con la base de datos de los establecimientos, usando la columna de `rbd`. Finalmente se añadió una columna extra con el año correspondiente para cada uno de los casos con la función `mutate`.

Como último paso se realizó la combinación de ambos paneles, los cuales contienen toda la información de uno de los años a estudiar, y de esta forma con la función `bind_rows` se combinaron ambos paneles, por lo que de esta forma tenemos anuales para cada establecimiento que se encuentre en ambos años, en caso contrario solo aparecerá el año con datos.

### 5.3.2. Panel de Escuelas

Para la creación de esta panel fue necesario agrupar los datos utilizando la función `group_by` usando las variables `agno` y `rbd` para realizar dicha agrupación.

También se crearon variables que contienen estadísticos descriptivos para cada variable creada con anterioridad las cuales son las que se mostrarán a continuación:

#### ■ Variables de promedio:

- `ingreso_promedio`: Variable que muestra el promedio de ingresos familiares para cada establecimiento para cada uno de los años.
- `prom_educ_padre`: Variable que muestra el promedio de años de estudio de los padres para cada establecimiento para cada uno de los años.
- `prom_educ_madre`: Variable que muestra el promedio de años de estudio de las madres para cada establecimiento para cada uno de los años.
- `prom_matematicas`: Variable que muestra el promedio obtenido por el establecimiento en la prueba de matemáticas.

- `prom_lectura`: Variable que muestra el promedio obtenido por el establecimiento en la prueba de lectura.
- `prop_extranjeros`: Variable que muestra la proporción de alumnos extranjeros en cada establecimiento para cada año de estudio.

■ **Variables de desviación estándar:**

- `de_ingreso`: Variable que muestra la desviación estándar en los ingresos familiares de cada establecimiento para cada año de estudio.
- `de_educacion_padre`: Variable que muestra la desviación estándar de los años de educación de los padres de los estudiantes de cada establecimiento para cada año de estudio.
- `de_educacion_madre`: Variable que muestra la desviación estándar de los años de educación de las madres de los estudiantes de cada establecimiento para cada año de estudio.

Con la creación de estas variables podremos realizar de mejor manera el análisis de composición socioeconómica de cada uno de los liceos, y de esta manera comprobar si los cambios son estadísticamente significativos en el tiempo o no, y también ver si hay cambios estadísticamente significativos entre cada uno de los grupos de estudio.

## 6. Metodología

### 6.1. Análisis Preliminar: Comparación de Medias (T-Tests)

Antes de abordar la estimación de efectos causales, es necesario desarrollar un diagnóstico descriptivo que permita identificar cambios relevantes en los resultados académicos y en la composición del alumnado de los Liceos Bicentenario en el periodo pre-reforma (2016) y post-reforma (2022). Para ello, esta investigación implementa un conjunto de pruebas t de diferencia de medias aplicadas sobre el panel de establecimientos construido en el capítulo anterior. Este panel resume, para cada escuela y año, indicadores promedio de desempeño académico y variables agregadas de composición socioeconómica y demográfica del estudiantado, lo que permite realizar comparaciones tanto entre grupos de establecimientos como entre periodos de tiempo.

La función metodológica de esta etapa es estrictamente exploratoria. Su propósito no es identificar un efecto causal del SAE, sino ofrecer una primera lectura empírica sobre cómo cambian las brechas de desempeño y composición antes y después de la reforma. En coherencia con los objetivos del estudio y con la estructura de resultados que se desarrollará posteriormente, este análisis se organiza en dos dimensiones complementarias.

La primera corresponde a un análisis transversal, orientado a comparar, dentro de un mismo año, a los Liceos Bicentenario con distintos grupos de referencia definidos por dependencia administrativa: establecimientos municipales no Bicentenario, establecimientos particulares subvencionados no Bicentenario y establecimientos particulares pagados. Esta comparación permite observar la posición relativa de los Liceos Bicentenario dentro del sistema escolar en 2016 y nuevamente en 2022.

La segunda corresponde a un análisis longitudinal, orientado a comparar, dentro de cada

grupo de establecimientos, la evolución de las medias entre 2016 y 2022. En este caso, el interés se concentra en establecer si los propios Liceos Bicentenario modificaron significativamente sus indicadores tras la implementación del SAE, y si ese cambio fue similar o distinto al observado en los demás grupos de comparación.

En conjunto, ambos ejercicios descriptivos permiten construir una base empírica más completa para interpretar los resultados posteriores. Mientras el análisis transversal muestra si los Liceos Bicentenario mantienen, reducen o pierden sus brechas relativas frente a otros tipos de establecimientos, el análisis longitudinal permite observar si dentro de cada grupo existieron transformaciones estadísticamente significativas entre el periodo pre y post reforma.

### 6.1.1. Procedimiento del análisis

El procedimiento empírico se implementa sobre la base es establecimientos definitiva, utilizando iteraciones automáticas en *RStudio* para aplicar pruebas t sobre el conjunto de variables de interés. La estrategia contempla dos tipos de contraste.

En primer lugar, para el **análisis transversal**, se compara en cada año la media de una variable  $X$  entre los Liceos Bicentenario y cada grupo de referencia. Formalmente, para cada grupo de comparación  $g$  y para cada año  $t$ , se contrasta la hipótesis nula:

$$H_0 : \mu_{Bic,g,t} = \mu_{Ctrl,g,t}$$

frente a la hipótesis alternativa:

$$H_1 : \mu_{Bic,g,t} \neq \mu_{Ctrl,g,t}$$

donde  $\mu_{Bic,g,t}$  representa la media de la variable de interés para los Liceos Bicentenario en el año  $t$ , y  $\mu_{Ctrl,g,t}$  representa la media del grupo de comparación correspondiente en ese mismo año. Este contraste permite examinar, por ejemplo, si en 2016 los Bicentenario exhibían puntajes o perfiles socioeconómicos distintos a los establecimientos municipales, y si esas brechas se mantuvieron o se redujeron en 2022.

En segundo lugar, para el **análisis longitudinal**, se compara la media de cada variable entre 2016 y 2022 dentro de un mismo grupo de establecimientos. Para cada grupo  $k$ , los cuales son Bicentenario, municipal no Bicentenario, subvencionado no Bicentenario y privado, se contrasta la hipótesis:

$$H_0 : \mu_{k,2016} = \mu_{k,2022}$$

frente a la hipótesis alternativa:

$$H_1 : \mu_{k,2016} \neq \mu_{k,2022}$$

donde  $\mu_{k,2016}$  y  $\mu_{k,2022}$  representan las medias de la variable de interés para el grupo  $k$  en los periodos pre y post reforma, respectivamente. Este contraste permite identificar si, al interior de cada tipo de establecimiento, ocurrieron cambios estadísticamente significativos en el desempeño académico o en la composición del alumnado entre ambos años de estudio.

Desde el punto de vista operativo, el procedimiento sigue tres etapas. Primero, se segmenta la base de datos para seleccionar el grupo o los grupos pertinentes según el tipo de prueba a realizar. Segundo, se identifica la variable de interés que será sometida a contraste. Tercero, se aplica la función `t.test()` para estimar la diferencia de medias, el estadístico de prueba y el

correspondiente valor- $p$ .

Las variables analizadas se agrupan en dos bloques principales. El primero corresponde a los **indicadores académicos**, medidos a través de las variables de promedio de matemáticas y lectura, que resumen el desempeño promedio del establecimiento en SIMCE. El segundo corresponde a los **indicadores de composición del alumnado**, donde se incluyen ingreso promedio, promedio de educación del padre y la madre y la proporción de extranjeros por curso. Estas variables fueron construidas específicamente en el panel de escuelas para representar el perfil socioeconómico y demográfico promedio de los estudiantes en cada establecimiento y año.

La combinación de ambos tipos de contraste es metodológicamente relevante para la investigación. Si los Liceos Bicentenario reducen su ventaja relativa en puntajes frente a otros grupos y, al mismo tiempo, muestran cambios longitudinales en su composición social, entonces la evidencia descriptiva sugiere que la implementación del SAE pudo haber alterado tanto su desempeño bruto como el perfil del alumnado que atienden. Sin embargo, esta etapa no permite por sí sola separar esos cambios de otros factores concurrentes, razón por la cual posteriormente se recurre a modelos econométricos de inferencia causal y a un enfoque multinivel complementario.

### **6.1.2. Criterios de Significancia**

La inferencia estadística de las pruebas  $t$  se basa en el  $p$ -valor asociado a cada contraste. Siguiendo el criterio convencional en economía aplicada, se considera que una diferencia de medias es estadísticamente significativa cuando  $p < 0,05$ . En ese caso, se rechaza la hipótesis nula de igualdad de medias y se concluye que la diferencia observada es poco compatible con una fluctuación puramente aleatoria.

No obstante, en esta investigación la interpretación de los resultados no depende únicamen-

te de la significancia estadística, sino también de la magnitud y dirección de la brecha estimada. En el análisis transversal, una brecha positiva en puntajes académicos indica una ventaja relativa de los Liceos Bicentenario respecto del grupo de comparación, mientras que una brecha negativa indica un desempeño inferior. De modo análogo, en variables socioeconómicas, una brecha positiva en ingreso promedio o escolaridad parental sugiere una composición relativamente más favorecida.

En el análisis longitudinal, en cambio, la lectura se centra en la evolución temporal interna de cada grupo. Un cambio negativo y estadísticamente significativo en puntajes puede interpretarse como una caída en rendimiento bruto entre 2016 y 2022, mientras que una reducción en ingreso promedio o escolaridad parental puede interpretarse como evidencia de una modificación en la composición del estudiantado. En el caso de la proporción de estudiantes extranjeros, una variación significativa da cuenta de cambios en la diversidad demográfica del establecimiento.

Por esta razón, las pruebas *t* cumplen en esta memoria una función de diagnóstico descriptivo. Su utilidad radica en mostrar si existen cambios y brechas estadísticamente relevantes antes de pasar a modelos de mayor complejidad. Sin embargo, dado que estas pruebas no controlan simultáneamente por heterogeneidad no observada entre escuelas ni por shocks comunes entre periodos, como la pandemia y el retorno a la presencialidad en 2022, sus resultados deben interpretarse como evidencia preliminar y no como estimaciones causales del efecto del SAE. Esa limitación es precisamente la que justifica el uso posterior del modelo de efectos fijos de dos vías y del modelo multinivel desarrollado en las secciones siguientes.

### **6.1.3. Alcance interpretativo del análisis preliminar**

A pesar de su utilidad, este análisis debe interpretarse con cautela. La prueba *t* permite identificar diferencias de medias, pero no constituye por sí sola una estrategia de identificación

causal. En consecuencia, una brecha significativa entre Liceos Bicentenario y otros tipos de establecimientos no implica necesariamente que dicha diferencia sea atribuible al SAE, ya que puede estar influida por múltiples factores observados y no observados, entre ellos el efecto acumulado de la pandemia, las trayectorias previas de cada establecimiento o cambios simultáneos en la composición del estudiantado.

Asimismo, dado que las comparaciones se realizan entre tipos de establecimientos con características estructuralmente distintas, la evidencia obtenida debe entenderse como descriptiva y no contrafactual. Su principal fortaleza radica en ordenar empíricamente el problema de investigación y mostrar cómo cambia la posición relativa de los Liceos Bicentenario dentro del sistema escolar chileno entre el periodo pre y post SAE.

Por ello, las pruebas  $t$  se utilizan en esta memoria como una etapa preliminar y complementaria. Su propósito es aportar una lectura inicial sobre la evolución de las brechas académicas y socioeconómicas, y preparar el terreno para la estrategia econométrica posterior basada en efectos fijos, la cual sí busca aproximarse al efecto causal medio del cambio institucional introducido por el SAE.

## **6.2. Modelo de Inferencia Causal: Efectos Fijos.**

Una vez establecido el diagnóstico descriptivo, la investigación avanza hacia una estrategia de inferencia causal orientada a estimar el efecto promedio asociado a la implementación del SAE sobre los resultados académicos de los Liceos Bicentenario. El problema econométrico central consiste en distinguir cuánto de la variación observada en el desempeño responde efectivamente al cambio institucional en el régimen de admisión, y cuánto puede atribuirse a diferencias históricas

entre establecimientos o a shocks comunes que afectan a todo el sistema educativo.

Para abordar este desafío, se utiliza un modelo de efectos fijos de dos vías (Two-Way Fixed Effects, TWFE) aplicado sobre el panel de establecimientos. Esta elección es coherente con la estructura de datos construida en la investigación, donde para cada establecimiento y año se dispone de variables agregadas de desempeño académico y composición del alumnado.

### **6.2.1. Lógica del Tratamiento y Adopción Escalonada**

La lógica del tratamiento se basa en la exposición diferencial de los establecimientos Bicentenario al nuevo régimen de admisión escolar. En términos sustantivos, la implementación del SAE alteró las condiciones de entrada de estos liceos al eliminar los mecanismos de selección académica que habían caracterizado históricamente su composición estudiantil. En consecuencia, el tratamiento relevante para esta sección no es simplemente “ser Bicentenario”, sino ser un establecimiento Bicentenario en el periodo posterior a la implementación del SAE.

En la práctica, el grupo tratado está constituido por los establecimientos que forman parte del conjunto Bicentenario y que, en el periodo post-reforma, operan bajo las reglas de admisión del SAE. El grupo de comparación está compuesto por establecimientos no Bicentenario observados en los mismos años. Esta formulación permite construir un contraste de diferencias en diferencias, donde el interés no radica en comparar niveles absolutos entre escuelas, sino cambios relativos entre periodos.

Dentro del grupo expuesto es posible distinguir dos subcohortes relevantes. Por una parte, se encuentran los Bicentenarios históricos, es decir, establecimientos que ya tenían dicha condición antes del periodo post-SAE. Por otra, los nuevos Bicentenarios, que adquieren esa categoría hacia el segundo periodo. Esta distinción es importante desde el punto de vista sustantivo, pues los

primeros enfrentan el shock de admisión manteniendo una identidad institucional consolidada, mientras que los segundos combinan el cambio en admisión con una transición hacia la categoría Bicentenario. Sin embargo, en la especificación principal ambas trayectorias se integran en un único efecto promedio del tratamiento, dejando la diferenciación por cohortes como una extensión natural para pruebas de robustez.

