



“Efectividad de los Liceos Públicos de Excelencia en Chile”

**TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN POLÍTICAS PÚBLICAS**

**Alumno:
Claudio Ignacio Allende González**

**Profesor Guía:
Juan Pablo Valenzuela Barros**

Santiago, Abril 2015

Efectividad de los Liceos Públicos Selectivos de Excelencia en Chile¹

Claudio Allende González

Profesor Guía: Juan Pablo Valenzuela

Abril, 2015

Resumen

Esta investigación buscó estimar cuál es la magnitud y las causas del efecto asociado a los liceos selectivos públicos de excelencia, a través de metodologías cuasi experimentales. Los resultados –al analizar dos cohortes de alumnos– muestran que existiría un considerable efecto positivo y significativo a favor de estos colegios –entre 0,2 y 0,3 desviaciones estándar– resultados que serían consistentes en términos de magnitud y sentido con el efecto encontrado para colegios selectivos en otros países en vías de desarrollo. Presentando, además, una alta heterogeneidad a lo largo de la distribución de puntajes (percentiles más bajos tienen un mayor efecto) como en los resultados por nivel socioeconómico (alumnos de menor NSE serían los que presentan un mayor efecto), así como también un efecto diferenciado entre lenguaje y matemática. Por otra parte, la evidencia indica que en la medida que el promedio de otras alternativas de educación secundaria van mejorando su desempeño, la brecha respecto de los liceos altamente selectivos se va atenuando en el tiempo.

Adicionalmente, se encontró que la selección de los mejores estudiantes (mayores capacidades, habilidad, motivación, etc.) y su agrupación sería lo que explica la efectividad de este tipo de establecimientos; esto plantea que el efecto de las metodologías de enseñanza y procesos internos parecieran estar siendo facilitados por el tipo de alumnos que reciben.

Los resultados de esta investigación relevan la importancia potencial que el efecto par puede tener en la modificación de las trayectorias educativas de los alumnos en un país como Chile.

¹ El autor agradece el apoyo del proyecto Anillo SOC-1104 y al proyecto Fondo BASAL FB003 CIAE

I. INTRODUCCIÓN	5
II. REVISIÓN DE LITERATURA	8
1) REVISIÓN DE LITERATURA INTERNACIONAL	8
2) REVISIÓN DE LITERATURA NACIONAL	14
III. DATOS	16
IV. METODOLOGÍA	18
1) DEFINICIÓN DEL TRATAMIENTO Y DEL GRUPO TRATAMIENTO, LOS LICEOS PÚBLICOS SELECTIVOS DE EXCELENCIA.	20
2) DEFINICIÓN DEL GRUPO DE CONTROL	22
I. CALIDAD DEL GRUPO DE CONTROL	23
3) ESTIMACIÓN DEL EFECTO PROMEDIO DEL TRATAMIENTO SOBRE LOS TRATADOS	23
A. ESTIMACIÓN DE LA PROBABILIDAD DE RECIBIR EL TRATAMIENTO, EL PROPENSITY SCORE	24
B. METODOLOGÍAS BASADAS EN LA ESTIMACIÓN DEL PROPENSITY SCORE	25
C. ESTIMACIÓN POR DIFERENCIAS EN DIFERENCIAS (DID) Y CHANGES IN CHANGES (CIC)	30
D. VARIABLES UTILIZADAS	31
E. ESTIMACIÓN DE LA VARIANZA DE ATT Y OTRAS CONSIDERACIONES	33
V. PRINCIPALES RESULTADOS	35
A). QUIÉNES ASISTEN A LOS LICEOS PÚBLICOS DE EXCELENCIA	35
B). ¿OPORTUNIDAD PARA LOS NIÑOS VULNERABLES?	35
C). ¿OPORTUNIDAD PARA LOS ESTUDIANTES DE ESCUELAS MUNICIPALES?	37
D). IMPACTO DEL TRATAMIENTO EN LA PRUEBA SIMCE DE LOS ESTUDIANTES QUE ASISTEN A LOS LSPE	38
A) EFECTO PROMEDIO DE LA PARTICIPACIÓN EN LOS LSPE SOBRE LOS TRATADOS, ESTIMADORES BASADOS EN EL PROPENSITY SCORE	38
B) EFECTO PROMEDIO DE LA PARTICIPACIÓN EN LOS LSPE SOBRE LOS TRATADOS, ESTIMADOR CHANGES IN CHANGES (CIC)	45
C) EFECTO PROMEDIO DE LA PARTICIPACIÓN EN LOS LSPE SOBRE LOS TRATADOS (ATT) DIFERENCIADO POR NIVEL SOCIOECONÓMICO.	47
D) ¿A QUÉ SE DEBEN LAS DIFERENCIAS OBSERVADAS A FAVOR DE LOS LSPE?	50
VI. CONCLUSIONES	55
VII. BIBLIOGRAFÍA	58

<u>ANEXO 1: LICEOS PÚBLICOS SELECTIVOS DE EXCELENCIA</u>	65
<u>ANEXO 2: MEJORA EN EL BALANCE DE LAS VARIABLES MEDIDAS ANTES DEL TRATAMIENTO</u>	66
<u>ANEXO 3: MEJORA EN EL BALANCE DE LAS VARIABLES MEDIDAS ANTES DEL TRATAMIENTO</u>	68
<u>ANEXO 4: ESTADÍSTICAS DESCRIPTIVAS PANEL 2002-2008</u>	70
<u>ANEXO 5: ESTADÍSTICAS DESCRIPTIVAS PANEL 2006-2012</u>	72
<u>ANEXO 6: DISTRIBUCIÓN DE LOS GRUPOS DE CONTROL Y TRATAMIENTO EN LA REGIÓN DE SOPORTE COMÚN</u>	74
<u>ANEXO 7: PROBABILIDAD DE INGRESAR A UN LSPE CON GRUPO DE CONTROL EX COMPAÑEROS EMPAREJADOS</u>	76
<u>ANEXO 8: ESTIMACIÓN DE ATT SOLO PARA LA BASE REDUCIDA CON GRUPO DE CONTROL EX COMPAÑEROS EMPAREJADOS</u>	78

I. Introducción

La selección académica de estudiantes en educación básica para cursar sus estudios desde 7mo básico en Liceos Selectivos Públicos de Excelencia (LSPE) ha sido bastante tradicional en el sistema escolar chileno, sin embargo, las nuevas políticas implementadas o propuestas por los gobiernos recientes han puesto en el debate esta práctica de gestión educativa.

En un contexto histórico, la creación de los primeros liceos públicos académicamente selectivos respondió a la primera oferta de educación pública en la creación de la República, pero luego a la baja cobertura de la educación secundaria a nivel nacional, luego –en las décadas del 30s y 40s– la creación de Liceos Experimentales, y posteriormente a la iniciativa de Liceos Montegrando (en los 90s), experiencias cuyo foco principal era la generación de nuevo conocimiento y desarrollo de las didácticas para mejorar el proceso de enseñanza-aprendizaje que deben realizar los docentes, así como experimentar modalidades que pudieran ser replicadas en otros establecimientos, equipos directivos y de docentes del país.-

Recientemente, el proyecto “Liceos Bicentenario de Excelencia” ha sido una política sin precedentes en la historia de Chile de expansión de este tipo de colegios. Esta consistió en la conformación de 60 Liceos Públicos –aunque algunos de ellos son particulares subvencionados- de Excelencia (denominados Bicentenario) entre los años 2011-2012, patrocinada por el Ministerio de Educación (MINEDUC), siendo la mayor parte de dependencia municipal. Estos colegios podían ser preexistentes –modificados para cumplir con las condiciones requeridas por el MINEDUC (entregar una cobertura mínima de 7mo básico a 4to medio, tener un equipo directivo de excelencia y comprometerse con altos resultados en las pruebas estandarizadas de evaluación académica más reconocidas en el país, SIMCE y PSU)– o completamente nuevos. A diferencia de las experiencias del siglo XX, el objetivo principal de esta política ha sido la selección de los estudiantes más talentosos de 6º básico, bajo la premisa de otorgar una oportunidad de movilidad académica y social a los estudiantes más desaventajados (vulnerables) y de alto potencial académico de las escuelas municipales.

Por otra parte, la nueva ley de Inclusión Escolar vigente desde el año 2016, que entre otras cosas, busca terminar con la selección existente en el sistema escolar (tanto académica como socioeconómica) legitima la selectividad académica existente en este tipo de colegios por al menos 5 años más, *“Respecto de aquellos establecimientos educacionales que de acuerdo a un reglamento dictado por el Ministerio de Educación, cumplan con características históricas, de rendimiento académico destacado dentro de su región, que sean gratuitos, que presenten una demanda considerablemente mayor a sus vacantes y que hayan establecido procedimientos de selección académica a la fecha de publicación de*

la ley, iniciarán sus procesos de admisión, de acuerdo a la gradualidad territorial a que se refieren los incisos anteriores y de conformidad a los siguientes porcentajes”, luego agregan, “Dichos establecimientos educacionales podrán admitir a sus estudiantes realizando sus pruebas de admisión de la siguiente forma: (i) El primer año para el 85% de sus cupos; (ii) El segundo año para el 70% de los cupos; (iii) El tercer año para el 50% de los cupos; (iv) El cuarto año para el 30% de los cupos; (v) El quinto año no se podrán realizar pruebas de admisión” (Extracto proyecto de ley que regula la admisión de los y las estudiantes, elimina el financiamiento compartido y prohíbe el lucro en establecimientos educacionales que reciben aportes del Estado. Proyecto de Ley en trámite, 28 de Enero de 2015).

Este tipo de políticas, aplicadas o propuestas en Chile durante el último tiempo, generan una interrogante que no ha sido completamente resuelta por la literatura, tanto nacional como internacional, sobre los potenciales efectos derivados de la masificación de este tipo de colegios muy selectivos académicamente y que no han sido consideradas al momento de proponer políticas como las mencionadas.

Dos críticas principales se han hecho a los Liceos Selectivos Públicos de Excelencia, principalmente debido a la masificación de estos a través de la política “Liceos Bicentenario”. La primera hace referencia a que este tipo de establecimientos promueve el mejoramiento educativo de un grupo reducido de estudiantes, pero ello se explicaría sólo por una estrategia hiperselectiva (“descreme”) de los estudiantes más talentosos, pero sin un valor agregado del propio establecimiento (La Tercera, 2010; CEPPE, 2010); la segunda crítica relevante es que esta estrategia de masificación de colegios hiperselectivos podría perjudicar a los colegios de donde provengan estos estudiantes más talentosos, es decir, al grupo de estudiantes que se quedarán en los colegios de donde provienen los seleccionados de los Liceos de Excelencia, puesto que ellos perderán compañeros talentosos con los cuales lograban aprendizajes de mayor calidad (eg. por medio del efecto par o de apoyo especializado de éstos a los estudiantes con mayores dificultades), así como directivos y docentes destacados.

En este sentido la evidencia empírica con respecto a las dos interrogantes planteadas –como se verá con mayor profundidad más adelante– es bastante reciente y muestra, para la primera interrogante, evidencia mixta –dependiendo del sistema educacional y metodología para la estimación utilizada– (Cullen y Jacob, 2009; Dobbie y Fryer, 2014; Abdulkadiroglu, Angrist y Pathak, 2014; Bifulco, Ladd y Ross, 2008), y para la segunda es aún más escasa, mostrando un efecto negativo y pequeño sobre aquellos que pierden a los mejores alumnos (Walsh, 2009; Altonji, Huang y Taber, 2014).

A pesar de las importantes críticas levantadas a la iniciativa y a la centenaria existencia en Chile de Liceos Públicos de Excelencia, la evidencia empírica a nivel nacional con respecto a la primera interrogante es extremadamente escasa y poco informativa sobre los efectos, positivos o negativos asociados a este tipo de establecimientos (Bucarey et al, 2014; Valenzuela y Allende, 2011); y con respecto a la segunda interrogante es totalmente inexistente.

Por lo cual, este trabajo pretende complementar la escasa evidencia respecto de la primera interrogante, utilizando para ello la evidencia acumulada del desempeño observado en los liceos públicos de excelencia más tradicionales y antiguos –muchas veces denominados emblemáticos-, utilizando para ello más de una cohorte de análisis, de tal forma de observar la robustez de dichos resultados.

La presente investigación se organiza de la siguiente forma, después de esta Introducción la sección 2 muestra la revisión bibliográfica; la sección 3, presenta los datos utilizados; la sección 4 describe la metodología utilizada en esta investigación; la sección 5 muestra los principales resultados; y finalmente, la sección 6 resume las principales conclusiones obtenidas en esta investigación.

II. Revisión de literatura

1) Revisión de literatura internacional

A nivel internacional se reconoce una amplia diversidad de mecanismos de separación o agrupación de los estudiantes² debido a sus capacidades académicas, en este sentido algunos de los métodos más conocidos y utilizados son denominados tracking o streaming³ y sorting⁴. El foco de esta investigación se centra en las modalidades donde los estudiantes son separados en diversos establecimientos de acuerdo a su desempeño académico previo y/o su desempeño en pruebas selectivas realizadas por estos establecimientos, en este sentido la literatura comparada sobre este tipo de políticas es bastante amplia y diversa en sus resultados.

Con respecto a la selección de alumnos en los sistemas escolares a nivel comparado la evidencia muestra que esta es bastante generalizada, solo cambiando la forma y alcance de dichas políticas; así en varios países de la OCDE existen políticas generalizadas de colegios públicos selectivos, caracterizados generalmente por la aplicación de mecanismos formales para seleccionar a los estudiantes de acuerdo a su desempeño académico alcanzado al final o a lo largo de la educación básica, los cuales ingresan a otros establecimientos donde cursarán su educación secundaria⁵, por ejemplo, la mayor parte de los países germanos aplican procesos de selección para determinadas escuelas públicas, entre ellos Alemania, Austria, Hungría, Suiza y Luxemburgo (Dupriez, Dumay y Vause, 2008); también en Holanda, Italia y Francia (Clark, 2010).

Existe otro grupo de países que se caracterizan por la identificación de un grupo acotado de colegios públicos, especialmente en educación secundaria, que imponen altas exigencias académicas para seleccionar a los estudiantes interesados, principalmente enfocadas en el rendimiento de los alumnos, por ejemplo, las Magnet Schools⁶ (Cullen y Jacob, 2009) y Exam Schools en Estados Unidos (Dobbie y Fryer, 2014; Abdulkadiroglu, Angrist y Pathak, 2014); Grammar Schools en el Reino Unido (Clark, 2010); las Selective Public Schools en Australia; las escuelas manejadas por la Comisión Metropolitana de Instituciones Públicas de Educación en México (Dustan, 2010); así como establecimientos pertenecientes a los gobiernos de China, Rumania, Singapur y Turquía, que distribuyen a los estudiantes

² Para una detallada revisión de estos mecanismos ver Dupriez (2010).

³ Práctica que consiste en separar estudiantes en distintos tipos de provisión que en la mayoría de los casos son académicas, técnicas o vocacionales (Dupriez, 2010).

⁴ Entendido como el grado en que los estudiantes con diferentes condiciones iniciales (habilidad, raza, sexo, nivel socioeconómico) son ordenados a través de las salas de clases en las escuelas (Kalogrides y Loeb, 2012; Dupriez, 2010).

⁵ A este modalidad se le denomina “modelo de separación”.

⁶ Planty, Provasnik y Daniel (2007) indican que las escuelas públicas secundarias altamente selectivas casi se habían duplicado entre 2002-2003 y 2007-2008, aunque aún eran menos del 4% del total de los establecimientos. Igual tendencia se presentaba entre las Magnet Schools. Estos procesos se explicaban como respuesta a los crecientes procesos de accountability basado en estándares a los establecimientos educacionales.

en escuelas de secundaria basados casi completamente en exámenes de admisión (Dobbie y Fryer, 2011), al igual que Trinidad y Tobago (Jackson, 2010). Todas estas escuelas se caracterizan por tener altos estándares de exigencia para ingresar a ellas (i.e. determinados por la rendición de un examen al momento de postular a cada colegio o por la aplicación de un examen generalizado), la alta competencia por los cupos limitados en cada escuela, la exigencia al interior de estas y por la premisa que los estudiantes serán beneficiados por la interacción con compañeros de alto rendimiento académico (Dobbie and Fryer, 2014).

En términos de resultados a nivel general, la comparación internacional entre los sistemas escolares altamente selectivos y aquellos menos intensivos en este tipo de políticas, a partir de los resultados de pruebas PISA y TIMSS, permite concluir que en aquellos países donde se aplican sistemas de selección temprana de estudiantes los resultados obtenidos por los estudiantes de menor nivel socioeconómico y mayores dificultades de aprendizaje son significativamente menores a los obtenidos por este grupo de estudiantes en sistemas escolares menos selectivos (Hanushek y Woessmann, 2010).

Adicionalmente, no se evidenciaría un mejoramiento global entre los estudiantes más destacados, y los efectos más sustantivos solo se presentarían en la equidad de los aprendizajes (Hanushek y Woessmann, 2010; OECD, 2010; Dupriez, 2010). Todo esto evidenciaría que este tipo de políticas no constituirían un mecanismo para mejorar el desempeño promedio del conjunto de los estudiantes del sistema escolar.

Con respecto a los efectos asociados a las escuelas públicas altamente selectivas, los casos particulares de Estados Unidos e Inglaterra⁷, muestran evidencia consistente en cuanto a que estas generarían un incremento de las diferencias socioeconómicas en la composición de estudiantes de los diferentes tipos de escuelas, llevando a una mayor segregación social y cultural (Dobbie y Fryer, 2014; Saporito, 1998; Clark, 2010). Aunque por otra parte, existiría una mayor integración por condiciones de raza y etnia (Rosell, 1992; Finn y Hockett, 2012b), aunque, es posible anticipar que este tipo de políticas no afectarían a una gran proporción de los estudiantes más vulnerables, por lo que en la medida que éstos se concentren en pocas escuelas y barrios (Neild, 2002) los efectos serán específicos, así estrategias para mejorar las oportunidades de los estudiantes más vulnerables debiesen ser más estructurales que la sola creación de algunos establecimientos selectivos (Neild y Balfanz, 2001).