### 6.2.2. Especificación Econométrica del Modelo

La especificación base del modelo de efectos fijos de dos vías se define como:

$$Y_{st} = \alpha + \delta(Bic_s \times Post_t) + \lambda'X_{st} + \eta_s + \tau_t + \varepsilon_{st}$$

donde  $Y_{st}$  representa el resultado académico promedio del establecimiento  $s$  en el año  $t$ , medido alternativamente mediante el puntaje SIMCE promedio en Matemática o Lectura.

La variable  $Bic_s$  es una variable dicotómica que identifica a los establecimientos pertenecientes al grupo Bicentenario. Por su parte,  $Post_t$  es una variable temporal que toma valor 1 en el periodo posterior a la implementación del SAE y 0 en el periodo pre-reforma. El término de interacción  $Bic_s \times Post_t$  constituye el parámetro de interés del modelo, cuyo coeficiente  $\delta$  captura el cambio diferencial de los establecimientos Bicentenario en el periodo posterior al SAE en comparación con el cambio experimentado por el grupo de referencia.

El término  $X_{st}$  corresponde a un vector de controles observables que varían en el tiempo a nivel de establecimiento y que aproximan la composición del alumnado. A su vez,  $\eta_s$  representa efectos fijos de establecimiento, los cuales absorben todas aquellas características inobservables que permanecen constantes en el tiempo para cada escuela, tales como infraestructura, localización,

cultura organizacional o reputación histórica. De forma complementaria,  $\tau_t$  corresponde a efectos fijos de año, que capturan shocks agregados comunes a todos los establecimientos, como el efecto sistémico de la pandemia y del retorno a la presencialidad observado en 2022. Finalmente,  $\varepsilon_{st}$  es el término de error idiosincrático.

La inclusión simultánea de  $\eta_s$  y  $\tau_t$  es la que transforma la especificación en un modelo de efectos fijos de dos vías. En este marco, la identificación del parámetro  $\delta$  proviene de la comparación entre la evolución temporal del grupo Bicentenario y la evolución temporal del grupo de comparación, una vez descontadas las diferencias permanentes entre establecimientos y los shocks comunes a todos ellos.

Para asegurar la validez de la inferencia, los errores estándar deben agruparse a nivel de establecimiento, dado que la estructura de panel implica la posible existencia de correlación serial dentro de cada unidad observacional a través del tiempo.

### **6.2.3. Control de Composición y el Desafío del Valor Añadido**

Un aspecto metodológico central de esta investigación es la distinción entre rendimiento bruto y valor añadido pedagógico. Los puntajes promedio observados en un establecimiento no reflejan únicamente la calidad de su gestión o de sus prácticas pedagógicas, sino también las características de entrada de los estudiantes que lo componen. En el caso de los Liceos Bicentenario, esta distinción adquiere especial relevancia, ya que la implementación del SAE modificó precisamente el mecanismo de conformación de la matrícula.

Desde esta perspectiva, el vector de controles  $X_{st}$  cumple una función sustantiva y no meramente técnica. Su inclusión busca aproximar los cambios en la composición social y académica del alumnado, de modo tal que el coeficiente de interés no recoja únicamente el efecto mecáni-

co de recibir estudiantes con perfiles distintos, sino el cambio promedio asociado a la condición Bicentenario bajo el nuevo régimen de admisión.

En esta investigación, los controles incorporan tres dimensiones. La primera es el nivel socioeconómico promedio del alumnado, aproximado mediante ingreso promedio del hogar y escolaridad promedio del padre y de la madre. La segunda es la heterogeneidad interna de la matrícula, capturada a través de medidas de dispersión como la desviación estándar del ingreso y de la escolaridad parental, ya que una mayor heterogeneidad puede alterar tanto la dinámica pedagógica como la intensidad del efecto par. La tercera dimensión corresponde a la diversidad demográfica, aproximada mediante la proporción de estudiantes extranjeros, con el objeto de distinguir cambios vinculados a la composición migratoria de los establecimientos.

El sentido de esta estrategia es claro: si los Liceos Bicentenario experimentan una caída en sus puntajes promedio después del SAE, esa caída no debe interpretarse automáticamente como una disminución de efectividad institucional. Parte de esa variación puede responder a la entrada de cohortes más heterogéneas o socialmente más diversas. El modelo de efectos fijos con controles permite aproximarse mejor a esa distinción y, por tanto, entregar una lectura más consistente con el debate entre selección, desempeño bruto y valor añadido.

#### **6.2.4. Limitaciones del Modelo de Efectos Fijos de Dos Vías**

Si bien el modelo de efectos fijos de dos vías constituye la estrategia principal de inferencia causal de esta memoria, su alcance debe entenderse en términos precisos. Su fortaleza radica en estimar un efecto promedio del cambio institucional, controlando por heterogeneidad inobservable fija entre establecimientos y por shocks temporales comunes. Sin embargo, esto no significa que resuelva completamente todos los desafíos de identificación presentes en el problema estudiado.

En primer lugar, la validez causal del estimador descansa en una versión del supuesto de tendencias paralelas: en ausencia del SAE, la trayectoria de los resultados académicos de los establecimientos Bicentenario habría evolucionado de manera similar a la del grupo de comparación, una vez controladas las diferencias observables e invariables. Dado que la investigación trabaja con un número acotado de cortes temporales, este supuesto no puede verificarse con la misma profundidad que en un panel con múltiples años pre tratamiento, por lo que la interpretación causal debe mantenerse prudente.

En segundo lugar, el año 2022 no refleja únicamente el periodo posterior a la implementación del SAE, sino también un contexto educacional extraordinario marcado por la pandemia, la interrupción de trayectorias escolares y el retorno a la presencialidad. Aunque los efectos fijos de año absorben el shock promedio común a todos los establecimientos, no eliminan necesariamente la posibilidad de impactos heterogéneos entre tipos de escuelas.

En tercer lugar, el modelo se estima a nivel agregado de establecimiento, por lo que no permite representar explícitamente la estructura jerárquica de los datos ni separar con precisión la variación atribuible al nivel individual de aquella correspondiente al nivel institucional. Esta limitación es particularmente importante cuando los objetivos de la investigación incluyen composición del alumnado y segregación escolar, dimensiones para las cuales resulta más apropiado un enfoque multinivel.

Finalmente, en presencia de cohortes tratadas con trayectorias potencialmente distintas, tales como Bicentenarios históricos y nuevos Bicentenarios, el coeficiente promedio puede resumir efectos heterogéneos de naturaleza diferente. Por esta razón, el modelo de efectos fijos debe complementarse con estimaciones de robustez y con el modelo multinivel desarrollado en la sección siguiente, el cual permite examinar con mayor profundidad los mecanismos de composición

y segregación que el TWFE no fue diseñado para capturar de manera directa.

### **6.3. Modelo Multinivel de Efectos Mixtos para el Análisis de Composición y Segregación Escolar**

#### **6.3.1. Fundamentación Metodológica del Modelo**

Con el fin de complementar el modelo de efectos fijos y profundizar en los objetivos asociados a la composición del alumnado y la segregación escolar, esta investigación incorpora un modelo multinivel de efectos mixtos. La elección de este enfoque responde a la naturaleza jerárquica de la base de datos utilizada, en la cual los estudiantes constituyen unidades de observación de primer nivel y los establecimientos educacionales representan unidades de segundo nivel.

En estudios de economía de la educación, esta estructura reviste especial relevancia, ya que los resultados individuales no dependen únicamente de las características propias del estudiante o de su hogar, sino también de atributos compartidos con sus pares y con el establecimiento al que pertenece. En otras palabras, dos estudiantes matriculados en el mismo liceo enfrentan un entorno organizacional, pedagógico y social común, por lo que asumir independencia total entre observaciones conduciría a una representación incompleta del proceso educativo.

Desde el punto de vista sustantivo, este modelo resulta especialmente conveniente para la presente investigación por tres razones. En primer lugar, permite aprovechar la información disponible a nivel estudiante sin perder de vista el contexto escolar en el cual se generan los aprendizajes. En segundo lugar, entrega una forma empíricamente consistente de analizar el llamado efecto par, toda vez que distingue entre características individuales y componentes compartidos al interior del establecimiento. En tercer lugar, posibilita una aproximación más sólida al estudio de

la segregación, al cuantificar qué proporción de la variación observada en variables académicas o socioeconómicas se concentra entre escuelas y qué proporción se explica dentro de ellas.

En este sentido, el modelo multinivel no reemplaza la estrategia principal de inferencia causal basada en efectos fijos, sino que la complementa. Mientras el modelo de efectos fijos permite estimar el impacto promedio asociado al cambio institucional introducido por el SAE, el modelo multinivel permite comprender con mayor profundidad cómo dicho cambio se vincula con la nueva composición del alumnado y con la distribución de la heterogeneidad entre establecimientos.

### **6.3.2. Estructura Jerárquica de los Datos**

La base consolidada para esta investigación permite trabajar con observaciones individuales de estudiantes, identificados por  $id_{alumno}$ , vinculados a sus respectivos establecimientos mediante la variable  $rbd$ . A partir de dicha estructura, es posible definir una organización jerárquica en la que cada estudiante  $i$  pertenece a un establecimiento  $s$  en el año  $t$ .

Esta estructura es coherente con la lógica del problema estudiado. La implementación del SAE no solo puede afectar los puntajes promedio observados en los establecimientos, sino también alterar la composición del grupo de pares, la dispersión socioeconómica al interior del liceo y la heterogeneidad del contexto de aprendizaje. Por ello, una modelación a dos niveles permite capturar de mejor manera los mecanismos intermedios mediante los cuales una política de admisión puede influir sobre los resultados escolares.

### **6.3.3. Especificación Econométrica**

La especificación base del modelo multinivel para resultados académicos se define como:

$$Y_{ist} = \alpha + \delta (Bic_s \times Post_t) + \beta' X_{ist} + \gamma' Z_{st} + \tau_t + u_s + \varepsilon_{ist} \quad (3)$$

donde:

- $Y_{ist}$  representa el resultado académico del estudiante  $i$ , perteneciente al establecimiento  $s$ , en el año  $t$ . En esta investigación, esta variable corresponde al puntaje SIMCE de Matemática.
- $Bic_s$  es una variable dicotómica que toma valor 1 si el establecimiento corresponde a un Liceo Bicentenario y 0 en caso contrario.
- $Post_t$  es una variable temporal que toma valor 1 para el periodo post-SAE y 0 para el periodo pre-SAE.
- $Bic_s \times Post_t$  es el término de interacción de interés, cuyo coeficiente  $\delta$  captura el cambio diferencial asociado a los Liceos Bicentenario en el periodo posterior a la implementación del SAE.
- $X_{ist}$  es un vector de controles a nivel estudiante y hogar, tales como ingreso familiar, escolaridad de los padres y condición de extranjero.
- $Z_{st}$  corresponde a variables observadas a nivel establecimiento y año, tales como composición promedio del alumnado, dispersión socioeconómica y demás indicadores agregados relevantes.
- $\tau_t$  representa efectos fijos de año, los cuales absorben shocks comunes a todos los establecimientos en cada periodo.

- $u_s$  es el intercepto aleatorio del establecimiento, que captura factores no observados compartidos por los estudiantes de una misma escuela.
- $\varepsilon_{ist}$  es el término de error idiosincrático individual.

Se asume que:

$$u_s \sim N(0, \sigma_u^2), \quad \varepsilon_{ist} \sim N(0, \sigma_\varepsilon^2) \quad (4)$$

y que ambos componentes son independientes entre sí.

#### 6.3.4. Interpretación del Coeficiente de Interés

El parámetro central del modelo es  $\delta$ , asociado a la interacción entre condición de Bicentenario y periodo posterior a la implementación del SAE. Su interpretación es análoga, en términos sustantivos, a la del estimador de diferencias en diferencias: mide en cuánto cambia el resultado esperado de los estudiantes de liceos Bicentenario en el periodo post-SAE, en relación con el cambio observado en el grupo de comparación, una vez controladas las características individuales, familiares y escolares incluidas en el modelo.

La diferencia respecto del enfoque tradicional radica en que, en este caso, la estimación reconoce explícitamente que los estudiantes se agrupan dentro de establecimientos y que existe una correlación intrínseca entre observaciones pertenecientes a una misma unidad escolar. Por ello, el modelo multinivel entrega una estimación más adecuada de la incertidumbre estadística y una representación más fiel del fenómeno educacional bajo estudio.

### **6.3.5. Aplicación del Modelo a los Objetivos Específicos**

El uso del modelo multinivel permite abordar de manera diferenciada tres dimensiones centrales de la investigación.

En primer lugar, respecto de los resultados académicos, el modelo permite estimar si la implementación del SAE se asocia a cambios en los puntajes SIMCE de Matemática y Lectura de los estudiantes pertenecientes a Liceos Bicentenario, controlando simultáneamente por el perfil socioeconómico del hogar y por las condiciones del establecimiento.

En segundo lugar, respecto de la composición del alumnado, el mismo enfoque puede aplicarse utilizando como variables dependientes indicadores de origen social y familiar, tales como ingreso del hogar, escolaridad del padre, escolaridad de la madre y condición de extranjero. En este caso, el interés ya no reside en explicar el rendimiento, sino en estimar si la composición del estudiantado de los Liceos Bicentenario cambió de forma sistemática tras la implementación del SAE.

En tercer lugar, respecto de la segregación escolar, el modelo ofrece una herramienta particularmente útil al permitir descomponer la varianza total de una variable entre el nivel individual y el nivel establecimiento. Esta descomposición permite evaluar si las diferencias socioeconómicas se encuentran altamente concentradas entre escuelas o si, por el contrario, existe una mayor mezcla social dentro de ellas.