En este sentido, Rao (2013) muestra que al generarse la condición de mayor inclusión dentro de las escuelas los niños más acomodados mejorarían o tendrían un impacto positivo en características asociadas con el comportamiento social (como la generosidad, sociabilización con compañeros o

⁷ Una investigación que revisa los procesos históricos y la evidencia acumulada para estos dos países puede encontrarse en Schmidt (2011).

personas más pobres, etc.). Sin embargo, esta concentración de estudiantes de bajo nivel socioeconómico (y bajo rendimiento), afectaría de forma pequeña pero negativa al resto de los alumnos en mediciones de rendimiento escolar como test estandarizados o pruebas de conocimiento (Rao, 2013), conclusión similar a la que muestra The Sutton Trust (2009) para los estudiantes más aventajados, quienes cuando se encuentran en establecimientos que concentran a estudiantes de bajo rendimiento (y muy posiblemente menor nivel socioeconómico), perderían parte de sus posibilidades de aprendizaje (The Sutton Trust, 2009). De esta forma, a medida que se “reparan” las escuelas con peores resultados, también se muestra como un gran desafío o interrogante el determinar ¿qué hacer con los estudiantes más talentosos, motivados, y de mejores rendimientos del país? (Finn y Hockett, 2012a).

Ahora específicamente con respecto a los resultados en pruebas estandarizadas de las escuelas públicas altamente selectivas, la evidencia internacional es mixta, así para Estados Unidos la literatura reciente que ha estudiado los efectos de esta política –utilizando regresiones discontinuas⁸–concluye que los estudiantes seleccionados no lograrían mejores resultados, estos serían muy pequeños o éstos incrementos serían marginales en los test estandarizados (Abdulkadiroglu, Angrist y Pathak, 2014; Dobbie y Fryer, 2011, 2014) resultados consistentes con estudios anteriores que utilizan metodologías menos robustas (Cullen, Jacob y Levitt, 2005) resultados que también se verían en otras mediciones tradicionales de resultados como tasas de graduación o asistencia escolar (Cullen, Jacob y Levitt, 2006), dichas conclusiones son consistentes incluso cuando la postulación está abierta a todos los estudiantes y la selección es realizada por mecanismos que aseguran una igual probabilidad de selección (por ejemplo, loterías) para todos los estudiantes interesados (Cullen, Jacob y Levitt, 2006; Cullen y Jacob, 2009; Ballou y Liu, 2009), en este mismo sentido Zhang (2008) tampoco encuentra ganancias asociadas a estudiantes seleccionados aleatoriamente en escuelas de China.

Consistente con la evidencia antes mostrada, para Inglaterra, Clark (2010)⁹ estima –nuevamente a través de una metodología de regresión discontinua– que las escuelas públicas selectivas tendrían a lo más un pequeño impacto en test estandarizados, así mismo Galindo-Rueda y Vignoles (2004) a través de matching propensity score¹⁰ encuentran que no existiría efecto sobre los estudiantes de habilidad baja o media, pero sugieren que existiría un efecto positivo sobre los estudiantes con mayor habilidad (top 20%) (Galindo-Rueda y Vignoles, 2004), esto último daría sustento al planteamiento realizado

⁸Metodología que, en pocas palabras, consiste en utilizar solo a los individuos seleccionados para ingresar y aquellos que no lo fueron, en el límite determinado por una o un grupo de variables de selección.

⁹ En Inglaterra ha existido durante los últimos veinte años diversos estudios respecto al efecto que tuvo la reducción de las Grammar Schools en los 60s y 70s hasta sólo 164 en la actualidad, sin embargo, se puede indicar que sus resultados son mixtos y muchos de ellos presentan importantes debilidades metodológicas (Pischke y Manning, 2006; Coe et al., 2008).

¹⁰ Metodología descrita más adelante, que consiste principalmente en encontrar la probabilidad de ser tratado con la finalidad de encontrar un grupo de control comparable.

por Finn y Hocket (2012b) quienes afirman que este tipo de colegios debiese estar enfocado a los estudiantes con mayores y mejores rendimientos del país.

Sin embargo, los resultados para países en vías de desarrollo –en general obtenidos con metodologías de regresión discontinua– muestran grandes efectos asociados a estos mecanismos, en este sentido, Pop-Eleches y Urquiola (2013), encuentran que existiría un efecto positivo asociado al ingreso a mejores escuelas en Rumanía; de la misma forma, Dustan (2010) muestra que asistir a una escuela de elite estaría asociado con mejores resultados en pruebas estandarizadas en México; y Jackson (2010) muestra que el asistir a mejores escuelas de Trinidad y Tobago estaría asociado con un gran efecto positivo en los puntajes obtenidos en un test realizado al terminar secundaria.

De igual forma, la literatura tampoco es concluyente en el impacto que pueden tener estos colegios en indicadores como ingreso a la universidad o graduación, así Dobbie y Fryer (2011, 2014) muestran para Estados Unidos que existiría solo un pequeño impacto sobre el ingreso a la universidad y la graduación de secundaria, en cambio Clark (2010) muestra que haber asistido a este tipo de escuelas tendría asociado un efecto importante en la matrícula en la universidad.

A pesar de la evidencia mixta sobre los efectos agregados sobre los test de evaluación, la evidencia indica que podrían existir efectos positivos de largo plazo. Clark (2010) demuestra que los estudiantes que asisten a las Grammar Schools tienen una mayor probabilidad de asistir a la universidad en comparación con aquellos que asisten a escuelas públicas tradicionales (*Comprehensive Schools*), efecto estadísticamente significativo luego de controlar por las diferencias de capital socioeconómico y cultural de los estudiantes y sus compañeros. En este sentido hay que ser prudentes al anticipar posibles efectos en el largo plazo “es muy probable que estos existan debido a que la exposición a compañeros de mejor rendimiento puede aumentar el *capital social* en formas que pueden ser importantes para los resultados futuros independientes de la matrícula universitaria, graduación o capital humano...el verdadero impacto de las *exam schools* solo podrá ser entendido con el paso del tiempo” (Dobbie y Fryer, 2014).

En este mismo sentido, otro efecto medible en el largo plazo es el ingreso, donde Dale y Krueger (2002) muestran para Estados Unidos que no existirían efectos positivos en los ingresos para aquellos que asistieron a escuelas selectivas, cosa que es corroborada en un estudio más reciente con otra cohorte en Dale y Krueger (2011). Sin embargo, tal como ellos mencionan, existen notables excepciones para ciertos subgrupos de población, personas de origen hispano o de color cuyos padres tenían baja escolaridad (presumiblemente bajo nivel socioeconómico) tendrían un retorno a la selectividad escolar

muy grande. En este sentido, Hoekstra (2009) muestra que asistir a una universidad selectiva tendría un gran impacto sobre los futuros salarios.

Si bien es cierto su existencia, la causalidad del efecto par – entendido como la influencia ejercida sobre cada estudiante por otros que se encuentran en su mismas clase o escuela en cuestión (Dupriez, 2010) – en este tipo de establecimientos no ha podido ser determinada con exactitud en su real sentido y magnitud sobre los resultados escolares (por ejemplo, test estandarizados) debido principalmente a problemas metodológicos (Angrist, 2014) muchos de los estudios antes citados dan cuenta que a pesar de sus datos, estimaciones y resultados es posible que existan otros efectos asociados con la asistencia a este tipo de escuelas que pueden ser positivos como negativos y que pueden ser directos o indirectos, como por ejemplo la reducción en problemas disciplinarios y arrestos (Cullen, Jacob y Levitt, 2006), la existencia de efectos directos derivados de la especialización en diversas áreas que proveerían este tipo de escuelas (Abdulkadiroglu, Angrist y Pathak, 2014), la reducción en el esfuerzo de los padres y el ordenamiento al interior de los colegios por parte de los profesores (Pop-Eleches y Urquiola, 2013). En este sentido, la decisión de elegir este tipo de escuelas no estaría enfocada solamente en temas netamente académicos, sino que se asociarían también con otro tipo de características o servicios no académicos entregados por este tipo de escuelas (Cullen, Jacob y Levitt, 2006).

La existencia de evidencia circunstancial sobre la efectividad, a lo que se suma la poca claridad existente sobre la causalidad de los resultados observados en este tipo de escuelas (los diseños de regresión discontinua solo mostrarían qué sucede en el límite obviando que pasa, por ejemplo, con los mejores alumnos), hace prácticamente imposible diferenciar si realmente el efecto o lo que les sucede a los estudiantes dentro de estas escuelas se debe a los altos estándares, currículum, profesores u otra característica de cada establecimiento, haciendo sentido las palabras de un director entrevistado “¿Los estudiantes lo hacen bien gracias a nosotros o a pesar de nosotros? No estamos seguros” (Finn y Hockett, 2012a, pp 193). O si sería reflejo de lo que los otros compañeros traen a la clase (Finn y Hockett, 2012a) haciendo referencia a los potenciales efectos derivados del nuevo ambiente en que se encontrarían al ingresar a estas escuelas que no necesariamente son cuantificables (por ejemplo, la reducción de la ansiedad de un alumno que no necesita temer más ser víctima o tratado como “nerd”, como afectan las expectativas de los profesores en lo que aprenden los alumnos, etc.) así la cultura de pares del colegio es muy probable que tenga otros efectos sobre los alumnos pero es difícil saber qué elementos tienen y cuál es su efecto (Finn y Hockett, 2012a).

Adicionalmente estás serían características cuyos efectos no necesariamente se verán en el corto plazo (test estandarizados, u otras mediciones tradicionales de efectividad) y que incluso en el largo plazo

pueden no ser medibles o cuantificables de forma correcta tal como la satisfacción personal o salud, relevan la importancia del potencial efecto par asociado con el ingreso a este tipo de colegios. En este sentido, existe evidencia sobre el potencial efecto negativo que podría tener sobre este tipo de estudiantes el asistir a estos colegios sobre la percepción de sí mismos, teoría conocida como efecto “pez grande en estanque pequeño” (big-fish-little-pond-effect) la cual ha demostrado empíricamente que los estudiantes que asisten a clases o escuelas con alumnos de altos rendimientos tendrían un auto-concepto o percepción de sí mismos menor que su contrafactual igual de capaz educado en ambientes de menor (o promedio) habilidad (Zell y Alicke, 2009; Seaton, Marsh y Craben, 2010; Marsh et al, 2014), efecto del cual Pop-Eleches y Urquiola (2013) encuentran evidencia observando que los estudiantes al ingresar a este tipo de escuelas y ser expuestos a mejores pares se dan cuenta que son más débiles sintiéndose marginalizados. Esta evidencia es de suma importancia debido a que la *identidad de una persona* se construye comparando el comportamiento, las emociones, los resultados, el estatus, y otras características con otras personas y otros grupos (Zell y Alicke, 2009), hecho ampliamente reconocido por los psicólogos sociales. Razón por la cual este tipo de políticas debe ser tomada y analizada con cuidado.

Finalmente, con respecto al análisis sobre los potenciales efectos negativos que podría generar la migración de estudiantes de excelencia hacia colegios públicos altamente selectivos sobre el desempeño escolar de los estudiantes que se mantienen en los establecimientos de origen, la evidencia y estudios empíricos son casi inexistentes (Altonji, Huang y Taber, 2014), ello debido a que el foco de la investigación empírica no ha estado concentrado en este grupo pasivo, así como también porque la investigación teórica (Epple y Romano, 1998, 2002; Nechyba, 2003) daba cuenta de potenciales efectos negativos de una política de este tipo, sin embargo, dicha conclusión es muy dependiente de los supuestos acerca de la función de producción de la calidad de los aprendizajes de los estudiantes (por ejemplo, la relevancia y forma funcional del efecto par sobre estos aprendizajes), lo cual implica que el resultado empírico de esta política podría tener un alto componente idiosincrático de cada país e incluso a nivel de cada distrito, comuna o escuela en particular¹¹ y probablemente sobreestime el efecto real que tenga dicha política (Walsh, 2009; Altonji, Huang y Taber, 2014). Sin embargo, la escasa evidencia con respecto a este tema muestra que el efecto del “descreme” de los mejores estudiantes tendría un efecto negativo, pero pequeño, entre los estudiantes no migrantes (Altonji, Huang y Taber, 2014), efecto que podría deberse a que la migración generalmente es de un solo estudiante.

¹¹ Por ejemplo, a pesar de una diversidad de evidencia que indica que el efecto par es significativo y con efectos más favorables para estudiantes vulnerables, Duflo, Dupas y Kremer. (2010) demuestran que en escuelas de Kenia la concentración de estudiantes más homogéneos en sus capacidades de aprendizaje beneficia más que una composición heterogénea, pues permite a profesores mejorar su atención a estudiantes con desafíos similares.

2) Revisión de literatura nacional

La literatura nacional sobre desempeño escolar se ha centrado en la evaluación de si mayores grados de competencia entre proveedores públicos y privados ha generado mejores resultados educativos para todo el sistema escolar, y si sus efectos son diferenciados entre aquellos estudiantes que escogen colegios privados respecto de los que escogen escuelas públicas. En general, la literatura ha demostrado que los efectos de la competencia han tenido un fuerte incremento en la segregación de los estudiantes y un incremento en la brecha de los aprendizajes de acuerdo al NSE de los estudiantes, asimismo; respecto de los aprendizajes, pareciera ser que en un contexto de equilibrio parcial, los colegios particulares (subvencionados y pagados) no logran un mejor desempeño que los públicos, una vez que se controla por las condiciones socioeconómicas de los estudiantes y sus compañeros de curso (Bellei, 2007; Lara, Mizala y Repetto, 2011).

En relación a los procesos de selección la evidencia a nivel internacional ha sido creciente en el último tiempo, pero en el sistema escolar chileno es muy escasa. Punto que en el último año ha cobrado relevancia debido a que dentro del nuevo proyecto de Ley de reforma a la educación enviado al Congreso en mayo de 2014, considera terminar definitivamente con este tipo de prácticas en el sistema escolar (con la excepción de los colegios públicos emblemáticos y particulares pagados). En este sentido, la evidencia empírica existente para Chile, ha mostrado que los procesos de selección a nivel escolar serían una práctica ampliamente utilizada por las escuelas en Chile, en todo los niveles de educación (hecho anecdótico debido a que la selección está prohibida por Ley hasta 6to año básico) y en colegios con o sin subvención del estado (Godoy, Salazar y Treviño, 2014), junto a esto al parecer no existiría una asociación directa entre selectividad y efectividad escolar en términos de valor agregado (Carrasco et al., 2014). Adicionalmente, los procesos de selección no estarían necesariamente relacionados con el exceso de demanda que existe en algunos establecimientos, razón esgrimida frecuentemente para promover y sustentar los procesos selectivos dentro del sistema escolar (Carrasco et al., 2014).

Específicamente, en relación al efecto de la selección realizada por los liceos públicos selectivos de excelencia en Chile la literatura es aún más escasa, así en una versión preliminar de esta investigación Valenzuela y Allende (2011) muestran que –para una cohorte analizada, utilizando una muestra reducida obtenida al aplicar matching 1 a 1–, existiría un efecto positivo pero que dicho efecto estaría asociado principalmente con el efecto par de la habilidad de los alumnos que ingresan a dichos establecimientos. Así mismo, Bucarey et al. (2014) muestran, mediante la estimación del efecto promedio local del tratamiento (LATE) para una determinada cohorte de estudiantes, que asistir al

Instituto Nacional estaría asociado con un incremento de 0,25 desviaciones estándares en la prueba de selección universitaria (PSU), sin embargo tal como ellos mencionan “...*la presente investigación no aborda el canal por el cual el Instituto Nacional genera mejores resultados, en particular no es posible señalar si el mayor puntaje en la PSU se debe a efecto par, tracking u otros aspectos.*” (Bucarey et al, 2014).

Finalmente, en otro tipo de análisis, Alcalde (2013) muestra que en términos de producción los Liceos Bicentenario presentarían un efecto positivo para el periodo 2009-2011, siempre haciendo notar que estos corresponden a resultados preliminares debido al poco tiempo de existencia de dichos liceos. En este sentido, los Liceos Bicentenario mostrarían una mejora notoria en resultados intermedios: retiro, asistencia, reprobación, variación en puntajes SIMCE y PSU (Carrasco et al, 2014b).

Dada la evidencia mixta a nivel internacional –efectos positivos en países en vías de desarrollo y pequeños o nulos efectos en países desarrollados– y la escasa a nivel nacional, se buscará contestar a la siguiente interrogante ¿Una política de selección exhaustiva de alumnos genera una mejora real en el rendimiento de los alumnos que ingresan a liceos selectivos de excelencia en Chile?, para luego –si dicho efecto existe –determinar ¿quiénes son los más beneficiados–en términos de nivel socioeconómico–con una política de este tipo? Una vez demostrada la existencia y la heterogeneidad de los resultados se buscará determinar ¿cuál es la razón de los mejores resultados?, para finalmente determinar si la efectividad mostrada por este tipo de colegios ¿se debe principalmente al tipo de alumnos (antecedentes familiares e individuales) que ingresa?