### **6.3.6. Medición de Segregación mediante el Coeficiente de Correlación Intraclase**

Para operacionalizar el análisis de segregación se utilizará el Coeficiente de Correlación Intraclase (ICC), obtenido a partir de un modelo nulo multinivel sin covariables:

$$Y_{ist} = \alpha + u_s + \varepsilon_{ist} \quad (5)$$

A partir de esta especificación, el ICC se calcula como:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} \quad (6)$$

Este coeficiente indica la proporción de la varianza total de la variable analizada que se explica por diferencias entre establecimientos. Un valor elevado del ICC sugiere que los estudiantes se encuentran fuertemente agrupados en escuelas con perfiles semejantes, lo cual es consistente con un escenario de mayor segregación. Por el contrario, una disminución del ICC entre 2016 y 2022 puede interpretarse como evidencia de una menor concentración entre establecimientos y, por ende, como una señal compatible con una reducción de la segregación.

En esta investigación, el ICC será estimado para variables tales como ingreso familiar, escolaridad parental y, complementariamente, para indicadores de rendimiento. De este modo, será posible evaluar no solo si cambió el promedio de la composición social del alumnado Bicentenario, sino también si se modificó el grado en que dicha composición se concentra entre distintos establecimientos.

### **6.3.7. Ventajas del Modelo Multinivel**

La incorporación del modelo multinivel presenta varias ventajas metodológicas en el contexto del presente estudio.

Primero, mejora la coherencia entre la teoría y la estimación empírica. El marco conceptual de esta memoria otorga un lugar central al efecto par, a la heterogeneidad del aula y al contraste

entre rendimiento bruto y valor añadido. Un modelo que distingue entre niveles individual y escolar se ajusta de manera más natural a esa lógica que uno puramente agregado.

Segundo, permite aprovechar de mejor forma la riqueza informativa de la base de datos. Dado que la investigación dispone de información individual de estudiantes y de variables agregadas del establecimiento, el modelo multinivel posibilita integrar ambos niveles sin sacrificar información relevante.

Tercero, fortalece la discusión sobre segregación. El modelo de efectos fijos es adecuado para estimar cambios promedio, pero no fue diseñado específicamente para cuantificar la proporción de desigualdad que se concentra entre escuelas. El enfoque multinivel, en cambio, sí permite realizar esa descomposición de manera explícita.

Cuarto, ofrece una lectura más completa del efecto del SAE. Mientras el modelo de efectos fijos responde principalmente a la pregunta por el efecto promedio del cambio institucional sobre el desempeño, el modelo multinivel permite examinar si dicho cambio vino acompañado por transformaciones en la composición del alumnado y en la estructura de la desigualdad escolar.

### **6.3.8. Alcances y Limitaciones del Modelo**

A pesar de sus ventajas, el modelo multinivel debe entenderse como una estrategia complementaria y no como un sustituto del modelo de efectos fijos utilizado para la identificación causal principal. Su principal fortaleza reside en describir y modelar la estructura de heterogeneidad entre estudiantes y establecimientos; por ello, sus resultados deben interpretarse en conjunto con la evidencia obtenida a partir del modelo de efectos fijos.

Asimismo, dado que el estudio considera dos años de observación, la variación temporal disponible es acotada. En consecuencia, el modelo multinivel no resuelve por sí solo el desafío de

separar completamente el efecto del SAE del shock asociado a la pandemia y al retorno a la presencialidad en 2022. Por esta razón, sus estimaciones deben leerse como evidencia complementaria para comprender mecanismos de composición y segregación, más que como una fuente aislada de identificación causal.

#### **6.4. Limitaciones de Identificación y Pruebas de Robustez**

La estrategia empírica adoptada en esta investigación permite aproximarse de manera consistente al efecto del SAE sobre los Liceos Bicentenario; sin embargo, es necesario explicitar ciertas limitaciones para una interpretación rigurosa de los resultados.

En primer lugar, la comparación entre 2016 y 2022 involucra no solo un cambio en el régimen de admisión escolar, sino también el efecto acumulado de la pandemia, la interrupción de trayectorias educativas y el retorno a la presencialidad. En consecuencia, el año 2022 debe entenderse como un periodo post-reforma en un contexto sistémico extraordinario, por lo que el coeficiente estimado no puede interpretarse como una medida completamente aislada del SAE.

En segundo lugar, la validez del enfoque de diferencias en diferencias descansa sobre el supuesto de tendencias paralelas. Dado que esta investigación trabaja con dos cortes temporales principales, no es posible verificar dicho supuesto con la misma solidez que en un panel con múltiples años pre tratamiento. Por ello, la interpretación causal de los resultados debe realizarse con cautela y en estrecha relación con la plausibilidad institucional del diseño.

En tercer lugar, la integración de datos mediante la vinculación entre alumnos, apoderados y establecimientos puede implicar pérdida de observaciones, especialmente cuando existen registros incompletos en los cuestionarios familiares. Si la ausencia de información no fuese aleatoria, ello

podría introducir un sesgo de selección en la muestra final analizada.

En cuarto lugar, algunas variables socioeconómicas fueron construidas a partir de categorías originales del cuestionario, lo que supone un proceso de aproximación cuantitativa que, si bien es útil para el análisis econométrico, puede introducir error de medición. En particular, la asignación de valores numéricos a tramos de ingreso y niveles educacionales debe entenderse como una operacionalización empírica del nivel socioeconómico y no como una medición exacta del mismo.

Con el objeto de evaluar la estabilidad de los resultados, se realizarán adicionalmente pruebas de robustez. Entre ellas, se consideran estimaciones separadas para Matemática y Lectura, especificaciones con y sin controles de composición, análisis restringidos a establecimientos presentes en ambos periodos, y estimaciones ponderadas por el número de estudiantes que rinden cada prueba. Estas verificaciones permitirán evaluar si los resultados principales se mantienen bajo distintas definiciones muestrales y modelaciones alternativas.

## 7. Resultados

### 7.1. Resultados preliminares de las Pruebas-T de tipo Longitudinal

En esta sección se presentan los resultados preliminares del análisis longitudinal, cuyo objetivo es identificar cómo evolucionaron entre 2016 y 2022 los indicadores académicos y de composición del alumnado dentro de cada grupo de establecimientos. En particular, este ejercicio permite observar si los Liceos Bicentenario experimentaron cambios significativos tras la implementación del SAE y contrastar esa evolución con la observada en liceos municipales, particulares subvencionados y particulares pagados. Dado su carácter descriptivo, estos resultados constituyen una primera aproximación al problema de investigación y entregan un contexto necesario para interpretar posteriormente los hallazgos de los modelos econométricos.

#### 7.1.1. Bicentenarios 2016 vs. 2022

Grupo	Variable	Media 2016	Media 2022	Cambio	¿Sig.?
Bicentenario	Prom. Matemática	299.91	248.51	-51.40	SÍ
	Prom. Lectura	275.97	240.70	-35.27	SÍ
	Ingreso Promedio	474.84	514.30	+39.46	SÍ
	Educ. Padre	11.83	11.30	-0.53	SÍ
	Educ. Madre	12.05	11.76	-0.29	SÍ
	Prop. Extranjeros	0.02	0.05	+0.03	SÍ

Tabla 6: Análisis longitudinal descriptivo: Bicentenario 2016 vs. 2022.

Los resultados de la Tabla 6 muestran que los Liceos Bicentenario experimentaron una caída importante y estadísticamente significativa en sus puntajes promedio entre 2016 y 2022, tanto en Matemática como en Lectura. La magnitud de esta disminución es especialmente alta en Matemática, donde el cambio alcanza -51,40 puntos, mientras que en Lectura la reducción es de -35,27 puntos. En términos descriptivos, esto sugiere que el periodo posterior a la implementación

del SAE coincide con una pérdida importante del rendimiento bruto observado en este grupo de establecimientos.

En paralelo, la composición del alumnado también muestra modificaciones significativas. Aunque el ingreso promedio aumenta, la escolaridad del padre y de la madre disminuye, mientras que la proporción de estudiantes extranjeros crece. En conjunto, estos resultados sugieren que el cambio observado en los Bicentenario no se limita al plano académico, sino que también involucra una transformación del perfil social y demográfico de sus estudiantes. Esta combinación entre caída en puntajes y mayor heterogeneidad de la matrícula es consistente con la hipótesis central de la investigación, según la cual el SAE habría alterado las condiciones de composición sobre las que anteriormente operaba el modelo de excelencia de estos liceos.

### 7.1.2. Municipales 2016 vs. 2022

Grupo	Variable	Media 2016	Media 2022	Cambio	¿Sig.?
Municipal	Prom. Matemática	226.42	226.93	+0.51	NO
	Prom. Lectura	227.63	229.07	+1.44	NO
	Ingreso Promedio	342.91	468.10	+125.19	SÍ
	Educ. Padre	9.98	10.90	+0.92	SÍ
	Educ. Madre	10.06	11.24	+1.18	SÍ
	Prop. Extranjeros	0.02	0.11	+0.09	SÍ

Tabla 7: Análisis longitudinal descriptivo: Municipal 2016 vs. 2022.

La Tabla 7 muestra que, a diferencia de los Liceos Bicentenario, los establecimientos municipales no presentan cambios estadísticamente significativos en sus puntajes promedio de Matemática y Lectura entre 2016 y 2022. Las variaciones observadas son pequeñas y no significativas, lo que sugiere una relativa estabilidad en el rendimiento bruto de este grupo durante el periodo analizado.

Sin embargo, sí se aprecian cambios significativos en la composición del alumnado. El ingreso promedio aumenta con fuerza, al igual que la escolaridad parental y la proporción de estudiantes extranjeros. Esto indica que, aunque el perfil social y demográfico de los establecimientos municipales también cambió en el periodo, esos cambios no se tradujeron en variaciones relevantes en los puntajes promedio. Comparado con el caso Bicentenario, este resultado es importante porque sugiere que la caída académica observada en estos últimos no fue un fenómeno homogéneo en todo el sector público, sino una trayectoria más específica dentro del grupo tratado.

### 7.1.3. Subvencionado 2016 vs. 2022

Grupo	Variable	Media 2016	Media 2022	Cambio	¿Sig.?
Subvencionado	Prom. Matemática	269.46	251.72	-17.74	SÍ
	Prom. Lectura	251.58	246.34	-5.24	SÍ
	Ingreso Promedio	583.46	693.39	+109.93	SÍ
	Educ. Padre	12.16	12.67	+0.51	SÍ
	Educ. Madre	12.23	12.93	+0.70	SÍ
	Prop. Extranjeros	0.01	0.05	+0.04	SÍ

Tabla 8: Análisis longitudinal descriptivo: Subvencionado 2016 vs 2022.

En los establecimientos particulares subvencionados, la Tabla 8 evidencia caídas estadísticamente significativas en ambas pruebas, aunque de menor magnitud que las observadas en los Liceos Bicentenario. En Matemática la disminución alcanza -17,74 puntos y en Lectura -5,24 puntos. Esto indica que también existió un deterioro del rendimiento bruto en este grupo entre 2016 y 2022, pero claramente menos pronunciado que en los Bicentenario.

Al mismo tiempo, el sector subvencionado presenta mejoras en ingreso promedio, escolaridad del padre, escolaridad de la madre y aumento en la proporción de estudiantes extranjeros. Por lo tanto, el patrón de cambio aquí combina una leve caída académica con un fortalecimiento relativo de ciertos indicadores de capital familiar. Este contraste resulta especialmente relevante

para la investigación, ya que muestra que no todos los grupos enfrentaron la misma combinación de deterioro académico y cambio composicional, reforzando la idea de que el caso Bicentenario responde a una dinámica particular dentro del sistema escolar.

#### 7.1.4. Privado 2016 vs. 2022

Grupo	Variable	Media 2016	Media 2022	Cambio	¿Sig.?
Privado	Prom. Matemática	321.77	307.55	-14.22	SÍ
	Prom. Lectura	278.68	276.71	-1.97	NO
	Ingreso Promedio	1989.72	2036.19	+46.47	NO
	Educ. Padre	17.11	17.12	+0.01	NO
	Educ. Madre	16.40	16.67	+0.27	SÍ
	Prop. Extranjeros	0.02	0.04	+0.02	SÍ

Tabla 9: Análisis longitudinal descriptivo: Privado 2016 vs. 2022.

La Tabla 9 muestra que los establecimientos particulares pagados también registran una caída significativa en Matemática, aunque bastante menor que la observada en los Liceos Bicentenario. En Lectura, en cambio, la disminución no resulta estadísticamente significativa, lo que da cuenta de una mayor estabilidad relativa de este grupo en comparación con los demás sectores analizados.

Desde el punto de vista de la composición, los cambios son más acotados. El ingreso promedio y la escolaridad del padre no presentan variaciones significativas, mientras que la escolaridad de la madre y la proporción de estudiantes extranjeros sí aumentan. En conjunto, estos resultados sugieren que el sector privado mantuvo una estructura social mucho más estable durante el periodo, lo que refuerza la interpretación de que los cambios observados en los Bicentenario deben leerse dentro de un contexto donde ciertos grupos conservaron una posición más sólida tanto en desempeño como en composición del alumnado.

## 7.2. Resultados preliminares de las Pruebas-T de tipo Transversal

La presente sección desarrolla el análisis transversal de las pruebas t, comparando en cada año a los Liceos Bicentenario con distintos grupos de referencia del sistema escolar chileno. Este ejercicio permite observar la posición relativa de estos establecimientos antes y después de la implementación del SAE, tanto en resultados académicos como en variables asociadas a la composición social y demográfica del alumnado. A diferencia del análisis longitudinal, aquí el foco no está en la evolución interna de cada grupo, sino en el cambio de las brechas entre Bicentenario y sus comparadores, lo que resulta clave para evaluar si estos liceos conservaron, redujeron o perdieron su ventaja relativa dentro del sistema.