III. Datos

Se utilizaron dos paneles de datos, correspondientes a los niveles 4º básico-2º medio (2002-2008; 2006-2012), en adelante al primer año del panel se le llamará año base y al año final será llamado año de tratamiento¹². Cabe destacar que las pruebas SIMCE no están construidas como un panel de datos realmente, sino que corresponden a pruebas censales realizadas todos los años, por lo que no se enfatiza en seguir realmente a los estudiantes a través del tiempo y se asume que por el hecho de ser censales debieran encontrarse, por esta razón la atrición o tasa de pérdida de observaciones encontrada en estos paneles de datos no está controlada y se da naturalmente (por ejemplo, aquellos alumnos que no asisten el día de la prueba no se les exige realizar el test otro día). Para que las estimaciones sean consistentes es necesario que el supuesto de valores perdidos aleatoriamente (MCAR o MAR por sus acrónimos en inglés) sean asumidos, de esta forma se asume que no existirían prácticas sistemáticas por parte de los colegios a dejar alumnos fuera de los test.

Adicionalmente, a los datos de los test realizadas se incorporaron antecedentes individuales y familiares provenientes de la encuesta complementaria que responden las familias de los estudiantes que rinden el SIMCE de ambos años. Así la muestra definitiva corresponderá a los estudiantes que fue posible encontrar en 2º medio en los paneles antes mencionados y que contaban con información sobre sus características individuales y escolares para el año base (2002, 2006) y de tratamiento (2008, 2012).

La Tabla I muestra el número de alumnos para los que fue posible obtener información de las pruebas SIMCE como de las bases secundarias de información utilizadas, y describe la pérdida de información existente.

Tabla I. Composición del Panel 2002-2008 y Panel 2006-2012

Panel 2002 -2008		Panel 2006 -2012	
Nº total matriculados en 4º básico 2002 (SIGE)*	290.755	Nº total de matriculados en 4º básico 2006 (SIGE)*	266.925
Nº rinden SIMCE año 2002	269.113	Nº rinden SIMCE año 2006	246.917
Nº cuentan con información de padres año 2002	248.577	Nº cuentan con información de padres año 2006	228.930
Nº total de matriculados en 2º medio 2008 (SIGE)*	253.952	Nº total de matriculados en 2º medio 2012 (SIGE)*	239.823
Nº rinden SIMCE año 2008	229.413	Nº rinden SIMCE año 2012	236.342
Nº cuentan con información de padres año 2008	193.500	Nº cuentan con información de padres año 2012	154.508

¹²Cabe destacar que el tratamiento para los paneles cuyo año final es 2º medio es en promedio de 4 años, y para el panel que termina en 8º básico en promedio es 2 años. No podemos determinar de forma exacta el año en que los estudiantes de estos paneles ingresan a los establecimientos, por lo que se asume que todos los estudiantes ingresan en 7º básico, año en que todos los colegios considerados ingresan a los respectivos establecimientos.

Nº total en panel alumnos	168.775	Nº total en panel alumnos	164.524
Nº con información completa 2002**	126.490	Nº con información completa 2006**	146.848
Nº con información completa 2008	134.060	Nº con información completa 2012	111.917

† Lo totales para cada año corresponden al total de alumnos para los cuales fue posible contar con la variable mrun (eliminando los duplicados). Para casos donde la variable mrun se encontró repetida esta fue recuperada desde el SIGE utilizando el mrun, rbd y curso respectivo.

†† Información completa se refiere al N obtenido utilizando TODAS las variables incluidas en el análisis.

*No se incluye datos con mrun repetido, solo considera los datos directos desde el SIGE, (no incluye alumnos que pueden estar en SIMCE y no en SIGE).

**Incluye promedio de notas obtenido desde el Sistema General de Información de Estudiantes (SIGE).

La reducción en la cantidad de información en los paneles o atrición, responde a que las pruebas SIMCE no están construidas realmente como paneles de datos en el sentido que no siguen a una misma cohorte a través del tiempo. Adicionalmente, la mayor pérdida de información en estos paneles se debe a la repitencia y abandono temprano de los alumnos (Valenzuela, Allende, Sevilla y Egaña, 2014).

Adicionalmente, debido a la forma en que se completa los cuestionarios para padres (estos son enviados a la casa y deben ser devueltos de forma voluntaria) el número de observaciones totales que tienen esta información se ve reducido drásticamente, siendo más evidente en 2º medio (especialmente en el año 2012).

Así, sumado esto a la pérdida de información debido a la no respuesta, la muestra finalmente construida y que contiene información para todas las variables utilizadas en esta investigación muestra que para el año 2002 fue posible contar con un 51% de las observaciones (64% para 2006) y para el año 2008 con un 69% (95% para el 2012).

IV. Metodología

El desafío principal de este estudio es estimar el efecto o impacto que ha tenido, en resultados de las pruebas estandarizadas SIMCE de 2° medio, la participación de dos cohortes de estudiantes en colegios clasificados como Liceos Selectivos Públicos de Excelencia (grupo de tratamiento).

En la literatura se utiliza frecuentemente dos tipos de estimadores para determinar el impacto de un tratamiento, el efecto promedio del tratamiento (ATE por su acrónimo en inglés) o la diferencia de resultados después de haber participado y no en un tratamiento que usualmente se identifica como

$$\tau_{ATE} = E[Y(1) - Y(0)]$$

donde $Y(1)$ e $Y(0)$ corresponden al resultado de haber sido expuesto al tratamiento o no respectivamente; y el efecto promedio del tratamiento sobre los tratados (ATT por acrónimo en inglés) que estaría dado por la diferencia en resultados solo para aquellos que fueron tratados

$$\tau_{ATT} = E[Y(1) - Y(0) | T = 1]$$

donde $T = 1$ denota que los individuos pertenecen solo al grupo de tratamiento.

El efecto promedio del tratamiento (ATE), en general busca responder a la pregunta ¿Cuál es el efecto esperado en los resultados si los individuos en la población fueran distribuidos aleatoriamente? (Caliendo y Kopeinig, 2008), así ATE será relevante cuando el interés de una política está enfocada en hacer obligatorio o masivo un programa voluntario (Bryson et al, 2002). En cambio, el ATT busca estimar el efecto promedio del tratamiento enfocándose directamente en los participantes (Caliendo y Kopeinig, 2008), buscando determinar el efecto solo sobre aquellos para los cuales el tratamiento fue diseñado¹³. Así, ATT será el parámetro que puede contestar la pregunta si un programa debe ser abandonado o no, ya que si el impacto promedio sobre los tratados es menor al costo del programa, existirá gran evidencia para eliminar dicho programa (Bryson et al, 2002).

Dado lo anterior, esta investigación estará enfocada en la búsqueda del efecto promedio del tratamiento sobre los tratados (ATT), debido a que la política analizada está dirigida a un determinado tipo de alumnos, aquellos con mayor habilidad y capacidades. Esto permitirá identificar cuál es el efecto de este tipo de colegios en los resultados obtenidos por los alumnos que cumplen con las condiciones de ingreso y si realmente las diferencias de puntajes observadas se deben a una mejora en sus resultados obtenida a partir del tratamiento, y adicionalmente permitirá encontrar respuesta a las interrogantes

¹³ A modo de ejemplo, ATT sería de interés en la evaluación de un programa de capacitación laboral diseñado para incrementar las habilidades de trabajadores con baja educación.

antes planteadas.

Para implementar esta última estimación, existe un problema, ya que una vez identificado el tratamiento para las unidades tratadas solo observamos los resultados bajo la condición de tratamiento y no la situación sin tratamiento; y al mismo tiempo para las potenciales unidades de control solo se observa el resultado sin tratamiento y no la situación con tratamiento (Steiner y Cook, 2013). Por lo tanto, para inferir si existe algún efecto asociado al programa no es posible utilizar directamente los resultados post aplicación del tratamiento, a lo que se suma la no existencia de un grupo de tratamiento definido de forma aleatoria (experimento). Razones por la cual métodos cuasi experimentales serán utilizados con la finalidad de simular un grupo de control que a la vez sea comparable al de tratamiento, de tal forma, el sesgo originado por la no aleatorización de la participación en el programa sea reducido considerablemente.

Al igual que en la mayor parte de los estudios revisados, dado el diseño del tratamiento sería apropiada la estimación del efecto promedio local del tratamiento¹⁴ (LATE por acrónimo en inglés) debido a que la asignación del tratamiento está determinada, completa o parcialmente, por el valor de un predictor (por ejemplo, una prueba de ingreso o alguna variable relevante utilizada en la selección de los estudiantes) que se encontraría en ambos lugares de la brecha generada por los tratados y no tratados (Imbens y Lemieux, 2008), dado esto un diseño de regresión discontinua podría ser aplicado permitiendo estimar el efecto local del tratamiento (Abdulkadiroglu, Angrist y Pathak, 2014; Dobbie y Fryer 2011, 2014; Pop-Eleches y Urquiola, 2013; Jackson, 2010; Dustan, 2010).

El problema, en este caso, es que no disponemos de las características ni los resultados de todos los procesos de selección realizados por los colegios considerados en este estudio, Bucarey et al (2014) disponen de estos registros para un año y estiman LATE pero solo para un liceo de la Región Metropolitana, específicamente el Instituto Nacional. Así la estrategia de estimación del efecto del tratamiento –utilizada en este estudio– corresponde a un *segundo mejor*, y es posible debido a que podemos estimar la probabilidad de haber sido tratado y a la existencia de mediciones de variables relevantes realizadas antes y después del tratamiento.

Antes de comenzar, es necesario definir concretamente cuál es el tratamiento que se busca evaluar y cuál es el objetivo de este, adicionalmente –y asociado a esto– se define qué entenderemos por Liceos Selectivos Públicos de Excelencia (LSPE), identificando así el grupo de tratamiento. Luego se define el grupo de control apropiado para estimar el efecto promedio del tratamiento sobre los tratados.

¹⁴Esta correspondería a un caso específico de ATT ya definido.

1) Definición del tratamiento y del grupo tratamiento, los Liceos Públicos Selectivos de Excelencia.

El tratamiento será definido como la selección e ingreso a un liceo selectivo público de excelencia (LSPE), liceos que destacan principalmente por su excelencia académica reflejada en los resultados de pruebas estandarizadas durante la educación básica y media (SIMCE) como en las pruebas de ingreso a la universidad (PSU). De esta forma, el objetivo del tratamiento se podría acotar¹⁵ a la obtención de excelencia académica por parte de los alumnos que ingresan a los colegios que se verá reflejada en la obtención de puntajes promedio destacados en las pruebas estandarizadas SIMCE y/o PSU.

De esta forma, la evaluación realizada buscó identificar *si los LSPE son capaces de generar una mejora real en el rendimiento de los alumnos, y si dichas mejoras se derivan de los procesos internos de este tipo de establecimientos (como pueden ser sus metodologías de enseñanzas, gestión u organización) o si dicho efecto se debe principalmente al tipo de alumnos que ingresan a través de los procesos de selección realizados.*

Una vez identificado el tratamiento es necesario determinar cuáles serán los liceos donde se aplica este tratamiento y que por ende serán definidos como liceos selectivos públicos de excelencia (LSPE). Esta identificación fue llevada a cabo, utilizando el panel 2002-2008, para luego utilizar los mismos colegios seleccionados en el análisis del panel 2006-2012.

El panel 2002-2008, cuenta con un total de 2.542 establecimientos que rindieron la prueba SIMCE de 2° medio en el año 2008 y que poseen información proveniente del cuestionario de padres y apoderados de la prueba SIMCE. De estos, solo 695 (27,3%) colegios en todo el país corresponden a liceos municipales, un 58,5% (1.486) a liceos particulares subvencionados, y el restante 14,2% a liceos particulares pagados.

Luego de entre los 695 colegios municipales que rindieron la prueba SIMCE de 2° medio en el 2008 en todo el país, se identificaron aquellos en donde los apoderados contestaron masivamente (más de un 70%) que el colegio aplicaba mecanismos de selección académica a los estudiantes al momento de postular. La definición de selección académica se basó en la respuesta afirmativa de los apoderados a las preguntas: *el colegio requería la entrega de certificado de notas del establecimiento de procedencia y la rendición de un examen escrito o de una prueba de conocimiento para ingresar al nuevo colegio*

¹⁵Se hace referencia a acotar el objetivo del tratamiento ya que estos establecimientos en su mayoría tienen objetivos mucho más amplios y que generalmente no son medibles o son muy difíciles de medir en el corto o largo plazo, como por ejemplo: “hacer florecer a la patria”, “darle honor a la patria”, “construir lazos afectivos”, “contribuir al bienestar de la sociedad”, “contribuir al desarrollo integral de las personas”, “generar ciudadanos responsables”, etc. (Objetivos extraídos desde los objetivos planteados por algunos de los colegios seleccionados).

público.

De los 695 colegios municipales solo aproximadamente un 16% (110) presentan la característica de selectividad académica antes descrita, su distribución a nivel país se muestra en la Tabla II.

Tabla II. Distribución de Colegios Públicos con Enseñanza Media de Acuerdo a Selectividad Académica (2008)

Región	N° de Colegios Municipales	N° Colegios Municipales selectivos
Región de Tarapacá	11	-
Región de Antofagasta	24	1
Región de Atacama	17	3
Región de Coquimbo	34	6
Región de Valparaíso	75	10
Región del Lib. General Bernardo O'Higgins	43	9
Región del Maule	59	7
Región del Biobío	100	19
Región de la Araucanía	51	3
Región de los Lagos	52	9
Región de Aysén	7	-
Región de Magallanes	11	2
Región Metropolitana	169	35
Región de los Ríos	30	5
Región de Arica y Parinacota	12	1
Total	695	110

Fuente: Elaboración Propia en base a Resultados SIMCE 2008

Para determinar que los liceos públicos selectivos contaran con la condición de “Excelencia” se restringió la muestra a aquellos que logran al menos 280 puntos promedio en alguno de los sectores de lenguaje o matemática en la prueba SIMCE de 2° medio del año 2008¹⁶, lo cual restringió la muestra a un total de 25 colegios.

Finalmente, la muestra fue restringida a aquellos liceos que imparten educación científico-humanista y que tuvieran una cobertura educativa desde 7° básico a 4° medio, excluyendo por ende a colegios completos, solo con educación media y que entregan enseñanza vocacional o técnico-profesional y los colegios denominados polivalentes. Este último filtro dejó un total de 17 establecimientos denominados Liceos Selectivos Públicos de Excelencia¹⁷ (LSPE). De estos, 13 se localizan en la Región Metropolitana, 2 en la de Los Lagos, y 1 en las regiones de O’Higgins y Maule; asimismo, 5 entregan educación solo a hombres, 5 solo a mujeres y 7 corresponde a colegios mixtos (Ver Anexo 1).

¹⁶ El promedio nacional del puntaje de estas pruebas en 2° medio es de 250 puntos.

¹⁷ Adicionalmente, se comprobó si dentro de estos colegios se encontraban los colegios llamados tradicionales.

2) Definición del grupo de control

Tal como se mencionó, la muestra corresponde a todos los estudiantes encontrados en los paneles y que posean información desde los cuestionarios de los padres para ambos periodos, adicionalmente ya fue definido el grupo de tratamiento como aquellos alumnos pertenecientes a colegios públicos selectivos de excelencia, definiendo la excelencia como la existencia de al menos un promedio en lenguaje y matemática superior a 280 puntos.

En el caso del grupo de control es necesario decidir qué queremos estimar para luego definir quiénes comprenderán el grupo de control, así como ya fue mencionado, esta investigación busca determinar el efecto solo sobre aquellos para los cuales el tratamiento fue diseñado que en este caso corresponderían a los mejores estudiantes del sistema educacional. Para determinar exactamente un grupo de control apropiado, la utilización de los resultados de los procesos de selección realizados en cada colegio sería lo óptimo (por ejemplo: el puntaje de corte para ingresar), pero como no lo poseemos, se utilizará una aproximación parcial que permite determinar *a los alumnos que potencialmente podrían ser afectados por esta política o podrían participar del tratamiento*.

El grupo de control comprenderá a la subpoblación de compañeros de curso en 4° básico (año base) de todos los alumnos tratados y que no se encuentran estudiando en liceos selectivos públicos de excelencia en el año de tratamiento (2° medio) de esta forma se elimina el efecto fijo asociado a las condiciones iniciales del establecimiento. Luego para determinar cuáles serán los alumnos que potencialmente pueden recibir el tratamiento, primero se restringe dicha subpoblación a todos aquellos que tengan un promedio de notas superior a 5,5 en la media de los promedios generales de notas correspondientes a 5° y 6° básico, esto se debe a que en general los LSPE en sus procesos de admisión le asignan una ponderación en el puntaje final a los antecedentes académicos –generalmente estos dos niveles– presentados por los alumnos¹⁸. Luego se restringe esta muestra a todos aquellos alumnos que en la prueba SIMCE de 4° básico de lenguaje y matemática hayan obtenido un nivel de logro elemental o adecuado¹⁹, este segundo filtro responde a la existencia de pruebas de ingreso tanto de lenguaje como matemática en los procesos de postulación a los LSPE así esta metodología considera como una aproximación de dichas pruebas a las pruebas SIMCE realizadas por los alumnos en 4° básico.

¹⁸ Específicamente toman alguno (o todos) los sectores: Lenguaje y Comunicación, Matemática, Historia, Geografía y Ciencias Sociales y Ciencias Naturales, como no se posee los promedios correspondientes a cada asignatura se utilizó solo el promedio general de notas.

¹⁹ Los Estándares de Aprendizaje son referentes que describen lo que los estudiantes deben saber y poder hacer para demostrar, en las evaluaciones SIMCE, determinados niveles de cumplimiento de los objetivos de aprendizaje estipulados en el currículum vigente (Ministerio de Educación, 2013). El MINEDUC define tres niveles de logro: (1) Insuficiente, estudiantes que no logran demostrar consistentemente que han adquirido los conocimientos y habilidades más elementales estipulados en el currículum para el periodo evaluado; (2) Elemental, estudiantes que han demostrado que han adquirido los conocimientos y habilidades más elementales estipulados en el currículum para el periodo evaluado; (3) Adecuado, que han adquirido los conocimientos y habilidades básicos estipulados en el currículum para el periodo evaluado. Estos niveles se definen en términos de los puntajes SIMCE obtenidos por los alumnos.

Se testeó como cambiaban los resultados obtenidos al modificar el promedio de notas al cual se restringió el primer grupo de control. Adicionalmente, se testeó un nuevo grupo de control que combinó matching exacto y multivariado para la elección del grupo de control. Estas pruebas se describen más adelante en el análisis de sensibilidad de los resultados obtenidos.

i. Calidad del grupo de control

Encontrar un grupo de control comparable es importante ya que una mala identificación puede llevar a conclusiones poco robustas. En este sentido, Imbens y Wooldridge (2009), muestran que reportar y analizar las diferencias normalizadas²⁰ entre los tratados y potenciales controles es una buena práctica y adicionalmente proporcionan una regla para un buen balance entre dos variables, esta corresponde a una diferencia normalizada de no más de un cuarto de desviación estándar.

Los resultados de las diferencias normalizadas se presentan en los Anexos 2 y 3, en estos es posible observar que los paneles completos generados para los años 2002-2008 como para 2006-2012 muestran poco balance en las variables explicativas consideradas, así 11 variables en el panel 2002-2008 y 10 en 2006-2012 –de un total de 31 variables consideradas– no cumplirían con la regla establecida, llegando a tener diferencias normalizadas que alcanzarían casi cuatro quintos de una desviación estándar.

Por el contrario, el ajuste observado para el grupo de control antes descrito²¹, los resultados muestran que el ajuste mejora considerablemente con ambos grupos construidos y las diferencias entre estos serían mínimas, en términos de ajuste de variables, ya que la diferencia en el número de observaciones es bastante considerable. En este sentido, solo 4 variables superarían la regla establecida²².

Así los resultados muestran que se obtiene una mejora considerable en el ajuste de las variables –medido a través de diferencias normalizadas– utilizando los grupos de control antes descritos en vez de paneles completos.

3) Estimación del efecto promedio del tratamiento sobre los tratados

Para implementar la estimación del efecto promedio del tratamiento sobre los tratados (ATT), se utilizará estimadores que combinan el uso del procedimiento conocido como matching propensity

²⁰ Las diferencias normalizadas corresponden a las diferencias en las medias para tratados y controles ponderada por la raíz cuadrada de la suma de sus varianzas, así se estima lo siguiente $\Delta_x = \frac{\bar{x}_1 - \bar{x}_0}{\sqrt{s_0^2 + s_1^2}}$

²¹ Se presentan las estimaciones solo para las observaciones dentro del área de soporte común para lenguaje ya que los resultados son prácticamente idénticos con el área de soporte común para matemática.

²² Dos de estas son utilizadas solo para estimar el propensity score (razones de ingreso) y las otras dos (puntajes SIMCE estandarizados) se encuentran muy cerca del punto de corte establecido.

score (Rosembaun y Rubin, 1983) y ajustes a través de regresión. Así se buscará abordar los problemas de las metodologías basadas exclusivamente en la estimación del propensity score, con la finalidad de alcanzar un mayor grado de robustez a la mala especificación de los modelos, común en el mundo real. Además, esto permitirá la utilización de una mayor proporción de muestra disponible.

A. Estimación de la probabilidad de recibir el tratamiento, el Propensity Score

El propensity score, $p(X_i)$, consiste en estimar la probabilidad que un individuo reciba el tratamiento condicional a un vector de variables explicativas, específicamente Rosembaun y Rubin (1983) muestran que si $p(X_i) \equiv \Pr\{T_i = 1|X_i\} = E\{T_i|X_i\}$ donde $0 < p(X_i) < 1$ entonces es posible argumentar que la condición de selección en observables²³ $Y_i(1), Y_i(0) \perp T_i | X_i$ se cumple para $p(X_i)$ de forma que $Y_i(1), Y_i(0) \perp T_i | p(X_i)$, donde $Y_i(1), Y_i(0)$ corresponden al resultado del individuo i si este fue expuesto al tratamiento o no, respectivamente. En otras palabras para encontrar un grupo de control a través de matching no es necesario condicionar las distribuciones de los resultados a un vector de características X para encontrar el mejor control, sino que solamente bastaría con controlar por la probabilidad de haber sido tratado lo que está dado por $p(X)$ (Burga, 2003), en este caso dicha probabilidad se estimó a través de un modelo Logit. Adicionalmente, estos estimadores requieren del cumplimiento del supuesto $0 < \Pr(T_i = 1|X_i = x) < 1$, o en otras palabras que para todos los valores de x existan unidades de control y de tratamiento (Imbens y Lemieux, 2008), dicho supuesto es conocido como traslapo (overlap). Este es muy probable que se cumpla dentro del área de soporte común, que corresponde al área de intersección de las distribuciones de $p(X_i)$ para los individuos tratados y no tratados, donde luego sólo aquellas observaciones dentro de dicha área serán utilizadas en los métodos de estimación seleccionados²⁴. Para la estimación de ATT, se puede utilizar una versión más débil de los supuestos antes descritos, (1) selección en observables débil²⁵ $Y_i(0) \perp T_i | X_i$ y (2) traslapo débil²⁶ $\Pr(T_i = 1|X_i = x) < 1$, de esta forma se alcanza el supuesto de “strong ignorability” requerido para la utilización del propensity score al estimar ATT.

La forma funcional del modelo para el propensity score será determinada utilizando el algoritmo de estimación propuesto por Dehejia y Wahba (1999; 2002), también llamado test de estratificación (Caliendo y Kopeinig, 2008). El cual funciona de la siguiente forma: (1) primero se estima el propensity score con una especificación parsimoniosa; (2) luego se divide las observaciones en estratos

²³ Este supuesto (unconfoundedness en inglés) además requiere que en las variables explicativas no deben existir características (no observadas) de los individuos asociadas con los resultados esperados y con el tratamiento (Imbens y Wooldridge, 2009).

²⁴ Estos dos supuestos es a lo que Rosembaun y Rubin (1983) llamaron strong ignorability o strongly ignorable treatment assignment, que es fundamental para la utilización del propensity score para realizar matching.

²⁵ Unconfoundedness for controls.

²⁶ Weak overlap.

basados en $\hat{p}(X_i)$ –este hace referencia al propensity score estimado para diferenciarlo del real, $p(X_i)$ – de forma que no existan diferencias significativas entre la media del grupo de tratamiento y control (Caliendo, 2008); (3) después se comprueba que no existan diferencias significativas para los tratados y controles en las variables explicativas utilizadas para cada estrato, el no cumplir esta condición implica aumentar el número de estratos; (4) si esto último no funciona se modifica la forma funcional (incorporando interacciones, nuevas variables o quitando variables) y se repite el proceso hasta que se logre un adecuado balance para las variables explicativas, seleccionando dicho modelo²⁷.

Finalmente, para cumplir el criterio de overlap en la estimación se utilizó el criterio min-max, el cual consiste en eliminar todas las observaciones donde el propensity score sea menor que el mínimo y mayor que el máximo del grupo opuesto (Caliendo, 2008); el cual se modifica para la estimación de ATT siendo suficiente asegurar que para cada participante se pueda encontrar un no participante cercano (Caliendo, 2008), así este criterio consistió en encontrar el menor valor de $p(X)$ entre los tratados $p_0 = \min_{i;T_i=1} \hat{p}(X_i)$, para luego eliminar todos los controles bajo este punto de corte (Dehejia y Wahba, 1999). La idea detrás de esta metodología consiste en que unidades de control con valores muy bajos del $p(X)$ pueden ser muy diferentes de las unidades tratadas como para ser incluidas en el análisis.

B. Metodologías basadas en la estimación del propensity score

La forma más usual de estimación a través de matching propensity score consiste en encontrar el efecto utilizando como contrafactual para cada tratado, la observación con el $\hat{p}(X_i)$ más cercano, permitiendo la construcción de un grupo de control y tratamiento muy similar en la probabilidad de ser tratado (Lara et al, 2011). El principal problema de esta metodología conocida como “vecino más cercano” y sus variaciones, es que depende mucho de la forma funcional utilizada para estimar $\hat{p}(X_i)$ lo que sucede al asumir modelos paramétricos para la estimación del propensity score y la media condicional, generando una excesiva dependencia de la forma funcional adoptada para el propensity score. Además obliga a utilizar un número menor de observaciones, por lo que la robustez de las estimaciones se ve reducida, por esta razón en esta investigación se utilizaron tres metodologías que cumplen con la propiedad conocida como doble robustez (DR).

Las metodologías que poseen la propiedad conocida como doble robustez requieren de la especificación de dos modelos: uno que describa la población y otro que describa el proceso por el cual los datos son seleccionados para producir la muestra observada (Kang y Shafer, 2007). Esta idea

²⁷ Este proceso, además, determinará un número de estratos óptimos asociados al propensity score que serán utilizados más adelante.

desarrollada por Robins, Rotnitzky y Zhao (1994, 1995) es una metodología que da una doble oportunidad para generar estimadores consistentes y eficientes, si el modelo para el propensity score es el real o si el modelo para los resultados predichos es el correcto, donde en el caso que ambas especificaciones sean correctas es posible obtener un estimador semiparamétrico eficiente (Emsley et al., 2008). La propiedad de doble robustez alcanzada por los métodos desarrollados por Robins, Rotnitzky y Zhao (1994, 1995), no son el único camino para lograr la doble robustez (Kang y Shafer, 2007), por lo cual, en este estudio se enfocará en estimadores que *tienen esta propiedad pero no necesariamente corresponden a la familia de estimadores propuestos por dichos autores*.

Para terminar, es necesario mencionar que los estimadores que cumplan con la propiedad DR no serían la panacea, debido a que en casos reales cuando ambos modelos están mal especificados este tipo de modelos pueden comportarse peor que los modelos más simples utilizados (Kang y Shafer, 2007). Este último punto hace necesario ser cuidadosos con los resultados obtenidos a partir de este tipo de modelos.

a. Regresión lineal Ajustada (ANCOVA ajustado)

El primer estimador del impacto del tratamiento sobre los tratados corresponde al estimador propuesto por Kang y Shafer (2007) y Shafer y Kang (2008), quienes para estimar ATE proponen utilizar una estimación ANCOVA simple (lineal en los parámetros), incorporando una serie de ajustes, de forma que este análisis se transforma de una de las peores formas para estimar ATE a una de las mejores (Shafer y Kang, 2008). Los ajustes propuestos por ellos e incorporados en esta investigación son: (1) considerar la posibilidad de no linealidades en variables explicativas importantes; (2) incluir variables dummy identificando estratos basados en el propensity score estimado, por ejemplo, 4 variables dicotómicas para distinguir entre 5 categorías; (3) incluir interacciones entre variables explicativas medidas en la línea base y el tratamiento; (4) realizar las estimaciones a través de métodos de regresión. Si bien es cierto, ellos plantean este método para estimar ATE, la estimación de ATT requiere supuestos menos restrictivos así la extensión de este método para estimar el efecto promedio sobre los tratados es directa.

Siguiendo las recomendaciones antes descritas, la incorporación de $\hat{p}(X_i)$ no es directa lo que incrementa el desempeño del análisis de regresión²⁸, en nuestro caso el número de bloques obtenidos desde $\hat{p}(X_i)$ serán aquellos obtenidos mediante el test de estratificación, por lo que se asegura que dichos bloques estén balanceados para tratados y controles. Este método presenta la propiedad DR al

²⁸ Esto aborda el problema asociado al método donde se incorpora el propensity score como una variable explicativa más, y que no lo hacen recomendable, ya que el propensity score, como tal, no tendría un significado sustantivo (Imbens y Wooldridge, 2009).

incorporar funciones del propensity score estimado en el modelo para explicar los resultados (Kang y Shafer, 2007) pero a diferencia de los estimadores propuestos por Robins, Rotnitzky y Zhao (1994, 1995), esta tiene la facultad de disminuir considerablemente la dependencia de la forma funcional utilizada para la estimación del propensity score.

Antes de continuar, es necesario diferenciar los métodos de regresión de un análisis de regresión simple, haciendo referencia a la diferencia entre lo que se entiende por un “coeficiente de regresión” que es un parámetro perteneciente a un modelo para la media condicional de un tratamiento dado un set de variables explicativas, en otras palabras, sería el parámetro en una regresión simple asociado a una variable explicativa que indica si el individuo fue afectado por el tratamiento o no. En cambio, un método de regresión, corresponde a una estimación que busca obtener la media poblacional de un determinado tratamiento (Shafer y Kang, 2008), este consiste en estimar un modelo para los tratados (Y_{i1}), distinto a un modelo para los no tratados (Y_{i0}) para luego, en el caso de ATT, estimar el impacto de la siguiente forma

$$ATT = \frac{1}{N_T} \sum_i T_i (\hat{Y}_{i1} - \hat{Y}_{i0}),$$

donde \hat{Y}_{it} corresponde a la predicción basada en el modelo para los tratados ($T = 1$), y para los controles ($T = 0$). Este tipo de estimación es preferida a la primera debido a que permite la utilización de diferentes formas funcionales para cada grupo y evita que se modele el impacto del tratamiento de forma paramétrica (Shafer y Kang, 2008). Adicional a lo anterior, la estimación utilizando métodos de regresión, incorpora automáticamente todas las interacciones entre la variable de tratamiento y las variables explicativas utilizadas, así para hacer equivalente un modelo de regresión debería incorporarse todas las interacciones entre la variable de tratamiento y las variables explicativas utilizadas.

b. Estratificación ajustada por análisis de regresión

La segunda metodología basada en propensity score, corresponde al método comúnmente llamado subclasificación o estratificación propuesto originalmente por Rosembaun y Rubin (1983), la idea tras este estimador es dividir en estratos la muestra basándose en los valores de $\hat{p}(X_i)$. Luego es posible analizar los datos dentro de cada estrato como si el $\hat{p}(X_i)$ fuera constante y los datos, dentro de cada estrato, pudieran ser interpretados como provenientes desde un experimento completamente aleatorizado (Imbens y Wooldridge, 2009). Con la finalidad de eliminar cualquier desequilibrio aún existente después de la estratificación se utilizará análisis de regresión en la estimación de cada estrato (Lunceford y Davidian, 2004), lo que además permitirá obtener estimadores considerablemente más

flexibles y robustos, siendo una de las alternativas más atractivas en la práctica (Imbens y Wooldridge, 2009), este tipo de estimación muestra una mayor precisión que otros métodos como ANCOVA o métodos que combinan propensity score con métodos de regresión (Shafer y Kang, 2008). De esta forma, el procedimiento de subclasificación utilizado para estimar ATT estima un forma funcional,

$$Y_i^{(j)} = \delta_0^{(j)} + \beta_T^{(j)} T_i + \beta_X^{(j)} X_i + \varepsilon_i^{(j)}$$

Donde T_i corresponde al tratamiento; X_i corresponde a un set de variables explicativas; y ε_i corresponde a un error aleatorio, para todos los alumnos $i = 1 \dots N$, y para cada estrato $j = 1 \dots J$. Luego el efecto promedio del tratamiento sobre los tratados se obtiene a través de un promedio ponderado definido como

$$\hat{\tau} = \sum_{j=1}^J \left(\frac{N_{j1}}{N_1} \right) \hat{\tau}_j$$

donde N_1 es el número total de individuos tratados, y N_{j1} representa el número de individuos tratados presentes en cada estrato.

Un punto crítico de esta metodología es la selección del número de estratos, como se menciona en Imbens y Wooldridge (2009); Kang y Shafer (2007); Stuart y Rubin (2007); Lunceford y Davidian (2004) es común la utilización de solo 5 estratos²⁹, donde dependiendo del tamaño de la muestra sería necesario utilizar un número mayor para refinar aún más el efecto de la estratificación (Lunceford y Davidian, 2004), pero no existe una regla definida para determinar cuál es el número adecuado de estratos (Stuart, 2010) por lo que es un desafío determinar el número de estratos óptimos. En nuestro caso particular se optó por utilizar los estratos estimados a través del test de estratificación propuesto por Dehejia y Wahba (1999; 2002), antes descrito, de esta forma se obtienen estratos mucho más refinados en el sentido que además de balancear $\hat{p}(X_i)$ para tratados y no tratados dentro de cada estrato, se asegura que las variables explicativas pre-tratamiento no sean estadísticamente distintas entre los dos grupos al interior de cada estrato, por lo que el supuesto de completa aleatorización al interior de cada estrato (Rosembaun y Rubin, 1983) es mucho más factible.

²⁹ Este número de estratos fue propuesto inicialmente por Rosembaun y Rubin (1983).

c. Regresión ponderada por el inverso de la probabilidad³⁰

El último estimador utilizado corresponde al propuesto en Wooldridge (2007; 2010) –conocido también como Wooldridge’s doubly-robust– estimador que también posee la propiedad de doble robustez dada la utilización de un modelo para predecir la probabilidad de ser tratado y uno distinto para los resultados. Acá se utilizó la implementación de este estimador disponible en el programa STATA 13, la cual consiste en tres pasos: (1) Estimar los parámetros del modelo para predecir la probabilidad de ser tratado y estimar los pesos a través del inverso de la probabilidad estimada (IPW); (2) utilizando los IPW estimados, se ajustan modelos de regresión ponderados para cada condición de tratamiento $T = 1, 0$ y luego obtienen los resultados predichos para cada sujeto; (3) Se estima las medias de cada condición de tratamiento, para luego estimar el ATT a través de la diferencia de medias de los resultados predichos para los tratados.

d. Test de validez

El supuesto de selección en observables (o unconfoundedness) no es testeable, debido a que no se tiene datos de $Y_i(0)$ para aquellos que reciben el tratamiento, ni de $Y_i(1)$ para aquellos pertenecientes al grupo de control (Imbens y Wooldridge, 2009), estos autores muestran que una de las formas de testear este supuesto puede ser estimando el impacto del tratamiento en un resultado que haya sido medido antes que el tratamiento fuera aplicado y que a la vez sea una variable relacionada al resultado post-tratamiento. La implementación de este test seguirá lo planteado por Lara, Mizala y Repetto (2011), quienes proponen estimar este “falso experimento” tomando el resultado SIMCE del año base como variable dependiente, y el tratamiento (pertenecer a un LSPE) más las variables pre tratamiento se utilizan como variables explicativas, este falso experimento será estimado para todos los métodos basados en propensity score acá utilizados. En este caso, si los grupos de control están correctamente especificados no será significativo el falso tratamiento, de lo contrario habrían indicios que condiciones no observables están explicando la diferencia y que el grupo control podría ser perfeccionado.

e. Heterogeneidad del efecto del tratamiento

Adicionalmente, las estimaciones antes descritas serán realizadas para la distribución por nivel socioeconómico (NSE) de los estudiantes. Para esto se estimará el ATT para cada quintil de nivel

³⁰Inverse-probability-weighted regression adjustment.

socioeconómico estimado³¹, lo que permitirá comprobar si existe algún efecto –y cuál es la magnitud de este– diferenciado por nivel socioeconómico en las pruebas SIMCE analizadas.