### 7.2.1. Bicentenarios vs. Municipales

Grupo	Variable	Media Ctrl	Media Bic	Brecha	¿Sig.?
Municipal	Prom. Matemática	226.42	299.91	+73.49	SÍ
	Prom. Lectura	227.63	275.97	+48.34	SÍ
	Ingreso Promedio	342.91	474.84	+131.93	SÍ
	Educ. Padre	9.98	11.83	+1.85	SÍ
	Educ. Madre	10.06	12.05	+1.99	SÍ
	Prop. Extranjeros	0.02	0.02	0.00	NO

Tabla 10: Análisis descriptivo comparativo: Municipal (2016).

Grupo	Variable	Media Ctrl	Media Bic	Brecha	¿Sig.?
Municipal	Prom. Matemática	226.93	248.51	+21.58	SÍ
	Prom. Lectura	229.07	240.70	+11.63	SÍ
	Ingreso Promedio	468.10	514.30	+46.20	SÍ
	Educ. Padre	10.90	11.30	+0.40	SÍ
	Educ. Madre	11.24	11.76	+0.52	SÍ
	Prop. Extranjeros	0.11	0.05	-0.06	SÍ

Tabla 11: Análisis descriptivo comparativo: Municipal (2022).

Esta sección compara a los Liceos Bicentenario con el resto de los establecimientos de administración municipal.

- **Resultados Académicos:** En 2016, los Bicentenarios superaban ampliamente a los municipales con brechas de +73.49 puntos en Matemática y +48.34 en Lectura. Para 2022, aunque siguen siendo superiores, la brecha se redujo drásticamente a +21.58 y +11.63 respectivamente.
- **Perfil Socioeconómico:** En 2016, los Bicentenarios tenían alumnos con ingresos significativamente mayores (+131,93 mil) y casi 2 años más de escolaridad en los padres que el promedio municipal. En 2022, estas diferencias socioeconómicas se acortaron (la brecha de escolaridad bajó a unos 0.5 años), reflejando una mayor similitud en el origen de los estudiantes.
- **Significado:** Las tablas muestran que, tras el SAE, los Liceos Bicentenarios se han vuelto socioeconómicamente más parecidos al resto de los liceos municipales, lo que coincide con una disminución notable en su ventaja de puntaje bruto.

### 7.2.2. Bicentenarios vs. Subvencionados

Grupo	Variable	Media Ctrl	Media Bic	Brecha	¿Sig.?
Subvencionado	Prom. Matemática	269.46	299.91	+30.45	SÍ
	Prom. Lectura	251.58	275.97	+24.39	SÍ
	Ingreso Promedio	583.46	474.84	-108.62	SÍ
	Educ. Padre	12.16	11.83	-0.33	SÍ
	Educ. Madre	12.23	12.05	-0.18	NO
	Prop. Extranjeros	0.01	0.02	+0.01	NO

Tabla 12: Análisis descriptivo comparativo: Subvencionado (2016).

Grupo	Variable	Media Ctrl	Media Bic	Brecha	¿Sig.?
Subvencionado	Prom. Matemática	251.72	248.51	-3.21	NO
	Prom. Lectura	246.34	240.70	-5.64	SÍ
	Ingreso Promedio	693.39	514.30	-179.09	SÍ
	Educ. Padre	12.67	11.30	-1.37	SÍ
	Educ. Madre	12.93	11.76	-1.17	SÍ
	Prop. Extranjeros	0.05	0.05	0.00	NO

Tabla 13: Análisis descriptivo comparativo: Subvencionado (2022).

Aquí se contrastan los resultados con los colegios particulares subvencionados (administración privada con fondos estatales).

- **Diferencias en Desempeño:** En 2016, los Bicentenarios rendían significativamente mejor que los subvencionados en ambas pruebas. Sin embargo, en 2022, la ventaja en Matemática desapareció (la diferencia de -3.21 no es estadísticamente significativa) y en Lectura los Bicentenarios pasaron a estar por debajo del grupo de control con una brecha de -5.64 puntos.
- **Composición Social:** Un dato relevante es que en 2016 los Bicentenarios ya tenían un ingreso promedio menor que los subvencionados. Para 2022, esta diferencia de ingresos se profundizó (-179,09 mil), al igual que la brecha en la escolaridad de los padres, donde los subvencionados ahora superan a los Bicentenarios por más de un año.
- **Significado:** Los resultados sugieren que los Bicentenarios han perdido su posición de ventaja académica frente a los subvencionados, a la par que su población escolar se ha vuelto comparativamente más vulnerable tras el fin de la selección.

### 7.2.3. Bicentenarios vs. Privados

Grupo	Variable	Media Ctrl	Media Bic	Brecha	¿Sig.?
Privado	Prom. Matemática	321.77	299.91	-21.86	SÍ
	Prom. Lectura	278.68	275.97	-2.71	NO
	Ingreso Promedio	1989.72	474.84	-1514.88	SÍ
	Educ. Padre	17.11	11.83	-5.28	SÍ
	Educ. Madre	16.40	12.05	-4.35	SÍ
	Prop. Extranjeros	0.02	0.02	0.00	NO

Tabla 14: Análisis descriptivo comparativo: Privado (2016).

Grupo	Variable	Media Ctrl	Media Bic	Brecha	¿Sig.?
Privado	Prom. Matemática	307.55	248.51	-59.04	SÍ
	Prom. Lectura	276.71	240.70	-36.01	SÍ
	Ingreso Promedio	2036.19	514.30	-1521.89	SÍ
	Educ. Padre	17.12	11.30	-5.82	SÍ
	Educ. Madre	16.67	11.76	-4.91	SÍ
	Prop. Extranjeros	0.04	0.05	+0.01	SÍ

Tabla 15: Análisis descriptivo comparativo: Privado (2022).

Esta comparativa se realiza frente a los establecimientos particulares pagados.

- **Brecha Académica:** En 2016, los colegios privados ya superaban a los Bicentenarios en Matemática (-21.86), aunque en Lectura la diferencia no era significativa. En 2022, la distancia se disparó: los privados superan a los Bicentenarios por 59.04 puntos en Matemática y 36.01 en Lectura.
- **Abismo Socioeconómico:** Las tablas de ambos años muestran una brecha económica masiva (superior a 1,5 millones de pesos de diferencia en ingreso promedio) y una escolaridad de los padres mucho mayor en el sector privado (alrededor de 5 a 6 años de diferencia).
- **Significado:** Mientras que en 2016 los Bicentenarios lograban competir de cerca con los privados en Lectura bajo el modelo de selección, en 2022 la brecha se volvió crítica en todas

las áreas, evidenciando que el perfil de los Liceos Bicentenarios se ha alejado totalmente de los estándares de los colegios pagados tras la implementación del SAE.

### 7.3. Resultados del Modelo de Efectos Fijos de Dos Vías

En esta sección se presentan los resultados del modelo de efectos fijos de dos vías, utilizado como estrategia principal de inferencia causal para estimar el efecto promedio del Sistema de Admisión Escolar (SAE) sobre el desempeño académico de los Liceos Bicentenario. En coherencia con la metodología, se reportan estimaciones separadas para Matemática y Lectura, junto con especificaciones sin controles y con controles de composición del alumnado. El coeficiente de interés corresponde al término de interacción entre condición Bicentenario y periodo posterior a la implementación del SAE, el cual captura el cambio diferencial observado en estos establecimientos respecto del grupo de comparación.

#### 7.3.1. Resultados para Matemática

Grupo	Variable	Coficiente	Error Est.	Estadístico t	¿Sig.?
Matemática	Sin controles	2.17	1.56	1.39	NO
	Con controles	0.06	1.48	0.04	NO

Tabla 16: Resultados del modelo de efectos fijos de dos vías: Matemática.

Esta sección presenta la estimación del efecto promedio del SAE sobre el puntaje SIMCE promedio de Matemática a nivel de establecimiento.

■ **Resultados Académicos:** En la especificación sin controles, el coeficiente de interés es positivo pero no estadísticamente significativo, con un valor de 2.17 puntos. Una vez incorporados los controles de composición, el coeficiente se reduce a 0.06 puntos y continúa siendo no significativo.

En consecuencia, no se observa evidencia robusta de un cambio diferencial atribuible al SAE en el puntaje promedio de Matemática de los Liceos Bicentenario.

■ **Rol de los Controles:** La fuerte reducción del coeficiente al pasar desde la especificación sin controles a la especificación controlada sugiere que parte importante de la variación inicial en Matemática se relaciona con diferencias en la composición del alumnado más que con un efecto institucional propiamente tal.

■ **Lectura del Resultado:** En términos descriptivos, el modelo indica que la trayectoria de Matemática de los Liceos Bicentenario no presenta una alteración diferencial estadísticamente detectable una vez descontadas las características permanentes de las escuelas, los shocks comunes entre periodos y la composición promedio de la matrícula.

### 7.3.2. Resultados para Lectura

Grupo	Variable	Coefficiente	Error Est.	Estadístico t	¿Sig.?
Lectura	Sin controles	-4.10	1.19	-3.45	SÍ
	Con controles	-4.87	1.15	-4.24	SÍ

Tabla 17: Resultados del modelo de efectos fijos de dos vías: Lectura.

Esta sección reporta la estimación del mismo modelo para el puntaje SIMCE promedio de Lectura.

■ **Resultados Académicos:** A diferencia de lo observado en Matemática, el coeficiente de interés en Lectura es negativo y estadísticamente significativo tanto sin controles como con controles. En la especificación sin controles, el efecto estimado es de -4.10 puntos, mientras que con controles alcanza -4.87 puntos.

■ **Persistencia del Efecto:** La permanencia del signo negativo y su mayor magnitud absoluta tras

la incorporación de controles indican que la caída relativa en Lectura no desaparece al ajustar por composición socioeconómica y demográfica del establecimiento. Esto sugiere que la diferencia no responde únicamente a cambios en el perfil de la matrícula.

■ **Lectura del Resultado:** El modelo de efectos fijos muestra que, dentro del periodo analizado, los Liceos Bicentenario presentan una disminución relativa en Lectura frente al grupo de comparación. Dado que la especificación incorpora efectos fijos por establecimiento y por año, esta estimación debe interpretarse como un cambio promedio diferencial y no como una simple comparación de niveles brutos entre escuelas.

### 7.3.3. Controles de composición y asociaciones relevantes

Grupo	Variable	Coefficiente	Error Est.	Estadístico t	¿Sig.?
Matemática	Prom. Educ. Padre	4.58	0.72	6.32	SÍ
	Prom. Educ. Madre	4.60	0.77	5.96	SÍ
	DE Ingreso	-0.01	0.00	-1.81	NO
	Prop. Extranjeros	54.65	6.37	8.57	SÍ
Lectura	Prom. Educ. Padre	3.20	0.63	5.06	SÍ
	Prom. Educ. Madre	1.19	0.63	1.88	NO
	DE Ingreso	-0.01	0.00	-1.32	NO
	Prop. Extranjeros	39.31	5.59	7.04	SÍ

Tabla 18: Coeficientes relevantes de los controles en el modelo de efectos fijos de dos vías.

La Tabla 18 resume algunos de los controles más relevantes de la especificación con composición.

■ **Capital Familiar:** En ambas pruebas, los promedios de escolaridad del padre y de la madre presentan asociaciones positivas con los puntajes, lo que es consistente con la idea de que el capital cultural del hogar incide sobre el rendimiento institucional promedio.

■ **Heterogeneidad de la Matrícula:** La desviación estándar del ingreso presenta signo negativo

en ambos modelos, aunque no alcanza significancia estadística bajo el umbral convencional. Esto sugiere que una mayor dispersión socioeconómica podría estar asociada a menores puntajes, aunque la evidencia en esta especificación es acotada.

■ **Diversidad Demográfica:** La proporción de estudiantes extranjeros aparece con coeficiente positivo y estadísticamente significativo tanto en Matemática como en Lectura. Este resultado debe leerse con cautela, pues no implica por sí solo una relación causal, pero muestra que la composición demográfica del establecimiento forma parte de la estructura de asociación relevante para entender el desempeño promedio.

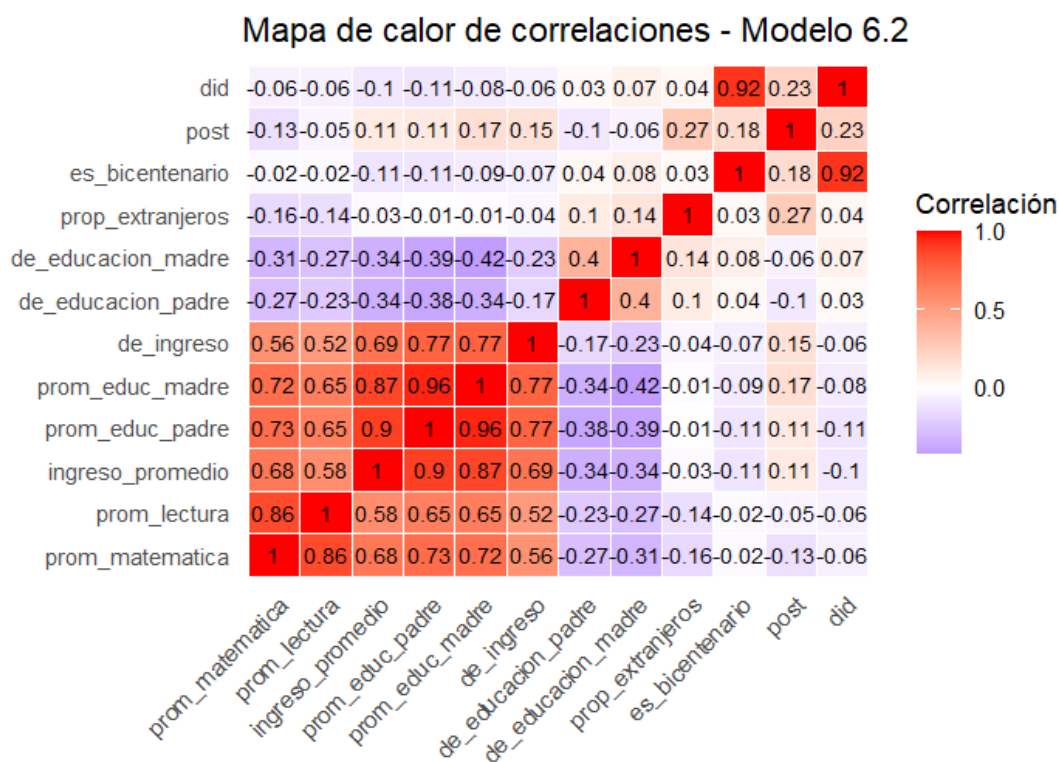


Figura 1: Mapa de Correlación para el Modelo de Efectos Fijos

El mapa de calor del modelo de efectos fijos muestra la estructura general de correlaciones entre las variables utilizadas en la especificación agregada por establecimiento. Se observan

asociaciones positivas elevadas entre los puntajes promedio y las variables de capital familiar, particularmente ingreso promedio y escolaridad de los padres, lo que es coherente con la relevancia de la composición socioeconómica en el rendimiento escolar. Asimismo, destaca la alta correlación entre la variable de tratamiento y la condición de establecimiento Bicentenario, lo cual es esperable por la propia construcción del término de interacción. En contraste, variables como la dispersión del ingreso y la proporción de estudiantes extranjeros presentan patrones menos intensos, lo que sugiere una relación más acotada o más compleja con el resto de los indicadores del modelo.