C. Estimación por Diferencias en Diferencias (DID) y Changes in Changes (CIC)

Dadas las limitantes de los métodos basados en el propensity score derivados principalmente de los supuestos utilizados (selección en observables y traslape), y dada la existencia de una medición pre-tratamiento y post-tratamiento de la variable de resultado (SIMCE) es posible aplicar la metodología de diferencias en diferencias (Ashenfelter y Card, 1985), que tiene la particularidad de permitir controlar por heterogeneidad no observada (Heckman et al., 1998) y por sesgos que permanecen constantes a través del tiempo.

Un supuesto fundamental de la metodología de diferencias en diferencias dice que en ausencia del tratamiento los resultados promedios de cada grupo (tratados y controles) deben mostrar trayectorias paralelas a través del tiempo (Abadie, 2005). Así, en este caso, donde algunas de las características individuales no observadas de aquellos seleccionados en los LSPE (tal como la motivación) no necesariamente eran las mismas antes del tratamiento para ambos grupos, puede llevar a violar el supuesto de trayectorias paralelas de los resultados. Por esta razón, se utilizó una versión generalizada del método de diferencias en diferencias que no asume el supuesto de trayectorias paralelas y bajo ciertos supuestos además sería consistente y asintóticamente normal, llamada changes in changes (CIC) propuesto por Athey e Imbens (2002; 2006). En palabras simples, el método compara grupos de cuantiles que tenían los mismos no observables en un periodo antes de tratamiento, asumiendo que los cuantiles que tienen el mismo ranking de no observables antes de tratamiento tendrán el mismo ranking después de tratamiento. Este supuesto permite que la distribución de no observables cambie a través de los grupos pero no a través del tiempo (Lara, Mizala y Repetto, 2011), así los no observables que pueden afectar las trayectorias de los resultados, independiente del tratamiento, serían consideradas en esta generalización del estimador de diferencias en diferencias.

Esta metodología, adicionalmente, permite estimar cuál es el efecto a lo largo de toda la distribución de puntajes, esto permitirá determinar a qué alumnos impacta –en términos de puntaje– con mayor intensidad el tratamiento en consideración. Lo anterior será realizado estimando el impacto del tratamiento (controlando por variables explicativas) para cada centil de la distribución de puntajes.

³¹ Variable construida a través de análisis factorial utilizando la educación del padre, de la madre y el ingreso del hogar antes del tratamiento.

Finalmente, es necesario enfatizar que esta metodología tiene como objetivo aumentar la validez externa de las estimaciones, debido a que los métodos basados en el propensity score solo tienen validez dentro del área de soporte común. Además de proporcionar una metodología capaz de estimar el efecto del tratamiento utilizando controles que pudieran haber sido afectados por el tratamiento – como los efectos pares en este caso– si es que se posee una medición de antes y después de aplicado el tratamiento.

D. Variables utilizadas

Las variables utilizadas son prácticamente las mismas en toda la investigación, solo existen algunas diferencias debido a los modelos que se utilizan (por ejemplo, la inclusión de variables de selección). De esta forma, los modelos basados en el propensity score requieren la especificación de dos modelos, en cambio los modelos CIC estimados requieren la utilización de variables que hayan sido medidas en dos tiempos, antes y después del tratamiento. Por esto, a continuación se describe por separado las variables utilizadas en cada modelo.

a) Modelos basados en el propensity score

Para estimar el efecto promedio del tratamiento sobre los tratados con las metodologías basadas en el propensity score, como ya fue mencionado, se requiere especificar dos modelos en este caso un modelo que determine la probabilidad de haber sido tratado y uno que describa la función de producción escolar.

De esta forma, para estimar la probabilidad de ser tratado se utilizó como variable dependiente la condición de asistencia a un LSPE o no. Las variables explicativas se pueden dividir en dos grupos (todas medidas antes del tratamiento). Las primeras corresponden a características individuales o del hogar; y segundo grupo corresponde a un conjunto de variables asociadas al establecimiento. Dentro de las variables a nivel individual o del hogar se encuentra el promedio del puntaje SIMCE en lenguaje y matemática estandarizado; la educación del padre y madre³²; el ingreso del hogar estandarizado; el número de libros en el hogar; sexo; las condiciones previas de haber repetido antes de cuarto básico y asistir a educación pre escolar; el promedio de las evaluaciones finales de 5° y 6° básico junto al cuadrado de este; las expectativas de los padres sobre la continuidad de estudios de los alumnos; y finalmente se incluyó variables que pudieran estar relacionadas directa o indirectamente con la elección

³²Variable continua que toma valores entre 0 y 21 años, construida a partir de una serie de variables dicotómicas existentes en el cuestionario de padres de la prueba SIMCE. Cabe destacar que para la cohorte 2002-2008, debido a la gran pérdida de datos sobre educación de los padres en el año 2002 (ocasionada por la forma en que se preguntó la educación de los padres) se imputó esta utilizando la variable educación construida para el año 2008, logrando recuperar aproximadamente 100mil datos de ambos padres.

del colegio, específicamente se incluyó cuál fue la razón –declarada por los padres en 2do medio– de ingreso de los alumnos a los colegios. El segundo grupo –variables del establecimiento– incluyó la dependencia institucional del establecimiento; la matrícula de educación básica; la tasa de repitencia de educación básica y la condición de ruralidad del establecimiento.

El modelo para la función de producción escolar –estimada a través de métodos de regresión o regresiones simples, dependiendo de la metodología– utilizó como variable dependiente el promedio SIMCE estandarizado de los sectores lenguaje y matemáticas en 2º medio. Las variables explicativas utilizadas nuevamente se pueden dividir en dos grupos (siempre medidas antes del tratamiento). El primero corresponde a las características del individuo o del hogar, dentro de estas se consideró la educación de la madre y padre; el ingreso del hogar estandarizado; género; el número de libros en el hogar; un indicador si repitieron antes de 4to básico; si asistieron a educación pre escolar (pre-kínder o kínder). El segundo grupo corresponde a un conjunto de variables asociadas, donde se incorporó la dependencia institucional del establecimiento; la condición de ruralidad de los establecimientos; el promedio por curso de la escolaridad de la madre y padre; el promedio por curso del ingreso monetario de las familias (estandarizado); la matrícula total de educación básica del colegio (estandarizada); el número de alumnos y tasa de repitencia de cada curso.

b) Modelos de diferencias en diferencias (CIC)

Como ya fue mencionado lo más importante al momento de plantear un modelo por diferencias en diferencias es que las variables utilizadas hayan sido medidas antes y después del tratamiento en cuestión. Para esto nuevamente se utilizó los dos grupos de variables antes mencionados, considerando características del individuo o del hogar ahora medidas antes y después del tratamiento, dentro de estas se consideró la educación de la madre y padre; el ingreso del hogar estandarizado; sexo; el número de libros en el hogar; un indicador si repitieron antes de 4to básico; si asistieron a educación pre escolar (pre-kínder o kínder). El segundo grupo nuevamente corresponde a variables asociadas al establecimiento medidas antes y después del tratamiento, dentro de las que se incorporó la dependencia institucional del establecimiento y la condición de ruralidad de los establecimientos.

Adicionalmente, se generaron modelos que además controlaron por variables de resultado intermedias³³, estas se pueden definir como variables que son afectadas por el tratamiento pero que no corresponden al objetivo real o final de la intervención o tratamiento realizado (más adelante se profundiza en este tema). Las variables identificadas como intermedias que fueron consideradas debido

³³ En inglés son conocidas como intermediate outcome, concomitant variables o post treatment variables.

a que se poseía mediciones antes y después del tratamiento, fueron el efecto par de la educación de la madre y del padre; el efecto par del ingreso del hogar; y el efecto par de la habilidad individual estimada.

Para estimar el efecto par de la habilidad individual, se utilizó la habilidad académica no observada de cada estudiante obtenida por medio de la estimación del residuo individual de la función de producción de los resultados de la prueba SIMCE en 4° básico, a través de un modelo lineal jerárquico de dos niveles de intercepto aleatorio³⁴. La elección del modelo multinivel descrito se debe a que asumir las características entre colegios como similares no sería realista, por lo que utilizar el mismo error e intercepto para cada colegio sesgaría las estimaciones³⁵. Un punto muy importante es que la habilidad académica no observada al estar siendo medida a través del residuo de una función de producción escolar, puede estar englobando muchos no observables, y no solo la habilidad académica, tal como puede ser la motivación de los alumnos, la perseverancia, y muchos otros no observables asociados con habilidades cognitivas y no cognitivas que pueden afectar el futuro académico de los estudiantes tanto en el corto como en el largo plazo.

E. Estimación de la varianza de ATT y otras consideraciones

Para estimar los errores estándar de todas las metodologías antes descritas se utilizó bootstrap con reemplazo a nivel de las escuelas, esto último dada la alta probabilidad que exista clustering a nivel de las escuelas (Lara et al., 2011). Tal como muestra Abadie e Imbens, 2006, la utilización de un estimador de matching con corrección de error removería el sesgo condicional asintótico de la varianza producido al utilizar bootstrap sobre estimadores de matching. Esta corrección se debe a la combinación de las ventajas y desventajas de los estimadores de matching y de estimadores de regresión (Abadie y Imbens, 2006), por esta razón es probable que los métodos acá utilizados también cumplan con esta característica y bootstrap nos provea de estimadores insesgados de la varianza.

Adicionalmente, en las estimaciones de cada año sólo se consideraron los establecimientos con 15 o más estudiantes que rinden válidamente cada prueba, de esta forma se reduce los problemas de volatilidad en los puntajes a través del tiempo, asociados a factores no observables de carácter aleatorio, situación que se incrementa entre los establecimientos de menor tamaño (Kane, Staiger y Geppert, 2002; Chay, McEwan y Urquiola, 2005)

³⁴ El modelo incluyó las variables: número de libros, ingreso del hogar estandarizado, educación del padre y de la madre, efecto par del ingreso, dependencia institucional, número de alumnos en el curso, matrícula total, y controles por la región a la que pertenece el colegio.

³⁵ La metodología multinivel, en pocas palabras, consiste en introducir un término de error específico para cada colegio lo que levanta el supuesto de independencia condicional de las observaciones.

V. Principales resultados

A). Quiénes asisten a los Liceos Públicos de Excelencia

Los estudiantes que asisten a los LSPE por definición son alumnos que poseen altos rendimientos durante prácticamente toda su enseñanza previa al ingreso a dichos establecimientos, cosa que se verifica al ver los puntajes SIMCE y el promedio de notas de 5to y 6to básico (Ver Anexo 4 y 5), Adicionalmente muestran una capacidad por sobre el promedio del país, que se ve reflejada en las expectativas de los padres sobre la educación que lograrán, donde un 81% en 2002 –y 85% en 2006– creen que sus hijos asistirán a la Universidad o tendrán algún grado superior.

Un punto importante es que al parecer las movilizaciones de los años 2006 y 2011, parecen no haber tenido el efecto que se hubiera esperado –coartando la entrada de los alumnos más hábiles a este tipo de colegios– debido a los largos periodos en que muchos LSPE permanecieron movilizados, así el tipo de alumnos que ingresa a estos establecimientos han continuado siendo estudiantes con habilidades muy superiores a la media del país.

Con respecto a las razones para ingresar a los LSPE, la creencia popular que mayoritariamente ingresarían hijos o familiares de alumnos o ex alumnos al parecer no tendría sustento, por lo menos en las dos cohortes analizadas, y primarían principalmente razones como el prestigio y la excelencia académica mostrada por éstos.

Finalmente, y un punto no menor, se observa que la cercanía a los establecimientos no sería una de las razones principales por las cuales suelen ingresar a los LSPE, en este sentido Bucarey et al, (2014) muestran que para el Instituto Nacional la mayor parte de los alumnos tendría asociado altos tiempos de viaje desde el hogar al establecimiento y viceversa.

B). ¿Oportunidad para los niños vulnerables?

Los estudiantes que asisten a Liceos Públicos de Excelencia no se concentra principalmente entre las familias más vulnerables del país (ver Tabla 2) análisis que sería consistente para ambas cohortes estudiadas; al contrario, aproximadamente un tercio de estos estudiantes cuando se encontraba en 4º básico pertenecía a las familias del quintil de mayores niveles educacionales de los padres y de ingresos del país, y otro tercio al quintil inmediatamente anterior; mientras que sólo un 6% de este grupo de estudiantes pertenecía al 20% de los estudiantes de menores condiciones socioeconómicas de ambas cohortes y un 12-13% al quintil siguiente. De esta forma, los estudiantes que asisten a los Liceos Selectivos Públicos de Excelencia corresponderían mayoritariamente a jóvenes de familias de clase media-alta, cuyo promedio en el Índice Socioeconómico –estandarizado a nivel nacional con

media cero y desviación estándar uno- es de 0,5, lo cual implica que el estudiante promedio de este tipo de colegios pertenecería al percentil 75 de la distribución socioeconómica del total de los estudiantes del país.

Estos resultados son consistentes con la experiencia de Estados Unidos, donde es poco frecuente que los estudiantes más vulnerables asistan a colegios públicos selectivos (Neild, 2002); ello se explica porque en sistemas escolares donde el capital socioeconómico y cultural está estrechamente vinculado con los logros educativos alcanzados en los primeros años de escolaridad, reflejado en la alta persistencia de los resultados a través del trayectoria educacional de los alumnos en Chile (Valenzuela, Allende, Sevilla y Egaña, 2014), es bastante improbable que la generación de establecimientos públicos de educación media que tengan un alto componente selectivo sean una oportunidad de mejoramiento para los niños más vulnerables, más bien, el acceso de este tipo de estudiantes se percibe como una situación excepcional.

Tabla III. Características Seleccionadas de los Estudiantes que Asisten a Liceos Públicos de Excelencia medidas en el año base para cada panel de datos (2002-2008; 2006-2012).

	Estudiantes Liceos Públicos de Excelencia (2008)	Estudiantes de 2º Medio de Otros Colegios (2008)	Test-t	Estudiantes Liceos Públicos de Excelencia (2012)	Estudiantes de 2º Medio de Otros Colegios (2012)	Test-t
Sexo (0 hombre; 1 Mujer)	0.49 (0.50)	0.52 (0.50)	-4.15***	0.49 (0.50)	0.52 (0.50)	-3.95***
<i>Capital Social, Cultural y Económico</i>						
Libros menor a 10	0.15 (0.35)	0.31 (0.46)	-23.15***	0.30 (0.46)	0.46 (0.50)	-18.23***
Libros mayor que 11 y menor a 50	0.43 (0.50)	0.42 (0.49)	1.07	0.46 (0.50)	0.38 (0.48)	9.43***
Libros mayor que 51	0.42 (0.49)	0.27 (0.44)	22.87***	0.24 (0.43)	0.16 (0.37)	12.28***
Educación madre	12.48 (2.97)	11.01 (3.73)	26.04***	13.10 (2.71)	11.68 (3.57)	22.00***
Educación padre	12.56 (3.19)	11.05 (3.86)	25.64***	13.26 (3.00)	11.79 (3.77)	21.57***
Nivel socioeconómico (NSE)**	0.51 (0.78)	0.14 (0.99)	19.80***	0.49 (0.73)	0.15 (1.01)	18.83***
<i>Quintiles de NSE</i>						
Quintil I de NSE†	6%	21%	-19.67***	6%	20%	-19.18***
Quintil II de NSE†	13%	20%	-8.96***	12%	20%	-11.21***
Quintil III de NSE†	24%	25%	-1.17	21%	20%	1.22
Quintil IV de NSE†	27%	14%	19.42***	34%	19%	19.99***
Quintil V de NSE†	30%	20%	12.95***	27%	20%	9.35***
<i>Dependencia institucional del establecimiento</i>						
Dependencia público	0.38 (0.49)	0.49 (0.50)	-14.38***	0.33 (0.47)	0.44 (0.50)	-12.33***
Dependencia particular Subvencionado	0.56 (0.50)	0.42 (0.49)	19.67***	0.64 (0.48)	0.47 (0.50)	18.67***
Dependencia particular pagado	0.05 (0.22)	0.09 (0.28)	-8.92***	0.03 (0.17)	0.09 (0.28)	-11.51***
<i>Razones de ingreso al establecimiento</i>						

Razón de ingreso: Cercanía del establecimiento	0.12 (0.32)	0.35 (0.48)	-0.23	0.11 (0.31)	0.37 (0.48)	-26.92***
Razón de ingreso: Prestigio del establecimiento	0.90 (0.30)	0.48 (0.50)	0.42			
Razón de ingreso: Buenos resultados en PSU	0.59 (0.49)	0.11 (0.31)	0.48			
Razón de ingreso: Otros miembros de la familia estudian o estudiaron en el establecimiento	0.19 (0.39)	0.31 (0.46)				
Razón de ingreso: Excelencia académica y resultados PSU/SIMCE				0.96 (0.19)	0.40 (0.49)	57.96***
Razón de ingreso: Hermanos estaban en establecimiento				0.11 (0.31)	0.17 (0.37)	-8.03***
Expectativas de los padres sobre educación futura						
Expectativa padres: Básica, media incompleta o técnico	0.09 (0.28)	0.31 (0.46)	-32.42***	0.06 (0.24)	0.23 (0.42)	-22.56***
Expectativa padres: media CH, IP o CFT [‡]	0.11 (0.31)	0.17 (0.38)	-10.68***	0.09 (0.29)	0.17 (0.37)	-10.94***
Expectativa padres: Universidad o postgrado	0.81 (0.39)	0.52 (0.50)	38.03***	0.85 (0.36)	0.60 (0.49)	27.75***
Nº de Estudiantes	4591	164184		3236	161288	

*Desviación estándar en paréntesis.