## 7.4. Resultados del Modelo Multinivel de Efectos Mixtos

Con el objeto de complementar la estrategia principal de identificación causal, esta sección presenta los resultados del modelo multinivel de efectos mixtos estimado sobre la base individual de estudiantes. Este enfoque permite incorporar simultáneamente variables a nivel alumno, hogar y establecimiento, además de un intercepto aleatorio por escuela, lo que resulta especialmente útil para analizar la composición del alumnado y la distribución de la heterogeneidad entre establecimientos.

### 7.4.1. Resultados académicos del modelo multinivel

Grupo	Variable	Coefficiente	Error Est.	Estadístico t	¿Sig.?
Académico	Puntaje Matemática	-1.23	0.52	-2.38	SÍ
	Puntaje Lectura	-4.91	0.47	-10.46	SÍ

Tabla 19: Resultados académicos del modelo multinivel de efectos mixtos.

Esta sección reporta el cambio diferencial estimado para estudiantes de Liceos Bicentenario en el periodo posterior al SAE, una vez controladas sus características individuales, familiares y

escolares.

■ **Resultados Académicos:** En Matemática, el coeficiente de interés es negativo y estadísticamente significativo, con un valor de -1.23 puntos. En Lectura, el efecto también es negativo, pero de mayor magnitud, alcanzando -4.91 puntos. En ambos casos, la evidencia apunta a una disminución relativa del rendimiento individual asociado a los Liceos Bicentenario en el periodo post-SAE.

■ **Comparación con el Modelo de Efectos Fijos:** El hallazgo para Lectura es consistente con el obtenido en el modelo agregado de efectos fijos, donde también se observó una caída negativa y significativa. En Matemática, en cambio, el modelo multinivel detecta un efecto negativo pequeño que no apareció con claridad en el modelo agregado, lo que sugiere que parte de la variación relevante se encuentra en el nivel individual y no únicamente en el promedio por establecimiento.

■ **Aporte del Enfoque Multinivel:** Esta especificación permite capturar mejor la estructura jerárquica de los datos y representar de manera más realista la relación entre características individuales y contexto escolar. Por ello, sus resultados complementan el análisis principal al mostrar cómo la composición del hogar y del establecimiento se vincula con el rendimiento del estudiante.

#### 7.4.2. Asociaciones relevantes del modelo multinivel académico

Grupo	Variable	Coefficiente	Error Est.	Estadístico t	¿Sig.?
Matemática	Ingresos del hogar	0.004	0.0002	17.76	SÍ
	Educ. Padre	1.07	0.04	30.37	SÍ
	Educ. Madre	1.01	0.04	27.58	SÍ
	Ingreso Promedio Est.	-0.003	0.001	-2.69	SÍ
	Prop. Extranjeros	41.72	2.59	16.13	SÍ
Lectura	Ingresos del hogar	0.002	0.0002	7.26	SÍ
	Educ. Padre	0.96	0.03	29.53	SÍ
	Educ. Madre	0.84	0.03	24.91	SÍ
	Ingreso Promedio Est.	0.00	0.001	-0.45	NO
	Prop. Extranjeros	23.97	2.29	10.45	SÍ

Tabla 20: Coeficientes relevantes del modelo multinivel para resultados académicos.

La Tabla 20 muestra algunas asociaciones relevantes entre variables individuales, contexto escolar y rendimiento.

■ **Variables del Hogar:** En ambas pruebas, el ingreso del hogar y la escolaridad parental presentan coeficientes positivos y estadísticamente significativos. Esto indica que los estudiantes con mayor capital económico y cultural tienden a obtener mejores puntajes individuales en ambas evaluaciones.

■ **Variables del Establecimiento:** En Matemática, el ingreso promedio del establecimiento aparece con signo negativo y significativo, mientras que en Lectura ese coeficiente no resulta significativo. En ambos modelos, la proporción de estudiantes extranjeros presenta una asociación positiva y significativa.

■ **Importancia del Contexto Escolar:** Estos resultados refuerzan la idea de que el rendimiento individual no depende únicamente de las condiciones del hogar, sino también del contexto compartido en el establecimiento. Por ello, el modelo multinivel resulta especialmente útil para articular el análisis de desempeño con la discusión sobre efecto par, composición y heterogeneidad escolar.

### 7.4.3. Resultados sobre composición del alumnado

Grupo	Variable	Coefficiente	Error Est.	Estadístico t	¿Sig.?
Composición	Ingreso del hogar	-35.73	4.13	-8.65	SÍ
	Educación del padre	-0.05	0.02	-2.25	SÍ
	Educación de la madre	0.05	0.02	2.06	SÍ
	Condición de extranjero	-0.01	0.00	-5.09	SÍ

Tabla 21: Resultados del modelo multinivel para composición del alumnado.

Esta sección utiliza el mismo enfoque multinivel, pero con variables dependientes de composición social y familiar del alumnado, con el fin de evaluar si el perfil del estudiantado Bicentenario cambió de manera sistemática tras la implementación del SAE.

■ **Cambios en el Perfil Socioeconómico:** El resultado más claro aparece en ingreso del hogar, donde el coeficiente estimado es de -35.73, indicando una caída estadísticamente significativa en esta dimensión. De forma similar, la escolaridad del padre presenta una disminución pequeña pero significativa de -0.05 años.

■ **Educación de la Madre:** La escolaridad de la madre muestra un coeficiente positivo, también significativo, de 0.05 años. Este resultado sugiere que los cambios en composición no fueron completamente homogéneos entre todas las dimensiones del capital familiar.

■ **Dimensión Demográfica:** La variable de condición de extranjero presenta un coeficiente negativo y significativo. Este resultado debe interpretarse con cautela, dado que se trata de una variable dicotómica modelada dentro de una estructura lineal multinivel, pero igualmente aporta información relevante sobre la composición demográfica del alumnado.

■ **Lectura Integrada:** En conjunto, estos resultados muestran que el periodo posterior al SAE no solo se asocia a variaciones en rendimiento, sino también a transformaciones observables en el perfil del estudiantado Bicentenario, especialmente en ingreso del hogar y escolaridad paterna.

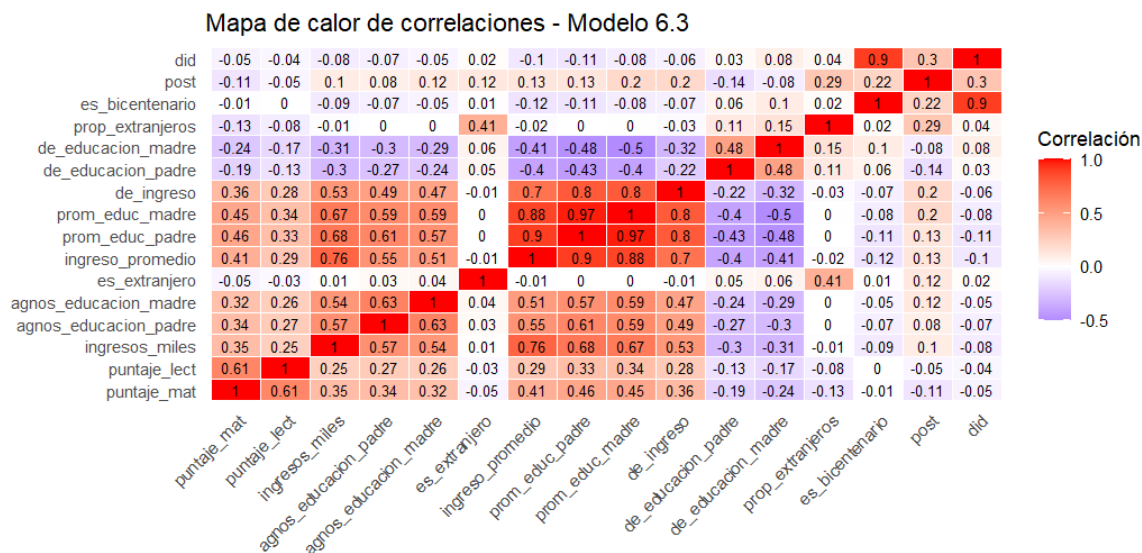


Figura 2: Mapa de Correlación para el Modelo Multinivel de Efectos Mixtos

El mapa de calor que se aprecia en la parte superior, correspondiente al modelo multinivel permite visualizar simultáneamente las correlaciones entre variables del estudiante, del hogar y del establecimiento. En la figura se aprecia un bloque de asociaciones altas entre los indicadores socioeconómicos individuales y sus equivalentes agregados a nivel escolar, mientras que los puntajes de Matemática y Lectura muestran una correlación positiva moderada entre sí y asociaciones también positivas con el capital económico y cultural del hogar. Del mismo modo, la variable de interacción del tratamiento presenta la relación esperada con la condición Bicentenario y el periodo post-SAE. En conjunto, la imagen refuerza la conveniencia del enfoque multinivel, ya que evidencia que el rendimiento individual se relaciona tanto con características propias del estudiante como con atributos compartidos del contexto escolar.

## 7.5. Resultados de segregación escolar: Coeficiente de Correlación Intraclase

La siguiente sección presenta los resultados del Coeficiente de Correlación Intraclase (ICC), calculado a partir de modelos nulos multinivel para distintas variables socioeconómicas y académicas. Este indicador permite estimar qué proporción de la varianza total se concentra entre establecimientos, por lo que una disminución del ICC entre 2016 y 2022 es compatible con una menor concentración entre escuelas.

Grupo	Variable	ICC 2016	ICC 2022	Variación	¿Cambio?
ICC	Ingresos del hogar	0.6747	0.6092	-0.0655	Disminuye
	Educación del padre	0.4146	0.3887	-0.0259	Disminuye
	Educación de la madre	0.3896	0.3423	-0.0473	Disminuye
	Puntaje Matemática	0.4804	0.3221	-0.1583	Disminuye
	Puntaje Lectura	0.3084	0.1947	-0.1137	Disminuye

Tabla 22: Coeficiente de Correlación Intraclase (ICC) por variable y año.

Esta tabla resume el grado de concentración entre establecimientos para variables de composición social y rendimiento académico antes y después del SAE.

■ **Segregación Socioeconómica:** Los ICC de ingresos del hogar y escolaridad parental disminuyen entre 2016 y 2022. En el caso del ingreso, el indicador pasa de 0.6747 a 0.6092, lo que sigue reflejando una alta concentración entre establecimientos, aunque menor que en el periodo previo. Las variables de educación del padre y de la madre también muestran reducciones, sugiriendo una menor segmentación entre escuelas según origen familiar.

■ **Segregación Académica:** Los puntajes SIMCE muestran una caída aún más marcada en su componente entre establecimientos. En Matemática, el ICC disminuye desde 0.4804 a 0.3221; en Lectura, desde 0.3084 a 0.1947. Esto implica que una proporción menor de la varianza total del

rendimiento se explica por diferencias entre escuelas en 2022.

■ **Lectura del Indicador:** Los resultados del ICC no implican desaparición de las desigualdades escolares, sino una menor proporción de esas desigualdades concentrada entre establecimientos. En este sentido, la evidencia es compatible con una reducción de la segregación socioeconómica y académica entre escuelas en el periodo posterior al SAE.

■ **Relación con los Resultados Previos:** La disminución del ICC conversa con los resultados del modelo multinivel para composición del alumnado y con la pérdida de ventaja relativa observada en Lectura. En conjunto, los hallazgos sugieren una menor concentración de perfiles semejantes entre escuelas, acompañada por transformaciones en el perfil social de los estudiantes y por cambios en los resultados académicos relativos.

## 8. Conclusiones

### 8.1. Conclusiones sobre el cambio en la composición del alumnado

Uno de los hallazgos más importantes de esta investigación es que la implementación del Sistema de Admisión Escolar sí se asocia a cambios relevantes en la composición del alumnado de los Liceos Bicentenario. Esta conclusión no surge de una sola tabla o de un solo modelo, sino de la convergencia entre la evidencia descriptiva, el modelo de efectos fijos y, especialmente, el modelo multinivel estimado a nivel de estudiante.

En los resultados longitudinales se observa que, entre 2016 y 2022, los Liceos Bicentenario experimentan una caída importante en sus puntajes promedio, junto con cambios estadísticamente significativos en variables del entorno familiar y demográfico. A nivel descriptivo, disminuyen los años de escolaridad promedio del padre y de la madre, y aumenta la proporción de estudiantes extranjeros. Aunque el ingreso promedio del establecimiento aumenta entre ambos años, este resultado debe interpretarse con cautela, ya que la variable de ingresos fue construida a partir de tramos y no representa una medición monetaria exacta. Por ello, el análisis de composición no puede descansar exclusivamente en una lectura aislada de esa variable.

La comparación transversal refuerza esta interpretación. En 2016, los Bicentenario se distinguían con claridad del sector municipal por sus mayores puntajes y por un perfil social relativamente más favorecido. En 2022, esa distancia se reduce de forma notoria. Las brechas con los municipales en ingreso promedio y escolaridad parental se estrechan, y al mismo tiempo los Bicentenario pierden buena parte de la ventaja académica que los caracterizaba. En términos sustantivos, esto sugiere que el SAE alteró las condiciones de entrada de estos establecimientos y debilitó la homogeneidad social que antes favorecía su funcionamiento bajo un esquema selectivo.

El modelo multinivel sobre composición del alumnado refuerza esta lectura. Allí, el ingreso del hogar presenta un coeficiente negativo y estadísticamente significativo de  $-35.73$ , y la escolaridad del padre también muestra una disminución pequeña pero significativa de  $-0.05$  años. Aunque la escolaridad de la madre aparece con signo positivo, con un coeficiente de  $0.05$ , el patrón agregado sigue apuntando a una transformación efectiva del perfil social de la matrícula posterior al SAE. Esto es metodológicamente relevante porque permite sostener que la discusión sobre resultados académicos no puede reducirse a una caída o mantención de puntajes, sino que debe considerar también el nuevo contexto social y familiar bajo el cual esos puntajes se producen.

Con todo, esta conclusión debe formularse con precisión. La investigación se propone analizar el perfil socioeconómico y académico del alumnado, pero empíricamente la dimensión mejor capturada es la socioeconómica y demográfica: ingreso del hogar, escolaridad parental y condición de extranjero. La dimensión académica de entrada no queda observada de forma directa en la base utilizada. Por ello, la afirmación más rigurosa es que el SAE modificó de manera clara el perfil social y familiar del alumnado Bicentenario, mientras que la transformación del perfil académico de entrada queda sugerida por la teoría del efecto par y por la pérdida de homogeneidad, más que medida directamente.

## **8.2. Conclusiones sobre los resultados descriptivos preliminares**

Los resultados descriptivos muestran una transformación muy marcada en la posición relativa de los Liceos Bicentenario dentro del sistema escolar chileno. En el análisis longitudinal, este grupo presenta una caída muy fuerte en sus puntajes promedio entre 2016 y 2022:  $-51,40$  puntos en Matemática y  $-35,27$  puntos en Lectura. Estas disminuciones son bastante mayores que las ob-

servadas en otros sectores. Los municipales, por ejemplo, no presentan cambios estadísticamente significativos en sus puntajes promedio; los subvencionados también caen, pero bastante menos; y los privados mantienen mayor estabilidad relativa, especialmente en Lectura.

La comparación transversal entrega una conclusión igualmente importante. En 2016, los Bicentenario mantenían amplias ventajas sobre los municipales y superaban también a los subvencionados en ambas pruebas. En 2022, esa ventaja se estrecha fuertemente frente al mundo municipal y desaparece, o incluso se revierte, frente al sector subvencionado. Respecto del sector privado, la distancia se amplía de forma considerable. Esto significa que el SAE no solo coincide con una caída absoluta en los resultados brutos, sino también con una pérdida de posición relativa en el mapa escolar chileno.

Desde un punto de vista interpretativo, estas pruebas descriptivas cumplen un papel fundamental, aunque no causal. Tal como se plantea en la metodología, las pruebas  $t$  permiten diagnosticar brechas y transformaciones, pero no separar por sí solas el efecto del SAE de otros shocks simultáneos, especialmente la pandemia y el retorno a la presencialidad en 2022. Por eso, el valor principal de esta evidencia no es demostrar causalidad, sino ordenar el problema empírico: primero, que los Bicentenario efectivamente cambiaron; segundo, que su cambio fue más intenso que el de otros grupos; y tercero, que dicho cambio ocurrió paralelamente a una mayor cercanía socioeconómica con sectores históricamente menos selectivos del sistema.

En consecuencia, los resultados descriptivos sí respaldan la idea de que el periodo post-SAE representa una etapa de redefinición para los Liceos Bicentenario. Lo que antes aparecía como un modelo de excelencia claramente diferenciado comienza a mostrar rasgos de convergencia con otros sectores, especialmente con el municipal, tanto en composición como en rendimiento bruto.

### 8.3. Conclusiones del modelo de efectos fijos de dos vías

El modelo de efectos fijos de dos vías constituye la estrategia principal de inferencia causal de esta memoria y permite extraer una conclusión central: la caída descriptiva observada en los Liceos Bicentenario no puede interpretarse de manera uniforme entre asignaturas. En Matemática, la evidencia causal promedio a nivel de establecimiento es débil una vez que se controla por composición. En Lectura, en cambio, la caída relativa persiste y aparece con solidez estadística.

En Matemática, el coeficiente de interés es 2.17 en la especificación sin controles y 0.06 en la especificación con controles, y en ninguno de los dos casos resulta estadísticamente significativo. Esto sugiere que, aunque descriptivamente los Bicentenario muestran una caída importante en sus puntajes brutos, el modelo no encuentra evidencia robusta de un cambio diferencial atribuible al SAE a nivel de establecimiento una vez descontadas las diferencias permanentes entre escuelas, los shocks comunes del periodo y la composición promedio de la matrícula. En otras palabras, la gran caída observada en Matemática parece estar fuertemente mediada por cambios en la composición del alumnado y por el contexto general del periodo analizado.

En Lectura, la conclusión es distinta. El coeficiente del tratamiento es -4.10 sin controles y -4.87 con controles, y ambos resultados son estadísticamente significativos. Además, la persistencia del signo negativo tras incorporar controles indica que la caída relativa en Lectura no desaparece al ajustar por composición socioeconómica y demográfica del establecimiento. Esto permite sostener que, dentro del marco analítico del modelo, sí existe evidencia consistente de una disminución relativa en Lectura para los Liceos Bicentenario en el periodo posterior al SAE.

Los controles del modelo también entregan información sustantiva relevante. En ambas pruebas, la escolaridad promedio del padre y de la madre muestra asociaciones positivas con el

rendimiento, lo que refuerza la importancia del capital cultural del hogar en los resultados académicos. La desviación estándar del ingreso aparece con signo negativo, aunque sin significancia estadística, mientras que la proporción de estudiantes extranjeros presenta coeficientes positivos y significativos en ambas asignaturas. Estos resultados no deben interpretarse causalmente por sí solos, pero sí muestran que la composición del establecimiento es una dimensión clave para comprender el desempeño promedio.

En síntesis, el modelo de efectos fijos permite concluir que el impacto del SAE sobre los Bicentenario no fue homogéneo entre áreas de aprendizaje. En Matemática, la evidencia causal promedio a nivel de establecimiento es frágil una vez controlada la composición. En Lectura, por el contrario, la caída relativa persiste y aparece con claridad. Esta diferencia entre asignaturas constituye uno de los hallazgos más importantes de la memoria y obliga a evitar conclusiones simplistas sobre una caída general e indistinta del desempeño académico.

#### **8.4. Conclusiones del modelo multinivel de efectos mixtos**

El modelo multinivel complementa de manera muy valiosa la estrategia principal, porque permite observar lo que ocurre cuando el análisis baja al nivel del estudiante y reconoce explícitamente la estructura jerárquica de los datos. Sus resultados académicos muestran un efecto negativo y significativo tanto en Matemática como en Lectura. En Lectura, el hallazgo es consistente con el modelo de efectos fijos: nuevamente aparece una disminución relativa importante. En Matemática, sin embargo, el modelo multinivel detecta un efecto negativo pequeño pero significativo que no había aparecido con claridad en el nivel agregado por establecimiento.

En términos concretos, el coeficiente de interés es -1.23 para Matemática y -4.91 para

Lectura. Esto sugiere que el SAE se asocia a una disminución relativa del rendimiento individual de los estudiantes de Liceos Bicentenario en el periodo posterior a su implementación, siendo el efecto más intenso en Lectura. La lectura sustantiva de esta diferencia es especialmente interesante: mientras el modelo agregado no detecta un efecto robusto en Matemática, el multinivel sí encuentra una señal negativa cuando se trabaja con información individual. Esto sugiere que parte de la variación relevante en esa asignatura podría estar ocurriendo dentro de los establecimientos y no solo en su promedio total.

Las asociaciones relevantes del modelo multinivel refuerzan, además, el peso del entorno familiar. En ambas pruebas, el ingreso del hogar y la escolaridad parental presentan coeficientes positivos y estadísticamente significativos. En Matemática, por ejemplo, el ingreso del hogar muestra un coeficiente de 0.004, mientras que la escolaridad del padre y de la madre alcanzan 1.07 y 1.01, respectivamente. En Lectura, los coeficientes correspondientes son 0.002, 0.96 y 0.84. Esto significa que el rendimiento individual sigue estando fuertemente relacionado con el capital económico y cultural del hogar, lo cual es coherente con la literatura revisada en el marco teórico y con la hipótesis de que el fin de la selección modifica también el contexto de pares en que se aprende.

En materia de composición, el multinivel también aporta una conclusión sustantiva: el periodo posterior al SAE no solo se relaciona con variaciones académicas, sino con transformaciones observables del estudiantado Bicentenario. Aquí destaca la caída en ingreso del hogar, con un coeficiente de -35.73, y la reducción en escolaridad paterna, con un coeficiente de -0.05. La señal sobre escolaridad materna, 0.05, y sobre condición de extranjero, -0.01, es más compleja y exige prudencia, porque no todas las dimensiones se mueven en una misma dirección y, además, algunas variables fueron modeladas linealmente pese a su naturaleza categórica o dicotómica. Por ello, la principal fortaleza de este modelo no está en ofrecer una conclusión simplista, sino en mostrar que

la transformación del alumnado posterior al SAE es real, aunque no perfectamente homogénea en todas sus dimensiones.

Por tanto, la conclusión general del modelo multinivel es doble. Primero, existe evidencia de una disminución relativa del rendimiento individual de los estudiantes de Liceos Bicentenario en el periodo post-SAE, especialmente en Lectura. Segundo, esa disminución ocurre paralelamente a una modificación de la composición social y familiar del alumnado. Esa combinación le da gran fuerza interpretativa al estudio, porque permite vincular resultados académicos y cambios de composición dentro de un mismo marco analítico. Al mismo tiempo, estos resultados deben leerse con cautela, porque el periodo 2022 también recoge el efecto acumulado de la pandemia y del retorno a la presencialidad, por lo que no puede interpretarse como una medición completamente aislada del SAE.

## **8.5. Conclusiones sobre segregación escolar**

La evidencia sobre segregación constituye otro de los aportes importantes de esta memoria. El Coeficiente de Correlación Intraclase disminuye entre 2016 y 2022 para todas las variables analizadas: ingresos del hogar, educación del padre, educación de la madre, puntaje de Matemática y puntaje de Lectura. Esto implica que una menor proporción de la varianza total de esas variables se explica por diferencias entre establecimientos en 2022.

En términos sustantivos, la evidencia es compatible con una reducción de la concentración entre escuelas y, por tanto, con una reducción parcial de la segregación socioeconómica y académica. Este hallazgo es consistente con el objetivo normativo central del SAE: disminuir la segmentación del sistema escolar chileno y promover una distribución menos concentrada de per-

files sociales entre establecimientos.

Sin embargo, la conclusión correcta no es que la segregación haya desaparecido. El ICC de ingresos del hogar sigue siendo alto en 2022, lo que significa que todavía existe una concentración importante de perfiles socioeconómicos entre establecimientos. Por consiguiente, el SAE parece haber contribuido a descomprimir parcialmente la distribución de estudiantes entre escuelas, pero no a eliminar las desigualdades estructurales que organizan históricamente el sistema escolar chileno.

Además, este resultado conversa bien con los demás hallazgos del estudio. La reducción del ICC coincide con el hecho de que los Bicentenario se acercan más al perfil del mundo municipal y pierden parte de su ventaja relativa en Lectura. Es decir, la menor segregación entre escuelas y la pérdida de homogeneidad en los Bicentenario parecen ser dos caras del mismo proceso: un sistema menos segmentado produce una distribución menos concentrada de perfiles, pero también exige a los liceos de excelencia operar en condiciones más heterogéneas que las que enfrentaban cuando podían seleccionar.

Desde una perspectiva de política pública, este hallazgo es particularmente importante porque sugiere que el SAE sí avanzó en su dimensión distributiva. No obstante, también muestra que ese avance ocurre junto con nuevas tensiones pedagógicas e institucionales. La reforma parece haber mejorado la mezcla entre escuelas, pero al mismo tiempo puso bajo presión a establecimientos que habían construido su identidad sobre una matrícula más homogénea.

## 8.6. Conclusiones respecto de las preguntas de investigación

Respecto de la primera pregunta de investigación, la conclusión es afirmativa: el perfil del alumnado de los Liceos Bicentenario cambió tras la implementación del SAE. La evidencia muestra modificaciones en ingreso del hogar, escolaridad parental y composición demográfica, además de un acercamiento relativo al perfil de los establecimientos municipales. La investigación, por tanto, sí respalda la idea de que el nuevo mecanismo de admisión alteró la composición social de estos liceos.

Respecto de la segunda pregunta, la respuesta debe ser matizada. Sí existe una relación entre el ingreso bajo reglas no selectivas y una disminución de los resultados académicos, pero esa relación no es idéntica en todas las áreas ni en todos los niveles de análisis. En Matemática, la evidencia descriptiva muestra una caída fuerte, pero el modelo de efectos fijos no encuentra un efecto causal robusto a nivel de establecimiento una vez controlada la composición; el multinivel, en cambio, sí detecta un efecto negativo pequeño a nivel individual. En Lectura, la evidencia es mucho más consistente: tanto el modelo de efectos fijos como el multinivel muestran una caída relativa significativa. En consecuencia, la respuesta más rigurosa es que el SAE se asocia con una pérdida académica más clara en Lectura, y con una señal más ambigua, aunque igualmente preocupante, en Matemática.