**Construido a través de análisis factorial utilizando la educación de ambos padres y el ingreso del hogar.

†Los quintiles fueron estimados desde el índice de nivel socioeconómico construido.

‡ Científico Humanista; Instituto Profesional y Centro de Formación Técnica, respectivamente.

*** Significancia al 1%; **Significancia al 5%; * Significancia al 10%

C). ¿Oportunidad para los estudiantes de escuelas municipales?

La construcción del panel de estudiantes de 4º básico a 2º medio también posibilita determinar si la matrícula de los Liceos Selectivos Públicos de Excelencia proviene especialmente de estudiantes que asistían previamente a las escuelas básicas municipales. La información de la Tabla III permite concluir que la mayor parte de los estudiantes de los Liceos Emblemáticos proviene de colegios particulares, incluso algunos de ellos de colegios particulares pagados –un 4% en 2002 y 3% en 2006–; por el contrario, sólo un 38% (33% en 2006) de ellos se educó en la educación municipal durante sus primeros años de escolaridad, implicando que la existencia de Liceos de Excelencia en el sector público no aumenta las posibilidades que los estudiantes del sector municipal accedan a un establecimiento público de calidad, puesto que los mecanismos de selección que éstos aplican parecen no considerar una discriminación positiva para aquellos provenientes del mismo sector administrativo, primando más bien mecanismos de selección comunes a todos los estudiantes, tales como información del desempeño académico previo (notas de 5to y 6to año básico en general) y de los resultados en pruebas de selección.

Por otra parte, estos resultados también permiten concluir que aquellas familias de mayores condiciones socioeconómicas que privilegian el ingreso temprano a establecimientos privados no descartan indefinidamente el ingreso posterior de sus hijos a establecimientos públicos, como son los Liceos Emblemáticos, esta evidencia es contraria a la existencia de un estigma generalizado sobre la

educación pública, y dicho interés se debería principalmente a razones de prestigio (90% en 2002) y excelencia académica (56% en 2002 y 96% en 2006). Esto da evidencia que una eventual mejora de la educación pública, con colegios prácticamente gratuitos³⁶ y de excelencia en general atraería alumnos provenientes de todos los sectores económicos.

Asimismo, es probable que la alternativa de los Liceos Selectivos Públicos de Excelencia también responda a las preferencias de un conjunto acotado de familias de alto nivel socioeconómico que, teniendo los recursos para acceder a la educación particular pagada o particular subvencionada con financiamiento compartido, valorizan la asistencia de sus hijos a establecimientos de mayor diversidad sociocultural, de orientación pública y laica, pero no descuidan el logro de un alto estándar de calidad académica. Adicionalmente, los resultados muestran que los alumnos no necesariamente corresponden a hijos o familiares de alumnos o ex alumnos de dichos colegios cuyo nivel socioeconómico es mayor y que tendrían una preferencia por dichos colegios, como podría pensarse.

Estos resultados anticipan que la ampliación en la oferta de Liceos Selectivos Públicos de Excelencia en el país, como son los Liceos Bicentenario de Excelencia, podrían no responder principalmente a una alternativa de educación secundaria de alta calidad para los estudiantes que inician su educación básica en los colegios públicos, ni tampoco para los estudiantes de menores condiciones sociales, sino que por el contrario, podrían ser una alternativa que acoge principalmente a los estudiantes de alto desempeño académico que provienen mayoritariamente de familias con mejores condiciones socioeconómicas y especialmente de colegios privados –subvencionados y particulares–.

D). Impacto del tratamiento en la prueba SIMCE de los estudiantes que asisten a los LSPE

a) Efecto promedio de la participación en los LSPE sobre los tratados, estimadores basados en el propensity score

i. Probabilidad de asistir a un LSPE

Tal como fue mencionado la forma funcional utilizada para cada cohorte, sector y muestras estimadas (grupo de control ex compañeros y ex compañeros emparejados) fueron obtenidas a través de la metodología conocida como test de estratificación, propuesta por Dehejia y Wahba (1999; 2002), por esta razón, las variables utilizadas en cada especificación pueden cambiar.

³⁶ Valenzuela, Bellei y De los Ríos (2009) concluyen que el financiamiento compartido se aplica en diversos liceos municipales del país, sin embargo, los montos mensuales cobrados son bajos (una séptima parte de los cobros exigidos en el promedio de los colegios particulares subvencionados).

Los resultados para ambas cohortes analizadas utilizando el grupo de control³⁷ se muestran en la Tabla IV. Los resultados entre los distintos sectores para cada cohorte se ve que son muy consistentes, en cambio los resultados entre ambas cohortes muestran algunas diferencias principalmente en las variables a nivel individual (educación de los padres y número de libros). Pero las variables más relevantes en la determinación de la probabilidad de pertenecer a un LSPE muestran ser consistentes entre ambas cohortes, dentro de estas se cuenta el resultado previo en las pruebas SIMCE, las razones de ingreso a los establecimientos informadas por los padres, las expectativas de los padres sobre la continuidad educacional y la tasa de repitencia de educación básica del colegio del año base. Destacando la relevancia que tendrían las razones de ingreso esbozadas por los padres especialmente el prestigio y los buenos resultados en pruebas estandarizadas (SIMCE o PSU). Además se releva el gran efecto que tendría sobre la probabilidad de ingreso a un LSPE el tipo de compañeros con los que se compartía antes del ingreso a estos colegios, reflejado principalmente en la tasa de repitencia de educación básica³⁸.

Tabla IV. Probabilidad de asistir a un LSPE

VARIABLES	2002-2008		2006-2012	
	Lenguaje	Matemática	Lenguaje	Matemática
Promedio SIMCE Lenguaje (estandarizado)	0.148*** (0.046)		0.235*** (0.053)	
Promedio SIMCE Matemática estandarizado		0.270*** (0.044)		0.270*** (0.053)
<i>Capital Social, Cultural y Económico</i>				
Educación madre	0.001 (0.012)	-0.000 (0.012)	0.010 (0.014)	0.010 (0.014)
Educación padre	0.016 (0.011)	0.018 (0.011)	0.045*** (0.014)	0.044*** (0.014)
Ingreso del Hogar (estandarizado)	-0.194*** (0.037)	-0.202*** (0.037)	-0.311*** (0.047)	-0.359*** (0.043)
Libros mayor que 11 y menor a 50	0.222*** (0.081)	0.216*** (0.081)	0.083 (0.072)	0.082 (0.072)
Libros mayor que 51	0.287*** (0.086)	0.276*** (0.086)	0.183** (0.087)	0.190** (0.087)
<i>Características Individuales</i>				
Repite antes de 4to básico (0 no; 1 si)	-0.397 (0.391)	-0.404 (0.388)	-0.086 (0.315)	-0.035 (0.317)
Asistió a educación pre escolar (0 no; 1 si)	0.058 (0.291)	0.088 (0.288)		
Promedio notas (5to y 6to básico)	-0.680 (2.581)	0.034 (2.584)	3.356 (3.135)	3.702 (3.122)
Promedio notas al cuadrado (5to y 6to básico)	0.205 (0.206)	0.142 (0.206)	-0.094 (0.251)	-0.124 (0.249)
<i>Razones de ingreso al establecimiento</i>				
Razón de ingreso: Cercanía del establecimiento	-0.789*** (0.072)	-0.774*** (0.072)	-1.135*** (0.088)	-1.118*** (0.088)
Razón de ingreso: Otros miembros de la familia estudian o estudiaron en el establecimiento	-0.145** (0.062)	-0.140** (0.062)		
Razón de ingreso: Prestigio del establecimiento	1.341***	1.335***		

³⁷ Recordar que este grupo de control considera a la subpoblación de compañeros de curso en 4º básico (año base) de todos los alumnos tratados y que no se encuentran estudiando en liceos selectivos públicos de excelencia en el año de tratamiento (2º medio), que tuvieron un promedio de notas superior a 5,5 en la media de los promedios generales de notas correspondientes a 5º y 6º básico, y que se encontraron en el nivel de logro elemental o adecuado en la prueba SIMCE de lenguaje y matemática en 4º básico.

³⁸ Los resultados del área de soporte común se muestran en el Anexo 6.

	(0.076)	(0.077)		
Razón de ingreso: Buenos resultados en PSU	1.908***	1.905***		
	(0.054)	(0.054)		
Razón de ingreso: Excelencia académica y resultados PSU/SIMCE			3.207***	3.198***
			(0.142)	(0.142)
Razón de ingreso: Asisten niños del mismo ambiente y realidad familiar			-0.330***	-0.319***
			(0.087)	(0.087)
Razón de ingreso: Hermanos estaban en establecimiento			-0.074	-0.078
			(0.091)	(0.091)
<i>Expectativas de los padres sobre educación futura</i>				
Expectativa padres: media CH, IP o CFT	0.312***	0.316***	0.251	0.235
	(0.114)	(0.114)	(0.156)	(0.155)
Expectativa padres: Universidad o postgrado	0.595***	0.589***	0.538***	0.534***
	(0.095)	(0.096)	(0.134)	(0.133)
<i>Características del establecimiento</i>				
Dependencia público	0.489***	0.503***	0.567***	0.574***
	(0.059)	(0.059)	(0.073)	(0.073)
Dependencia particular pagado	-1.109***	-1.146***	-0.430**	
	(0.134)	(0.134)	(0.169)	
Matricula educación básica del colegio (estandarizada)	0.025*	0.022*	0.026*	0.025*
	(0.013)	(0.013)	(0.015)	(0.015)
Tasa de repitencia educación básica del colegio	6.380***	6.260***	7.640***	7.834***
	(1.346)	(1.354)	(1.318)	(1.311)
Ruralidad (0 Urbano; 1 Rural)	-0.204	-0.167	0.248	0.249
	(0.175)	(0.175)	(0.177)	(0.177)
Constante	-8.269	-10.374	-23.008**	-24.041**
	(8.092)	(8.101)	(9.790)	(9.752)
Número de Observaciones	17,050	17,036	11,875	11,895
R-cuadrado	0.327	0.328	0.321	0.321

Errores estándar robustos en paréntesis

*** Significancia al 1%; **Significancia al 5%; * Significancia al 10%

ii. Estimación del efecto del tratamiento sobre los tratados

Los resultados del ATT se presentan en las Tablas V y VI, en estas se ve que para lenguaje el efecto medido implica que los puntajes de los alumnos pertenecientes a los LSPE sería aproximadamente entre 0,17 y 0,22 desviaciones estándares mayor que el grupo de control, esto representaría casi un quinto de desviación estándar que en términos de puntaje SIMCE, asistir a los LSPE estaría asociado con un aumento de casi 10 puntos en la prueba de lenguaje.

Para el sector matemática, se ve un efecto bastante diferenciado entre las dos cohortes analizadas llegando a casi 0,25 desviaciones estándar el efecto en la cohorte 2002-2008, que se reduciría a prácticamente la mitad (0,11 – 0,13 dependiendo de la metodología) en la cohorte 2006-2012, y que representaría un aumento aproximadamente entre 7 y 13 puntos en la prueba SIMCE para los alumnos que asisten a los LSPE.

Si bien es cierto, se observa una diferencia entre las cohortes que es bastante considerable en el caso de matemática, es necesario mencionar que para ninguna de las metodologías utilizadas estas mostraron ser significativas.

Con respecto al test de validez es posible observar que para ambas cohortes y sectores analizados no se encuentra efecto alguno del tratamiento sobre el test medido antes en el año base (2002 y 2006), esto adicionalmente es consistente para todas las metodologías utilizadas. De esta forma, es posible afirmar que es muy probable que el supuesto de selección en observables requerido para las estimaciones realizadas a través del propensity score se cumpla y por ende los resultados obtenidos no estén influenciados por variables no observables.

Tabla V. Efecto promedio del tratamiento sobre los tratados sector Lenguaje (ATT), modelos basados en propensity score.

	Panel SIMCE		Panel SIMCE	
	2002-2008	2006-2012	2002-2008	2006-2012
	Efecto estimado (ATT)		Test de validez	
Estratificación	0.215*** (0.067)	0.175*** (0.046)	0.016 (0.017)	0.015 (0.019)
ANCOVA ajustado	0.211*** (0.067)	0.183*** (0.051)	0.021 (0.018)	0.024 (0.019)
Wooldridge's double-robust	0.199*** (0.061)	0.169*** (0.052)	0.009 (0.010)	0.013 (0.010)

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

Tabla VI. Efecto promedio del tratamiento sobre los tratados sector Matemática (ATT), modelos basados en propensity score.

	Panel SIMCE		Panel SIMCE	
	2002-2008	2006-2012	2002-2008	2006-2012
	Efecto estimado (ATT)		Test de validez	
Estratificación	0.249*** (0.078)	0.125*** (0.049)	0.025 (0.021)	-0.004 (0.020)
ANCOVA ajustado	0.252*** (0.084)	0.126** (0.058)	0.029 (0.021)	0.004 (0.020)
Wooldridge's double-robust	0.234*** (0.076)	0.107* (0.057)	0.011 (0.012)	-0.010 (0.009)

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

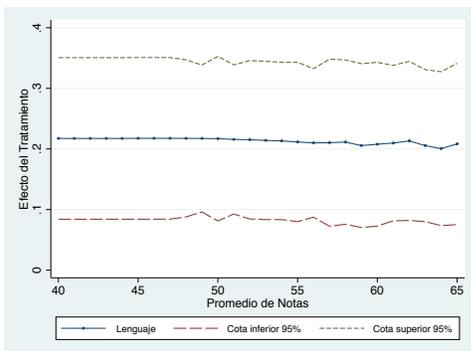
iii. Sensibilidad del efecto del tratamiento sobre los tratados

A continuación se muestran los resultados obtenidos para el efecto promedio del tratamiento sobre los tratados modificando el grupo de control utilizado.

1) Sensibilidad al filtro promedio de notas

En esta sección se muestra los resultados obtenidos al modificar el filtro del promedio de notas establecido con la finalidad de asemejar los procesos de admisión de los LSPE. Este promedio fue fijado en 5,5 así se testeará cual es el cambio obtenido en el ATT estimado si se utiliza un como filtro el rango de promedios 4,0 hasta 6,5 de esta forma es posible verificar si el promedio antes fijado influye y de qué forma en los resultados obtenidos. Los resultados obtenidos se muestran en las figuras I y II, en estas es posible observar que prácticamente hasta el promedio 6,0 el efecto promedio del tratamiento se mantiene estable (esta es más notorio en la cohorte 2006-2012) reduciéndose a medida que el promedio aumenta. A esto se suma que el promedio de los estudiantes tratados en la cohorte 2002-2008 es de 0,36 y para la cohorte 2006-2012 es de 6,3, esto quiere decir que la gran reducción en el número de estudiantes tratados existente desde el promedio 6,0 en adelante puede ser la razón de la variación en el ATT detectada, la cual es más notoria para la cohorte 2006-2012.

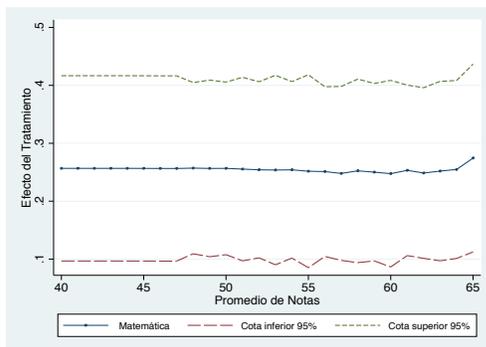
Figura I: Efecto del tratamiento para distintos promedios de corte, resultados para lenguaje



(a) cohorte 2002-2008

(b) cohorte 2006-2012

Figura II: Efecto del tratamiento para distintos promedios de corte, resultados para matemática



(a) cohorte 2002-2008

(b) cohorte 2006-2012

2) Selección de un segundo grupo de control

Para corroborar los resultados obtenidos, adicionalmente, se utilizó un segundo grupo de control, con la finalidad de dar una mayor robustez y validez interna a los resultados. Así para mejorar el overlap y refinar el grupo de control se construyó una nueva muestra a través de matching (Rubin, 2006; Ho et al., 2007), técnica recomendada cuando se busca estimar solo ATT (Imbens y Wooldridge, 2009).

i. Metodología

La metodología consiste en encontrar para cada unidad en el grupo de tratamiento al menos una unidad no tratada desde el total de potenciales controles que sea idéntica o por lo menos similar en todas las características observadas impuestas antes del tratamiento (Steiner y Cook, 2013), cabe mencionar que esta muestra se construye antes de utilizar las técnicas de análisis basadas en el propensity score. El método para encontrar la muestra mencionada combinó dos técnicas distintas: *matching exacto* y *matching multivariado*. Para evitar la existencia de controles muy distintos se utilizó matching con reemplazo, esto indica que un control ya utilizado puede servir de control para otra unidad de tratamiento.

El matching exacto consiste en encontrar los controles que coincidan exactamente en un set de variables predefinidas a las unidades de tratamiento, acá se utilizó las variables: género, asistencia a educación pre-escolar y si habían repetido algún grado antes de 4to básico. El matching multivariado utilizado consistió en estimar una distancia que cuantificara el grado de disimilitud entre pares de observaciones dadas ciertas variables, así el(los) contrafactuales son elegidos desde aquellos que muestran un menor grado de disimilitud. Para estimar dicha distancia la métrica Mahalanobis fue

seleccionada³⁹ (medida que está basada en el inverso de la matriz de covarianzas) y la estrategia de emparejamiento de observaciones fue el matching 1:M, la cual consiste en que cada unidad de tratamiento es emparejada a sus M vecinos más cercanos de acuerdo a la distancia estimada (Steiner y Cook, 2013). Con esta finalidad se utilizó una metodología que consta de dos partes, primero, se seleccionó un número grande de vecinos (100), para luego seleccionar los M vecinos más cercanos utilizando un punto de referencia (caliper) sobre la distancia estimada (Cochran y Rubin, 1973), el punto de referencia fue definido en términos de desviaciones estándar con respecto a las variables originales, permitiendo en este caso una diferencia estandarizada de las variables de hasta 0,25 para las variables consideradas. Este caliper fue elegido basado en Imbens y Wooldridge (2009, pp 24) quienes mencionan que los métodos de regresión serían sensibles a las especificaciones si la diferencia normalizada entre las variables de control y tratamiento excede 0,25, debido a que los métodos acá utilizados –descritos más adelante– se basan en ajustes a través de regresiones se utilizó este valor para definir los controles.

El modelo general para el matching multivariado consideró las siguientes variables: puntaje SIMCE de los sectores lenguaje y matemática, y un índice de nivel socioeconómico⁴⁰. A la muestra obtenida a través de este procedimiento se hará referencia como “grupo de control emparejado”.

ii. Resultados

Primero es necesario mencionar, que al igual que con el grupo de control antes descrito, el ajuste observado para este grupo mejora considerablemente la diferencia con respecto a la base completa, la principal diferencia es solo con respecto al número de observaciones de cada grupo.

En términos de resultados, lo primero que se puede observar es que la probabilidad de ser tratado y las variables más relevantes encontradas utilizando el primer grupo de control se mantienen al utilizar este segundo grupo de control más acotado.

Con respecto a los resultados del ATT, los resultados obtenidos en las Tablas IV y V, son bastante consistentes con lo obtenido al utilizar el nuevo grupo de control que cuenta solo con los ex compañeros emparejados, resultados mostrados en el Anexo 8. La diferencia más notable es que el efecto medido en lenguaje para la cohorte 2002-2008 aumentaría y variaría entre 0.24 y 0.25 desviación estándar, prácticamente equiparando el efecto obtenido en el sector de matemática para la

³⁹Dicha medida se define en el set de variables medidas en la línea base (antes del tratamiento), donde una medida de distancia igual a 0 implica que las dos unidades son idénticas en todas las variables consideradas, mientras que una distancia distinta de cero sugiere una diferencia en al menos una de las variables consideradas. La distancia entre distintas observaciones depende de las unidades en que estén medidas por lo que para evitar problemas con la escala utilizada para medir las distintas variables se estandarizaron todas las variables incluidas en el análisis.

⁴⁰ Variable construida a través de análisis factorial utilizando la educación del padre, de la madre y el ingreso del hogar antes del tratamiento.

misma cohorte. El resto de las estimaciones también muestran aumentos pero de menor consideración a lo ocurrido en el sector de lenguaje.

En general, es posible mencionar que el efecto del tratamiento sobre los tratados (ATT) –estimado a través de metodologías cuasi experimentales basadas en el propensity score–antes descrito son consistentes (entre los grupos de control utilizados) y relevan el hecho que los LSPE tendrían un efecto significativo en ambos sectores (lenguaje y matemática) y que dicho efecto sería de una magnitud considerable. Pero al parecer existirían grandes diferencias/variaciones entre las cohortes analizadas, especialmente en los resultados obtenidos para el sector de matemática, sin embargo dichas diferencias no mostraron ser estadísticamente significativas⁴¹. Una razón de esto puede deberse a que los estudiantes en el grupo de control han reducido la brecha con respecto a los estudiantes de los LSPE, en otras palabras, los colegios donde ingresaron los estudiantes similares a aquellos que entran a los LSPE muestran mejores resultados en la cohorte 2006-2012. En este sentido, cobra relevancia lo planteado por Illanes (2014), “en la medida en que el sistema educacional fuera de los colegios selectivos es de peor calidad, el efecto del colegio selectivo será más grande”.

Finalmente, es necesario mencionar que las estimaciones realizadas, además de depender del supuesto de selección en observables, están limitadas al área de soporte común. Así, si bien es cierto, debido a que las dos muestras analizadas llevan a resultados similares es posible afirmar que estas tendrán validez interna o, en otras palabras, dentro del área de soporte común. Por esto aún resta determinar si las estimaciones del ATT tienen la validez externa (fuera del área de soporte común).

b) Efecto promedio de la participación en los LSPE sobre los tratados, estimador Changes in Changes (CIC)

Tal como ya fue mencionado el estimador CIC corresponde a una generalización del estimador de diferencias en diferencias que permite, entre otras cosas, utilizar un mayor número de observaciones y, por lo tanto, no reducir las conclusiones al área de soporte común como sucede con las metodologías basadas en el propensity score, cosa que permitiría mostrar si las estimaciones tienen validez externa.

Los resultados, se muestran en la Tabla VII, en estos se ve que para el sector lenguaje el efecto sería consistente para ambas cohortes alcanzando entre 0,20 y 0,22 desviaciones estándar. Así, si bien es cierto que los resultados son similares a los ya obtenidos, estos serían más estables que los obtenidos por los métodos basados en el propensity score.

⁴¹ La formula para el test realizado corresponde al test planteado en Peternoster et al (1998), de esta forma el test utilizado fue $\frac{b_1 - b_2}{\sqrt{SE_{b_1}^2 - SE_{b_2}^2}}$

El sector de matemática muestra un efecto aún mayor que el detectado con los métodos basados en el propensity score para ambas cohortes, variando entre 0,21 y 0,31 desviaciones estándar, lo que correspondería a un efecto que variaría entre cerca de 10 y 15 puntos en la prueba SIMCE. En comparación con los resultados antes obtenidos, nuevamente se muestran diferencias entre las cohortes analizadas, y lo más importante se ve aumento de casi 0,1 desviación estándar entre las estimaciones, sin embargo, se mantiene la reducción en el efecto estimado en la cohorte 2006 respecto a la del año 2002. Pero ninguna de estas diferencias mostró ser significativa.

Tabla VII. Modelos CIC para estimación de ATT

	Panel SIMCE			
	2002-2008	2006-2012	2002-2008	2006-2012
Changes in Changes	Lenguaje		Matemática	
CIC sin controles	0.157 (0.104)	0.140** (0.070)	0.238* (0.129)	0.102 (0.110)
CIC con todos los controles	0.203** (0.096)	0.216*** (0.065)	0.309*** (0.110)	0.212** (0.092)

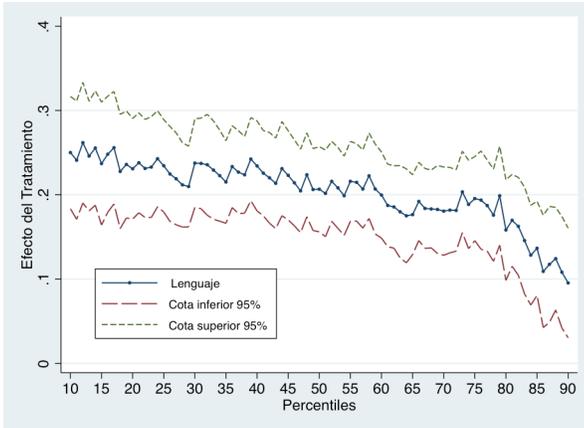
Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

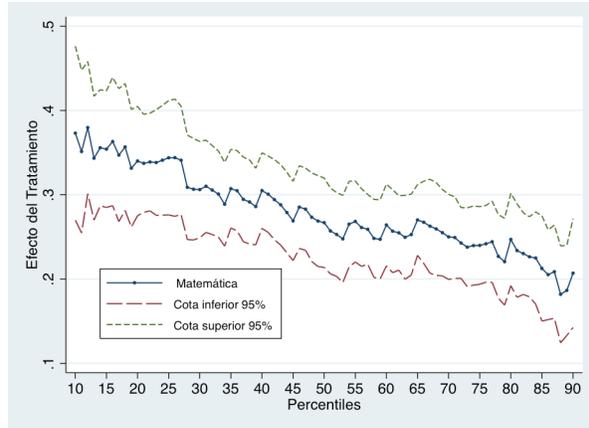
Como fue mencionado una de las características del modelo CIC es que permite estimar cuál es el efecto a lo largo de toda la *distribución de puntajes* esto permitirá determinar a qué alumnos impacta – en términos de puntaje– con mayor intensidad el tratamiento en consideración.

Las figuras III y IV muestran los resultados desagregados para cada centil, sector de aprendizaje y cohorte analizada, resultados obtenidos controlando por todas las variables explicativas. En general, los resultados muestran que todos los alumnos que ingresan a los LSPE se beneficiarían de asistir a estos establecimientos, donde prácticamente todos los efectos sería significativos. Adicionalmente, se observa que los alumnos pertenecientes a lugares más bajos de la distribución de puntajes (bajo el percentil 30) serían los más beneficiados donde dicho efecto muestra ir disminuyendo a medida que se avanza en la distribución de puntajes.

Figura III: Efecto del tratamiento para la cohorte 2002-2008 a lo largo de la distribución (entre centil 10 y 90) de puntajes SIMCE.

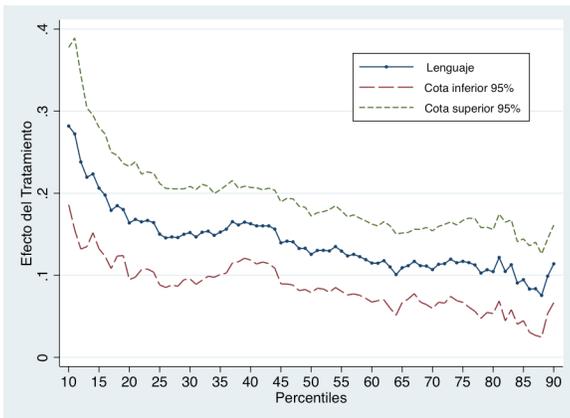


(a) resultados para lenguaje

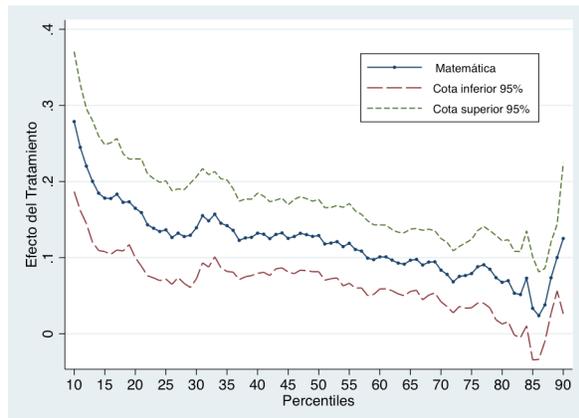


(b) resultados para matemática

Figura IV: Efecto del tratamiento para la cohorte 2006-2012 a lo largo de la distribución (entre centil 10 y 90) de puntajes SIMCE.



(a) resultados para lenguaje



(b) resultados para matemática

c) Efecto promedio de la participación en los LSPE sobre los tratados (ATT) diferenciado por nivel socioeconómico.

Las Tablas VIII y IX muestran los resultados del ATT diferenciado por nivel socioeconómico (NSE) este ejercicio permitirá mostrar si existe un efecto diferenciado para el tipo de alumno que ingresa a los LSPE o si el efecto antes detectado es relativamente constante.

Los resultados muestran que para lenguaje existiría una diferencia entre las cohortes, específicamente para los dos primeros quintiles siendo cerca de 0,42 desviaciones estándares el efecto para la cohorte 2006-2012, y solo 0,2 desviaciones estándares para la cohorte 2002-2008, no siendo significativo solo este último resultado. En el caso de matemática, el efecto para ambas cohortes muestra ser similar e ir en el mismo sentido, así para el grupo perteneciente a las familias del menor NSE existiría un efecto que variaría entre 0,31 y 0,37 desviaciones estándar que representa hasta cerca de 18 puntos adicionales en la prueba SIMCE.

Tabla VIII. Efecto promedio del tratamiento sobre los tratados diferenciado por NSE sector Lenguaje (ATT), modelos basados en propensity score

		Panel SIMCE		Panel SIMCE	
		2002-2008	2006-2012	2002-2008	2006-2012
		Efecto estimado (ATT)		Test de Validez	
ANCOVA ajustado					
	Quintil 1	0.20 (0.13)	0.42*** (0.09)	0.02 (0.07)	0.10 (0.08)
	Quintil 2	0.22*** (0.08)	0.16* (0.09)	0.06 (0.05)	0.08 (0.05)
	Quintil 3	0.28*** (0.07)	0.19*** (0.07)	0.00 (0.03)	0.03 (0.03)
	Quintil 4	0.19** (0.08)	0.14** (0.06)	-0.01 (0.03)	0.02 (0.03)
	Quintil 5	0.17** (0.07)	0.14** (0.06)	0.05* (0.03)	0.01 (0.03)
Wooldridge's double-robust					
	Quintil 1	0.21* (0.12)	0.38*** (0.10)	-0.03 (0.05)	0.03 (0.05)
	Quintil 2	0.24*** (0.07)	0.19** (0.09)	0.05 (0.03)	0.00 (0.03)
	Quintil 3	0.27*** (0.06)	0.18*** (0.07)	0.01 (0.02)	0.00 (0.02)
	Quintil 4	0.17*** (0.07)	0.13** (0.07)	0.01 (0.02)	0.04** (0.02)
	Quintil 5	0.16** (0.07)	0.17** (0.09)	0.00 (0.02)	0.02 (0.02)

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

Tabla IX. Efecto promedio del tratamiento sobre los tratados diferenciado por NSE sector Matemática (ATT), modelos basados en propensity score

	Panel SIMCE		Panel SIMCE	
	2002-2008	2006-2012	2002-2008	2006-2012
	Efecto estimado (ATT)		Test de Validez	
ANCOVA ajustado				
Quintil 1	0.36** (0.15)	0.32*** (0.12)	-0.02 (0.08)	0.08 (0.07)
Quintil 2	0.22** (0.11)	0.21** (0.10)	-0.02 (0.05)	0.02 (0.05)
Quintil 3	0.28*** (0.09)	0.12* (0.07)	0.04 (0.03)	-0.04 (0.04)
Quintil 4	0.25*** (0.09)	0.10 (0.07)	0.01 (0.04)	0.01 (0.03)
Quintil 5	0.21** (0.10)	0.05 (0.06)	0.05 (0.04)	0.01 (0.03)
Wooldridge's double-robust				
Quintil 1	0.37*** (0.13)	0.31*** (0.10)	0.02 (0.05)	0.02 (0.04)
Quintil 2	0.26*** (0.08)	0.21** (0.10)	0.02 (0.03)	0.00 (0.03)
Quintil 3	0.25*** (0.08)	0.13* (0.08)	0.01 (0.02)	0.01 (0.02)
Quintil 4	0.24*** (0.09)	0.07 (0.08)	0.01 (0.02)	-0.02 (0.02)
Quintil 5	0.20** (0.10)	0.04 (0.05)	0.04* (0.02)	0.00 (0.02)

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

En general, es posible observar que existiría un efecto diferenciado, el cual se concentraría principalmente en los alumnos de menor NSE – principalmente para el quintil 1 de la distribución– existiendo grandes diferencias entre el quintil 1 y 2 de la distribución tanto en lenguaje como en matemática para ambas cohortes analizadas.

Adicionalmente se observa que existiría un efecto asociado a la cohorte específica analizada, en este sentido una razón por la cual se observan las diferencias en los resultados obtenidos entre las cohortes es que los colegios que elegirían los estudiantes de control para la cohorte 2006-2012 han mejorado su desempeño a través del tiempo. En este sentido, esta evolución temporal explicaría que en la cohorte del año 2006 los efectos de participar en LSPE serían inexistentes para los estudiantes de los quintiles

3, 4 y 5 en el sector de matemática, mientras que su efecto también se habría reducido en los mismos grupos en lectura. Sin embargo, es necesario mencionar que ninguna de las diferencias entre las cohortes mostró ser significativas.

En resumen, estos resultados dan cuenta que el ingreso a este tipo de colegios podría ser una fuente de movilidad individual en términos de resultados⁴², representando una posibilidad cierta de mejora para los alumnos de menor NSE con las mejores habilidades del sistema escolar (resilientes). Posibilidad que muy probablemente no se encuentre disponible en el resto de los colegios a los cuales suelen –o pueden – acceder dadas sus condiciones sociales. Así, es necesario hacer énfasis que la mayor proporción de alumnos que ingresa a este tipo de colegios no pertenecería a familias de menor NSE (solo cerca de un 6% de la matrícula), sino que la mayor proporción de alumnos pertenecería al NSE medio o medio-alto.

d) ¿A qué se deben las diferencias observadas a favor de los LSPE?

En esta sección se trata de dar respuesta a una de las hipótesis centrales de esta investigación que se puede resumir en la pregunta ¿a qué se deben las diferencias observadas a favor de los LSPE?, pregunta para la cual ya se esbozó una potencial causa al observar el efecto diferenciado entre las cohortes. Para esto, primero se introduce el efecto del tratamiento sobre variables intermedias, y luego se identifica una de las posibles razones que explicaría las diferencias observadas a favor de los LSPE antes estimadas.