Respecto de la tercera pregunta, la evidencia sugiere que el SAE sí contribuyó a reducir la segregación escolar, pero solo de manera parcial. Los ICC caen en todas las dimensiones observadas, lo que es compatible con una menor concentración de perfiles sociales y académicos entre establecimientos. No obstante, la persistencia de niveles altos de concentración, sobre todo en ingresos, indica que el problema no fue resuelto completamente. Por ello, no sería correcto afirmar

que el SAE eliminó la segregación; más bien, la redujo en cierta medida y redistribuyó parte de la heterogeneidad entre escuelas.

Respecto de la cuarta pregunta, la investigación concluye que mantener estándares de excelencia académica en un sistema no selectivo sí es posible, pero con mayores tensiones y con una redefinición inevitable de lo que se entiende por excelencia. Los Bicentenario todavía conservan una ventaja respecto del sector municipal en 2022, pero esa ventaja es mucho menor que antes. Frente al sector subvencionado dejan de mostrar superioridad clara, y frente al sector privado se alejan todavía más. Por tanto, la excelencia basada en rendimiento bruto alto y matrícula homogénea se debilita bajo un sistema no selectivo. Lo que queda abierto es si puede emerger una excelencia distinta, más asociada a valor añadido que a selección previa.

En conjunto, las preguntas de investigación pueden responderse de manera coherente con los objetivos específicos del estudio. El SAE se asocia con una transformación de la composición del alumnado, con una pérdida relativa de rendimiento más clara en Lectura que en Matemática, y con una reducción parcial de la segregación escolar. Esa combinación muestra que la reforma tuvo efectos simultáneos sobre equidad, composición y resultados.

## **8.7. Comparación de resultados con las expectativas previas**

En términos generales, sí: una parte importante de los resultados era esperable a la luz del marco teórico y de la metodología propuesta. El documento se construye sobre la idea de que la selección previa contribuía a la homogeneidad social y académica de los Bicentenario, y de que el efecto par podía jugar un rol importante en sus resultados. También se reconoce desde el comienzo que el año 2022 está afectado por el shock de la pandemia y por el retorno a la presencialidad, lo

que ya anticipaba un contexto general de caída en puntajes. Desde esa perspectiva, era razonable esperar que el fin de la selección coincidiera con una mayor heterogeneidad del alumnado y con una pérdida de parte de la ventaja histórica de estos liceos.

Sin embargo, no todo era completamente predecible. Lo más interesante es que los resultados no apuntan a un colapso uniforme del modelo Bicentenario, sino a un efecto diferenciado según asignatura y nivel de análisis. No era obvio, por ejemplo, que en Matemática el modelo de efectos fijos fuera a mostrar ausencia de significancia mientras el multinivel sí detectara una señal negativa pequeña. Tampoco era evidente que, junto con la pérdida relativa en resultados, apareciera una reducción del ICC en todas las variables analizadas. Esto muestra que el SAE no produjo un efecto unidimensional.

Por eso, la mejor síntesis es que los resultados eran parcialmente esperables. Eran esperables en cuanto a la dirección general del cambio: más heterogeneidad, menor distancia con el mundo municipal y tensiones para sostener la excelencia basada en selección. Pero no eran del todo previsible en su forma específica, ni en la magnitud distinta de los efectos entre Matemática y Lectura, ni en la coexistencia de pérdida académica relativa y reducción parcial de la segregación. Precisamente allí radica uno de los aportes del estudio: mostrar que la discusión no es binaria entre inclusión o deterioro, sino más compleja y simultánea.

## **8.8. Conclusiones generales**

En conjunto, la investigación permite concluir que el Sistema de Admisión Escolar sí alteró de manera significativa el funcionamiento observable de los Liceos Bicentenario. Lo hizo modificando la composición social de su alumnado, reduciendo parte de su ventaja académica relativa y

disminuyendo parcialmente la segregación entre escuelas. Estos tres procesos aparecen conectados entre sí y deben leerse como parte de una misma transformación institucional.

La evidencia también muestra que no basta con observar puntajes promedio para evaluar el efecto de una política de admisión. Una parte de la caída en rendimiento bruto está asociada a cambios en la composición del alumnado, y el propio diseño metodológico de la memoria insiste en distinguir entre desempeño bruto y valor añadido. En este sentido, el estudio no demuestra simplemente que los Bicentenario empeoraron, sino que comenzaron a operar bajo condiciones muy distintas a las que les dieron origen. Eso obliga a interpretar sus resultados con mayor profundidad y a revisar cómo se define la excelencia escolar en contextos de mayor diversidad.

Desde una perspectiva de política pública, la conclusión de fondo es que el SAE parece haber avanzado en equidad distributiva, pero ese avance no está exento de costos, tensiones y desafíos para los establecimientos que habían construido su identidad sobre la selección. El verdadero problema no parece ser solo si bajaron o no bajaron los puntajes, sino si el sistema educativo chileno cuenta con herramientas pedagógicas, organizacionales y de apoyo suficientes para sostener altos estándares en un escenario de admisión más inclusivo.

En definitiva, esta memoria aporta evidencia para una discusión que muchas veces se ha planteado de forma simplificada. Sus resultados sugieren que la tensión entre equidad y calidad no desaparece con la reforma, sino que cambia de forma. El SAE parece haber contribuido a una distribución menos segmentada de los estudiantes, pero al mismo tiempo exige repensar las condiciones bajo las cuales puede sostenerse una educación pública de excelencia en un contexto de mayor heterogeneidad social y académica.

## 9. Limitaciones

### 9.1. Limitaciones del Diseño Temporal

La primera gran limitación del estudio es su estructura temporal. La comparación principal se realiza entre 2016 y 2022, lo que ofrece una lógica clara de pre y post reforma, pero al mismo tiempo deja a la investigación con un número acotado de cortes para examinar tendencias previas y posteriores. Esto limita la capacidad de verificar con mayor solidez el supuesto de tendencias paralelas que sustenta la interpretación causal del enfoque de diferencias en diferencias y del modelo de efectos fijos de dos vías. Con dos años principales de análisis, la inferencia debe formularse con cautela.

A ello se suma que 2022 no representa solo el periodo posterior a la implementación del SAE. También corresponde a un contexto extraordinario marcado por la pandemia, la interrupción de trayectorias escolares y el retorno a la presencialidad. Aunque la metodología incorpora efectos fijos de año para absorber shocks comunes, esa estrategia no elimina completamente la posibilidad de impactos heterogéneos entre tipos de establecimientos. Por ello, el coeficiente estimado debe entenderse como un efecto asociado al periodo post-SAE en un contexto sistémico complejo, y no como una medida puramente aislada de la reforma.

Esta limitación temporal afecta tanto la interpretación de los resultados descriptivos como la de los modelos econométricos. En particular, obliga a evitar afirmaciones demasiado tajantes sobre causalidad estricta y a presentar los hallazgos como evidencia consistente, pero condicionada por un entorno histórico excepcional.

## 9.2. Limitaciones de Identificación Causal

Si bien el modelo de efectos fijos de dos vías es la estrategia principal de inferencia causal del estudio, su alcance tiene límites que el propio diseño reconoce. La identificación descansa en comparar el cambio relativo de los Bicentenario respecto del grupo de comparación, una vez controladas diferencias fijas entre establecimientos y shocks comunes del tiempo. Sin embargo, esta estrategia no garantiza por sí sola que toda diferencia residual pueda atribuirse exclusivamente al SAE.

Persisten posibles fuentes de heterogeneidad no capturadas, especialmente si distintos tipos de liceos reaccionaron de manera desigual al contexto post-pandemia. Por ejemplo, establecimientos con distintas capacidades de gestión, recursos pedagógicos o trayectorias institucionales previas pudieron enfrentar de forma diferente tanto la vuelta a la presencialidad como los desafíos asociados a cohortes más heterogéneas. En ese sentido, el coeficiente estimado resume un cambio diferencial, pero no elimina por completo todas las posibles explicaciones concurrentes.

Además, dentro del grupo Bicentenario coexisten trayectorias potencialmente distintas, como los establecimientos históricos y los nuevos Bicentenario. La propia metodología señala que ambas trayectorias se integran en un único efecto promedio en la especificación principal. Eso significa que el parámetro estimado puede estar resumiendo experiencias institucionales heterogéneas bajo una sola medida promedio. En consecuencia, los resultados son informativos y relevantes, pero no deben leerse como si describieran una única realidad homogénea para todos los Bicentenario del país.

### **9.3. Limitaciones de la Base de Datos y de la Construcción Muestral**

Otra limitación importante proviene de la construcción misma de la base analítica. La investigación integra información de estudiantes, cuestionarios de padres y establecimientos a través de procesos de unión entre archivos. Ese procedimiento es metodológicamente correcto y necesario para responder a los objetivos del estudio, pero puede implicar pérdida de observaciones cuando existen registros incompletos, especialmente en los cuestionarios familiares.

Si esa pérdida de casos no fuese aleatoria, entonces podría introducir un sesgo de selección en la muestra final analizada. En otras palabras, es posible que los estudiantes con información completa no representen exactamente al universo total de quienes rindieron SIMCE en los años considerados. Aunque esta limitación no invalida el análisis, sí obliga a interpretar los resultados con prudencia y a reconocer que la muestra efectiva depende de la calidad y completitud de los registros disponibles.

También debe considerarse que la investigación trabaja con información secundaria, por lo que la calidad del análisis está condicionada por la calidad de la medición original. Esto significa que cualquier error de registro, inconsistencia en codificaciones o diferencia entre versiones de las bases puede trasladarse, al menos parcialmente, al resultado final.

### **9.4. Limitaciones asociadas a la Transformación de Variables**

Una limitación adicional del estudio se relaciona con la forma en que fueron operacionalizadas algunas variables socioeconómicas a partir de categorías cualitativas originales del cuestionario SIMCE. En particular, variables como ingreso familiar y escolaridad parental no se observan como magnitudes continuas exactas, sino como tramos o categorías, por lo que fue necesario trans-

formarlas en valores numéricos para hacer posible su uso en los modelos econométricos.

En el caso del ingreso familiar, la variable `ingresos_miles` fue construida mediante la asignación aleatoria de un valor dentro del intervalo correspondiente a la categoría de ingreso reportada por la familia. Si bien esta estrategia permite aproximar cuantitativamente el nivel socioeconómico del hogar, introduce una fuente de error de medición, ya que no recupera el ingreso exacto observado, sino una realización artificial dentro de un rango previamente definido.

Además, en este procedimiento se impuso un mínimo de 60 mil pesos y un máximo de 3 millones de pesos. Esta decisión implica que la variable no representa con precisión ni la cola inferior ni, especialmente, la cola superior de la distribución del ingreso. En particular, para los establecimientos particulares pagados, donde es más probable la presencia de hogares con ingresos por sobre el tope definido, esta construcción puede generar una subestimación del ingreso promedio efectivo del grupo. En consecuencia, las comparaciones basadas en ingreso promedio deben interpretarse como aproximaciones del nivel socioeconómico relativo de los establecimientos y no como mediciones exactas del ingreso monetario familiar.

Una consideración similar aplica a las variables de escolaridad parental, cuya conversión desde categorías educacionales a años de estudio también simplifica trayectorias formativas heterogéneas. Por ello, estas variables deben interpretarse como proxies razonables del capital cultural del hogar, útiles para el análisis estadístico, pero no exentas de pérdida de precisión respecto de la realidad individual de cada familia.

## 9.5. Limitaciones de Operacionalización Conceptual

Existe además una limitación de coherencia entre el alcance conceptual prometido y lo que efectivamente puede observarse en los datos. El problema de investigación y los objetivos mencionan el análisis del perfil socioeconómico y académico de los estudiantes, pero en la práctica la medición empírica se concentra mucho más en la dimensión socioeconómica y demográfica: ingreso del hogar, escolaridad parental y condición de extranjero.

El estudio no incorpora una medida directa del desempeño o preparación académica previa al ingreso al liceo, por lo que la dimensión académica del perfil de entrada queda más inferida que observada. Esto no constituye una falla metodológica grave, pero sí una delimitación importante del alcance real de las conclusiones. En rigor, la investigación capta mejor cómo cambió el perfil social del alumnado Bicentenario que cómo cambió su perfil académico de entrada.

A esto se suma que el planteamiento del problema menciona la posibilidad de considerar percepciones de actores involucrados, pero la investigación desarrollada es plenamente cuantitativa. Esta diferencia entre formulación inicial y alcance empírico no invalida el trabajo, pero sí conviene explicitarla con claridad, para evitar que el lector espere una dimensión cualitativa que finalmente no forma parte del análisis realizado.

## 9.6. Limitaciones del Modelo de Efectos Fijos de Dos Vías

El modelo de efectos fijos presenta fortalezas claras, pero también limitaciones asociadas a trabajar con información agregada por establecimiento. Al estimarse sobre un panel de escuelas, no puede representar explícitamente la estructura jerárquica de los datos ni separar con precisión la variación que ocurre entre estudiantes de una misma escuela de aquella que ocurre entre escuelas.

Esta es una restricción importante en una investigación cuyo marco teórico otorga gran importancia al efecto par, a la heterogeneidad del aula y a la composición del alumnado. En ese sentido, el modelo captura adecuadamente cambios promedio a nivel institucional, pero no permite examinar con la misma precisión cómo se distribuyen esos cambios dentro de cada establecimiento.

Además, existe una pequeña tensión entre la formulación teórica inicial y la implementación empírica posterior. En el marco teórico, la estrategia de diferencias en diferencias se presenta originalmente a nivel de estudiante, mientras que en la metodología aplicada la especificación principal se desarrolla a nivel agregado de establecimiento. La adaptación es defendible y coherente con la estructura del panel finalmente construido, pero conviene reconocerla explícitamente para reforzar la transparencia metodológica del estudio.

## **9.7. Limitaciones del Modelo Multinivel**

El modelo multinivel de efectos mixtos resuelve parte del problema anterior al trabajar con información individual y escolar al mismo tiempo, pero tampoco elimina todas las restricciones de la investigación. El propio documento señala que este enfoque debe entenderse como complementario y no como sustituto de la estrategia principal de identificación causal. Su fortaleza está en modelar la heterogeneidad y la segregación, no en aislar de manera definitiva el efecto del SAE frente a todos los shocks concurrentes del periodo.

También hay que considerar que algunas variables incluidas en este modelo, como la condición de extranjero, son dicotómicas y se estiman dentro de una estructura lineal multinivel. Esto obliga a interpretar ciertos coeficientes con especial cautela, especialmente cuando se trata de resultados de composición y no de rendimiento. Del mismo modo, aunque el multinivel permite

cuantificar mejor la variación entre establecimientos, sigue estando condicionado por el hecho de que el estudio solo dispone de dos años principales de observación.

Por ello, el valor del modelo multinivel en esta memoria es principalmente analítico e interpretativo. Aporta profundidad para comprender mecanismos de composición y segregación, pero no debe ser presentado como una solución definitiva a todos los desafíos de identificación causal presentes en el estudio.

## **9.8. Cierre del Capítulo**

En síntesis, las limitaciones identificadas no invalidan los resultados de la memoria, pero sí delimitan con precisión su alcance. La investigación ofrece evidencia consistente y valiosa sobre cómo el SAE se relaciona con cambios en la composición del alumnado, con variaciones en el desempeño relativo de los Liceos Bicentenario y con una reducción parcial de la segregación entre escuelas. Sin embargo, esos hallazgos deben interpretarse como estimaciones prudentes y contextualizadas, condicionadas por la estructura temporal disponible, por el contexto post-pandemia y por las restricciones propias de las bases de datos utilizadas.

Reconocer estas limitaciones no debilita el estudio; por el contrario, lo vuelve metodológicamente más sólido. Explicitar aquello que los datos y los modelos permiten afirmar, y también aquello que todavía no permiten responder por completo, fortalece la credibilidad del análisis y abre espacio para futuras investigaciones que profundicen en dimensiones que aquí solo pudieron abordarse parcialmente.

En particular, futuras extensiones podrían incorporar más años de observación, distinguir con mayor claridad entre Bicentenario históricos y nuevos, incluir medidas más directas del perfil



académico de entrada y complementar el enfoque cuantitativo con evidencia cualitativa sobre la experiencia institucional de las comunidades escolares. De ese modo, sería posible enriquecer aún más la discusión sobre cómo compatibilizar inclusión, diversidad y excelencia en la educación pública chilena.

## Anexos

Variable	Código SIMCE	Rango de Imputación Aleatoria (Miles de \$)
Ingreso Familiar	1	Entre 60 y 99
	2	Entre 100 y 200
	3	Entre 201 y 300
	4	Entre 301 y 400
	5	Entre 401 y 500
	6	Entre 501 y 600
	7	Entre 601 y 800
	8	Entre 801 y 1.000
	9	Entre 1.001 y 1.200
	10	Entre 1.201 y 1.400
	11	Entre 1.401 y 1.600
	12	Entre 1.601 y 1.800
	13	Entre 1.801 y 2.000
	14	Entre 2.001 y 2.200
	15	Entre 2.201 y 3.000

Tabla 23: Criterios de imputación estocástica para la variable Ingreso Familiar.

Variable	Código SIMCE	Años de Escolaridad Asignados / Imputados
Escolaridad Parental***	1	0 años (Asignación directa)
	2 al 14**	Valor del código menos 1 (De 1 a 12 años)*
	15	Entre 12 y 13 años (Aleatorio)
	16	Entre 14 y 15 años (Aleatorio)
	17	Entre 14 y 16 años (Aleatorio)
	18	Entre 17 y 19 años (Aleatorio)
	19	Entre 18 y 21 años (Aleatorio)
	20	Entre 22 y 25 años (Aleatorio)

Tabla 24: Criterios de conversión e imputación estocástica para los años de escolaridad de los padres.

\*Nota: Los códigos del 2 al 14 corresponden a cursos exactos de enseñanza básica y media, por lo que se asignó el valor lineal correspondiente sin aleatorización.

\*\*Nota: Los códigos 13 y 14 ambos asignan 12 años, ya que uno es enseñanza media científico-humanista y la otra es enseñanza media en escuela técnica.

\*\*\*Nota: Los parámetros son los mismos para la escolaridad del padre y de la madre.

Variable	prom_matematica	prom_lectura	ingreso_promedio	prom_educ_padre	prom_educ_madre	de_ingreso	de_educacion_padre	de_educacion_madre	prop_extranjeros	es_bicentenario	post	did
did	-0.06	-0.06	-0.1	-0.11	-0.08	-0.06	0.03	0.07	0.04	0.92	0.23	1
post	-0.13	-0.05	0.11	0.11	0.17	0.15	-0.1	-0.06	0.27	0.18	1	0.23
es_bicentenario	-0.02	-0.02	-0.11	-0.11	-0.09	-0.07	0.04	0.08	0.03	1	0.18	0.92
prop_extranjeros	-0.16	-0.14	-0.03	-0.01	-0.01	-0.04	0.1	0.14	1	0.03	0.27	0.04
de_educacion_madre	-0.31	-0.27	-0.34	-0.39	-0.42	-0.23	0.4	1	0.14	0.08	-0.06	0.07
de_educacion_padre	-0.27	-0.23	-0.34	-0.38	-0.34	-0.17	1	0.4	0.1	0.04	-0.1	0.03
de_ingreso	0.56	0.52	0.69	0.77	0.77	1	-0.17	-0.23	-0.04	-0.07	0.15	-0.06
prom_educ_madre	0.72	0.65	0.87	0.96	1	0.77	-0.34	-0.42	-0.01	-0.09	0.17	-0.08
prom_educ_padre	0.73	0.65	0.9	1	0.96	0.77	-0.38	-0.39	-0.01	-0.11	0.11	-0.11
ingreso_promedio	0.68	0.58	1	0.9	0.87	0.69	-0.34	-0.34	-0.03	-0.11	0.11	-0.1
prom_lectura	0.86	1	0.58	0.65	0.65	0.52	-0.23	-0.27	-0.14	-0.02	-0.05	-0.06
prom_matematica	1	0.86	0.68	0.73	0.72	0.56	-0.27	-0.31	-0.16	-0.02	-0.13	-0.06

Tabla 25: Matriz de Correlaciones completa para el Modelo de Efectos Fijos de Dos Vías (Modelo 6.2).

Variable	puntaje_mat	puntaje_lect	ingresos_miles	agnos_educacion_padre	agnos_educacion_madre	es_extranjero	ingreso_promedio	prom_educ_padre	prom_educ_madre	de_ingreso	de_educacion_padre	de_educacion_madre	prop_extranjeros	es_bicentenario	post	did
did	-0.05	-0.04	-0.08	-0.07	-0.05	0.02	-0.1	-0.11	-0.08	-0.06	0.03	0.08	0.04	0.9	0.3	1
post	-0.11	-0.05	0.1	0.08	0.12	0.12	0.13	0.13	0.2	0.2	-0.14	-0.08	0.29	0.22	1	0.3
es_bicentenario	-0.01	0	-0.09	-0.07	-0.05	0.01	-0.12	-0.11	-0.08	-0.07	0.06	0.1	0.02	1	0.22	0.9
prop_extranjeros	-0.13	-0.08	-0.01	0	0	0.41	-0.02	0	0	-0.03	0.11	0.15	1	0.02	0.29	0.04
de_educacion_madre	-0.24	-0.17	-0.31	-0.3	-0.29	0.06	-0.41	-0.48	-0.5	-0.32	0.48	1	0.15	0.1	-0.08	0.08
de_educacion_padre	-0.19	-0.13	-0.3	-0.27	-0.24	0.05	-0.4	-0.43	-0.4	-0.22	1	0.48	0.11	0.06	-0.14	0.03
de_ingreso	0.36	0.28	0.53	0.49	0.47	-0.01	0.7	0.8	0.8	1	-0.22	-0.32	-0.03	-0.07	0.2	-0.06
prom_educ_madre	0.45	0.34	0.67	0.59	0.59	0	0.88	0.97	1	0.8	-0.4	-0.5	0	-0.08	0.2	-0.08
prom_educ_padre	0.46	0.33	0.68	0.61	0.57	0	0.9	1	0.97	0.8	-0.43	-0.48	0	-0.11	0.13	-0.11
ingreso_promedio	0.41	0.29	0.76	0.55	0.51	-0.01	1	0.9	0.88	0.7	-0.4	-0.41	-0.02	-0.12	0.13	-0.1
es_extranjero	-0.05	-0.03	0.01	0.03	0.04	1	-0.01	0	0	-0.01	0.05	0.06	0.41	0.01	0.12	0.02
agnos_educacion_madre	0.32	0.26	0.54	0.63	1	0.04	0.51	0.57	0.59	0.47	-0.24	-0.29	0	-0.05	0.12	-0.05
agnos_educacion_padre	0.34	0.27	0.57	1	0.63	0.03	0.55	0.61	0.59	0.49	-0.27	-0.3	0	-0.07	0.08	-0.07
ingresos_miles	0.35	0.25	1	0.57	0.54	0.01	0.76	0.68	0.67	0.53	-0.3	-0.31	-0.01	-0.09	0.1	-0.08
puntaje_lect	0.61	1	0.25	0.27	0.26	-0.03	0.29	0.33	0.34	0.28	-0.13	-0.17	-0.08	0	-0.05	-0.04
puntaje_mat	1	0.61	0.35	0.34	0.32	-0.05	0.41	0.46	0.45	0.36	-0.19	-0.24	-0.13	-0.01	-0.11	-0.05

Tabla 26: Matriz de Correlaciones completa para el Modelo Multinivel de Efectos Mixtos (Modelo 6.3).

## Referencias

- Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-proofness versus efficiency in matching with indifference: Redesigning the nyc high school match. *American Economic Review*, 99(5):1954–1978.
- Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. (2005). The boston mechanism. *The Quarterly Journal of Economics*, 120(2):675–713.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Bellei, C. (2013). El estudio de la segregación socioeconómica y académica de la educación chilena. *Estudios Pedagogicos*, 39.
- Carrasco, A. and Honey, N. (2019). Nuevo sistema de admisión escolar y su capacidad de atenuar la desigualdad de acceso a colegios de calidad: al inicio de un largo camino. *Estudios en Justicia Educacional*, 1.
- Carrasco, A., Villalobos, C., Gutiérrez, G., Honey, N., Dibona, G., Pereira, S., Rostagno, F., Buzeta, S., and Treviño, E. (2024). Evaluación contextualizada e histórica de los liceos bicentenario de excelencia. Centro de Políticas y Prácticas en Educación (CEPPE UC) y Centro de Justicia Educacional.
- Correa, J., Epstein, N., Epstein, R., Escobar, J., Rios, I., Aramayo, N., Bahamondes, B., Bonet, C., Castillo, M., Cristi, A., Epstein, B., and Subiabre, F. (2022). School choice in chile. *Operations Research*, 70(2):1066–1087.

- Eyzaguirre, B. and Fontaine, L. (1999). ¿qué mide realmente el SIMCE? *Estudios Públicos*, (75):107–161. Invierno 1999.
- Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.
- González, R., Valenzuela, J. P., and Bellei, C. (2017). Segregación educativa en el sistema chileno desde una perspectiva comparada. *El primer gran debate de la reforma educacional: Ley de Inclusión Escolar*.
- Gómez, M. C. (2014). En camino a la excelencia y calidad en un liceo bicentenario. *Revista de Estudios Cotidianos*, pages 207 – 232.
- Himmel, E. (1997). Impacto social de los sistemas de evaluación del rendimiento escolar: el caso de Chile. In Álvarez, B. and Ruiz-Casares, M., editors, *Evaluación y Reforma Educativa: Opciones de Política*. PREAL, Santiago.
- Hoxby, C. M. (2000). Does competition among public schools benefit students and taxpayers? *The American Economic Review*, 90(5):1209–1238.
- Hsieh, C.-T. and Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics*, 90(8-9):1477–1503.
- Lubienski, C., Gulosino, C., and Weitzel, P. (2009). School choice and competitive incentives: Mapping the distribution of educational opportunities across local education markets. *American Journal of Education*, 115.

OCDE (2017). *Educación en Chile: Evaluaciones de Políticas Nacionales de Educación*. OECD Publishing.

Padilla, G., Suazo, C., and Rodríguez, C. (2023). Seguimiento al sistema de admisión escolar en Chile: ¿cómo ha funcionado el aseguramiento de la inclusión? *Revista de la Escuela de Ciencias de la Educación*, 1(18).

Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, volume 1. sage.

Roth, A. E. and Sönmez, T. (2005). The evolution of the college admissions market. *Journal of Political Economy*, 113(4):781–810.

Sillard, M., Garay, M., and Troncoso, I. (2018). Análisis al nuevo sistema de admisión escolar en Chile: Región de Magallanes como experiencia piloto. *Calidad en la Educación*, pages 112 – 136.

Valenzuela, J. P., Bellei, C., and De Los Ríos, D. (2010). Segregación escolar en Chile. *Fin de ciclo*, pages 209–229.

Villar, P. (2020). Efectividad de liceos bicentenario de excelencia en Chile y su acceso a la elite de la educación superior. Master's thesis, Universidad de Chile - Facultad de Economía y Negocios.

Villaseca, M. (2022). Análisis general del sistema de admisión escolar (sae) 2019-2021. *acciónE-ducación*.