- i. Determinación de la efectividad de los procesos de selección, el efecto del tratamiento sobre variables intermedias

Una variable de resultado intermedia, tal como fue mencionado, se puede definir como una variable afectada por el tratamiento pero que no corresponde al objetivo real o final de la intervención o tratamiento realizado⁴³, usualmente incorporan características propias como también del tratamiento que se busca evaluar (Frangakis y Rubin, 2002). Controlar por este tipo de variables suele ser importante e incluso puede ser considerado un objetivo de la investigación ya que suelen contener información relevante sobre la muestra analizada (por ejemplo, el no cumplimiento en experimentos aleatorios), pero la incorporación de estas en la estimación de efectos causales –como, por ejemplo, a

⁴² Recordar que Chile se caracterizaría por una alta persistencia de los resultados a través de la trayectoria educacional de los alumnos (Valenzuela, Allende, Sevilla y Egaña, 2014).

⁴³ Un ejemplo de este tipo de variables es el no cumplimiento en experimentos aleatorios, que suele ser muy importante incorporar como control para estimar efectos causales. En dicho caso la variable no cumplimiento puede englobar características de eficacia del programa como características propias de los individuos (Frangakis y Rubin, 2002), que son interesantes de ser analizadas.

través de métodos basados en el propensity score– puede generar estimadores afectados por el llamado *sesgo de selección post-tratamiento* (Rosenbaum, 1984; Frangakis y Rubin, 2002).

Dada la definición del tratamiento, en el contexto de los LSPE es posible identificar un set de variables que pueden ser calificadas como variables intermedias. Específicamente, se hace referencia a los efectos pares estimados para el año 2008 (efecto par de la habilidad individual⁴⁴, efecto par de la educación de la madre y padre, efecto par del ingreso del hogar). La identificación de estas variables como intermedias se debe a que la selección de alumnos por parte de los LSPE hace que se agrupen padres y alumnos con determinadas características en dichos establecimientos pero esto no corresponde al principal objetivo del tratamiento⁴⁵, así estas variables estarían mostrando solamente la eficacia de los procesos de selección de dichos establecimientos. Por ende, dicha agrupación correspondería solo a un primer paso para conseguir el objetivo final del tratamiento, la excelencia académica.

Tal como fue mencionado, esta metodología permitirá determinar la efectividad de los procesos de selección llevados a cabo en estos establecimientos, al identificar el *efecto del tratamiento sobre variables intermedias*. Más formalmente, se definirá el efecto del tratamiento sobre variables intermedias, como el efecto asociado al tratamiento que se diferencia de las características propias de una variable afectada por el tratamiento (variable intermedia) pero que no corresponde al objetivo o resultado final de un tratamiento o proceso de selección realizado.

La estimación del efecto del tratamiento sobre variables intermedias se puede realizar con mediciones de corte transversal (por ejemplo, con metodologías tipo propensity score) aplicable en estudios donde no existen mediciones antes y después del tratamiento de dichas variables (como es el caso del no cumplimiento en experimentos aleatorios). Por otro lado, la robustez de las estimaciones aumenta considerablemente si existen mediciones antes y después de aplicado el tratamiento para dichas variables, datos con los cuales es posible aplicar metodologías de tipo diferencias en diferencias (DID) para estimar el efecto. En este caso, fue posible contar con información antes y después del tratamiento para las variables intermedias identificadas y se utilizó la metodología CIC –que se adapta de mejor

⁴⁴Recordar que la llamada habilidad estimada, corresponde a un proxy estimado a través de un residuo de una función de producción escolar. De esta forma, la acá llamada habilidad puede estar englobando muchos no observables tal como la motivación de los alumnos, la perseverancia, y muchos otros no observables asociados con habilidades cognitivas y no cognitivas.

⁴⁵Si así fuese, el objetivo de los colegios debería cambiarse desde la obtención de excelencia académica a uno similar a “generar colegios donde se reúnan los mejores alumnos y los padres con las mejores características observables”.

forma a las condiciones del tratamiento– permitiendo identificar el sentido y la magnitud del efecto de la selección sobre estas variables⁴⁶.

Las Tablas X y XI muestran los resultados para ambas cohortes, en estas se observa que las cuatro variables analizadas (efecto par de la habilidad individual⁴⁷, efecto par de la educación de la madre y padre, y efecto par del ingreso del hogar) muestran tener un efecto significativo en ambas cohortes mostrando una consistencia con respecto a las magnitudes estimadas. Así, el efecto par de la habilidad muestra que tendría entre 0,5-0,6 desviaciones estándar de diferencia entre el grupo tratado y de control; de la misma forma el efecto par de la educación de los padres muestra que para el grupo tratado en promedio los padres tendrían entre 1,6 y 1,8 años adicionales de educación; finalmente, el efecto par del ingreso también muestra ser significativamente mayor en los colegios tratados (entre 0,24 y 0,28 desviaciones estándares).

Tabla X. Modelos CIC para estimación del efecto del tratamiento sobre variables intermedias. Panel 2002 – 2008

	Panel SIMCE 2002-2008			
Changes in Changes	ep habilidad	ep educ madre	ep educ padre	ep ingreso
CIC con todos los controles	0.536*** (0.048)	1.781*** (0.205)	1.809*** (0.226)	0.241*** (0.058)

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

Tabla XI. Modelos CIC para estimación del efecto del tratamiento sobre variables intermedias. Panel 2006 – 2012

	Panel SIMCE 2006-2012			
Changes in Changes	ep habilidad	ep educ madre	ep educ padre	ep ingreso
CIC con todos los controles	0.617*** (0.065)	1.643*** (0.237)	1.713*** (0.264)	0.283*** (0.076)

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

Los resultados obtenidos demuestran que el tratamiento analizado sería efectivo, en el sentido que los procesos de selección realizados realmente aseguran el ingreso de alumnos con mejores habilidades no observadas y que a la vez se asociaría a padres con mejor educación e ingresos. En este sentido, el

⁴⁶ Las variables incorporadas en el modelo CIC acá utilizado fueron: Número de libros en el hogar, educación del padre y madre, ingreso del hogar estandarizado, sexo, si repitió antes de 4º básico, si asistió a educación pre escolar, dependencia institucional del establecimiento y condición de ruralidad del establecimiento.

⁴⁷ Se presentan solo los resultados del efecto par de la habilidad estimada a través del puntaje SIMCE del sector de lenguaje en el año base, esto se debe a que los resultados para ambos sectores son similares, estos resultados están disponibles a quien los solicite.

número y efecto de aquellos que ingresan a través de “pituto” (Bucarey et al, 2014) a este tipo de establecimientos se puede asegurar que no sería considerable.

Cabe destacar que, si bien es cierto, este tipo de variables no determinarían completamente los resultados educacionales, sí tienen asociado un efecto protector en el sentido que ayudarían a permanecer en los mejores niveles de rendimiento a través de las trayectorias educacionales (Valenzuela, Allende, Sevilla y Egaña, 2014).

- ii. Efecto promedio de la participación en los LSPE sobre los tratados, ajustado por variables intermedias medidas antes y después del tratamiento.

La metodología de Diferencias en Diferencias (DID) y por ende su generalización conocida como Changes in Changes (CIC), requieren que las variables utilizadas en la estimación del efecto del tratamiento hayan sido medidas antes y después del tratamiento. En este caso, existen cuatro variables que se pueden clasificar como intermedias, las cuales ya se demostró que estarían afectadas por el tratamiento, pero corresponderían a variables que proporcionarían información relevante para responder la pregunta, ¿a qué se deben las diferencias observadas a favor de los LSPE?, y por ende a una de las hipótesis de esta investigación: los resultados obtenidos por estos colegios se derivan de procesos internos de los establecimientos o dicho efecto se debe principalmente al tipo de alumnos que ingresa a estos establecimientos.

En este caso, las variables intermedias relevantes identificadas – específicamente el efecto par de la educación de la madre y del padre; efecto par del ingreso del hogar; y efecto par de la habilidad no observada– permiten determinar si las características familiares o específicas de los alumnos (efecto par de la habilidad) son la razón detrás del ATT estimado, y por ende si dichos resultados se asociarían principalmente al proceso de selección. En caso contrario, los resultados indicarían que existe “algo” al interior de estos establecimientos, posiblemente algún proceso interno como sus metodologías de enseñanzas, gestión u organización que estarían generando las mejoras observadas en el rendimiento.

Así estas variables fueron incorporadas al modelo CIC estimado –donde se utilizó el puntaje SIMCE estandarizado de lenguaje y matemática como variable dependiente– que incluyó adicionalmente los controles: Número de libros en el hogar, educación del padre y madre, ingreso del hogar estandarizado, sexo, si repitió antes de 4° básico, si asistió a educación pre escolar, dependencia institucional del establecimiento y condición de ruralidad del establecimiento.

Como las variables intermedias consideradas fueron medidas antes y después del tratamiento es posible incluirlas como controles, además es necesario incluirlas debido a que el tratamiento actúa mediante un

efecto en la selección de los mejores alumnos haciéndose necesario controlar por este tipo de variables para explicar el beneficio observado, o podrían existir efectos complementarios. Así los resultados de la estimación del efecto del tratamiento sobre los tratados en el contexto de diferencias en diferencias (CIC), se muestran en la Tabla XII.

Tabla XII. Modelos CIC para estimación de ATT incluyendo variables intermedias

	Panel SIMCE			
	2002-2008	2006-2012	2002-2008	2006-2012
Changes in Changes	Lenguaje		Matemática	
CIC con todos los controles + efecto par familia ^{††}	0.063 (0.081)	0.115** (0.058)	0.155 (0.100)	0.07 (0.076)
CIC con todos los controles + efecto par habilidad	0.007 (0.076)	0.026 (0.060)	0.021 (0.095)	-0.058 (0.086)

Errores estándar en paréntesis

^{††}Se incluye como controles el efecto par de la educación de la madre y padre, y el efecto par del ingreso del hogar

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

Los resultados muestran que al incluir los controles asociados al tipo de familias que ingresan a los LSPE, las estimaciones del efecto del tratamiento no serían del todo consistentes entre los sectores y cohortes analizadas, detectando un efecto positivo del tratamiento para lenguaje que variaría entre 0,06 y 0,12 desviaciones estándar siendo significativo solo en el panel 2006-2012 y para matemática variaría entre 0,07 y 0,16 desviaciones estándares (solo significativo en el panel 2002-2008). Estos resultados indicarían que las características de las familias por sí solas no explicarían la totalidad de los resultados obtenidos por los establecimientos.

Por el contrario, al incluir el efecto par del proxy de la habilidad no observada, los resultados solo relevan una conclusión: el efecto promedio del tratamiento sobre los tratados estimado en la sección anterior que variaba para lenguaje entre 0,17 y 0,22 desviaciones estándar y entre 0,23 - 0,31 para matemática (este corresponde solo a los resultados de la cohorte 2002-2008) estaría explicado principalmente por el efecto par de la habilidad de los estudiantes seleccionados para ingresar a los LSPE.

VI. Conclusiones

Esta investigación ha tratado de contribuir a la investigación nacional, muy escasa, sobre los efectos asociados a los colegios selectivos, enfocándose específicamente en los Colegios Selectivos Públicos de Excelencia. Si bien es cierto, los datos y la metodología utilizadas en esta investigación son cuasi experimentales y, como fue mencionado, corresponderían a un *segundo mejor* –debido al diseño del tratamiento–, los resultados obtenidos mostraron una consistencia tanto interna como externa para ambas cohortes, y se encontrarían aproximadamente entre 0,20 y 0,30 desviaciones estándar –que equivale a aproximadamente a un aumento entre 10 y 15 puntos en las pruebas estandarizadas– encontrándose algunas variaciones entre las cohortes analizadas.

Una de las razones de la reducción en el tiempo del desempeño entre los estudiantes de los LSPE es que los estudiantes similares a éstos (controles) tenderían a elegir colegios de mejor desempeño también, así como también porque los colegios no emblemáticos están mejorando su desempeño promedio a través del tiempo. Sin embargo, es necesario mencionar que las diferencias observadas a través del tiempo no mostraron ser estadísticamente significativas.

En general, las metodologías utilizadas muestran una gran consistencia en términos de magnitud y sentido con el efecto encontrado para colegios selectivos en países en vías de desarrollo (Pop-Eleches y Urquiola, 2013; Dustan, 2010; Jackson, 2010) y también para el caso chileno (Bucarey et al, 2014). Resultados que relevan la importancia potencial que el efecto par puede tener en las trayectorias educativas de los alumnos en un país como Chile.

Adicionalmente, los resultados dan cuenta que el ingreso a este tipo de colegios podría ser una fuente de movilidad en término de resultados⁴⁸, representando una posibilidad cierta de mejora para los alumnos de menor NSE con las mejores habilidades del sistema escolar (resilientes). Posibilidad que muy probablemente no se encuentre disponible en el resto de los colegios a los cuales suelen –o pueden – acceder dadas sus condiciones sociales.

En este sentido, es necesario destacar que los alumnos que asisten a los LSPE corresponden a estudiantes con capacidad, habilidad y otras características no observables (por ejemplo, la motivación) superiores al promedio del país –incluso pudiendo calificar, en algunos casos, dentro de los mejores

⁴⁸ Recordar que Chile se caracterizaría por una alta persistencia de los resultados a través de la trayectoria educacional de los alumnos en Chile (Valenzuela, Allende, Sevilla y Egaña, 2014).

estudiantes del país—. Pertenecientes en su mayoría a familias de nivel socioeconómico medio o medio-alto, provenientes principalmente de colegios particulares (tanto subvencionados como pagados). Por esta razón, si bien es cierto, los resultados revelan que este tipo de colegios pueden ser una fuente de movilidad escolar –como social– para estudiantes de menor nivel socioeconómico, esto afecta solo a un pequeño porcentaje de su matrícula, pues solo un 6% de los alumnos en ambas cohortes pertenece al quintil menor de la distribución por nivel socioeconómico.

A diferencia de las investigaciones tanto nacionales como internacionales, la riqueza de los datos existentes en Chile utilizados en esta investigación, permite responder parcialmente a una de las preguntas centrales de esta investigación ¿cuál es la razón de los mejores resultados observados en los LSPE? Al analizar dos cohortes de estos colegios parece ser que el efecto del descreme o selección de los mejores estudiantes (mayores capacidades, habilidad, motivación, etc.) sería lo que explica la diferencia observada a favor de este tipo de colegios; así, una vez controlado por el efecto par de la habilidad de los estudiantes, el efecto encontrado se hace prácticamente cero y no significativo. Sin embargo, es importante mencionar que la habilidad académica no observada estimada en este estudio, al estar siendo medida a través del residuo de una función de producción escolar, puede estar englobando muchos no observables, y no solo la habilidad académica, tal como puede ser la motivación de los alumnos, la perseverancia, y muchos otros no observables asociados con habilidades cognitivas y no cognitivas que pueden afectar el futuro académico de los estudiantes tanto en el corto como en el largo plazo, de esta forma saber qué diferencia a estos alumnos del resto en el sistema escolar es una pregunta que esta investigación no puede responder y queda abierta para futuras investigaciones.

Adicionalmente, es necesario mencionar que el resultado encontrado no quiere decir que los colegios realmente no agreguen valor a través de sus procesos internos como son sus metodologías de enseñanza, gestión u organización desarrolladas, por el contrario, muestra que el tipo de alumnos que ingresa a estos establecimientos en promedio *facilitarían y amplificarían* la efectividad de estas escuelas, o en otras palabras, el efecto de las metodologías de enseñanza y procesos internos sobre el resultado de los alumnos –y por ende del establecimiento– pareciera ser facilitado por la selección realizada.

Finalmente, al igual que toda la evidencia a nivel internacional, es necesario reconocer que esta investigación no es capaz de abordar efectos que van más allá de los resultados de corto plazo –observados en pruebas estandarizadas u otras mediciones tradicionales de efectividad escolar– y que incluso en el largo plazo pueden no ser medibles o cuantificables de forma correcta tal como la

satisfacción personal o salud, que pueden ser derivados de los potenciales efectos asociados al nuevo ambiente del cual los alumnos son parte (Finn y Hockett, 2012a). Factores que podrían relevar aún más la importancia del efecto par, que podría estar asociado con el ingreso a este tipo de colegios.

Para terminar, los resultados obtenidos muestran que los liceos definidos como selectivos públicos de excelencia (muy similares a los llamados liceos emblemáticos), tendrían un desafío no mayor para mantener su efectividad si fuesen prohibidos definitivamente los procesos de selección. Esto, debido a que es muy probable que sus procesos internos, tanto en gestión como en metodologías de enseñanza, estén optimizados para funcionar, y ser efectivos, con un determinado tipo de estudiantes –aquellos a los que en este trabajo se ha hecho referencia como con mayor habilidad no observada– y padres –que muy probablemente correspondan a padres más interesados o involucrados en la educación de sus hijos–. Por esta razón, es necesario tener en consideración qué es o qué se quiere mantener, potenciar e incluso preservar con este tipo de colegios, ¿se busca dar una opción real de movilidad escolar a los alumnos de menor NSE en el sistema escolar?, o ¿solo a aquellos con mayores habilidades?, ¿se busca preservar un tema histórico –republicano– asociado a este tipo de colegios?, o ¿se busca mejorar la educación municipal en su conjunto a través de la generación de este tipo de colegios?

Finalmente, un punto importante planteado por Finn y Hockett (2012a, 2012b) es ¿qué hacer con los niños y jóvenes con más talento académico? ¿pueden las escuelas selectivas hacerse cargo de la educación de aquellos alumnos con mayor potencial o talento académico? En este sentido, es necesario tener en consideración que agrupar a los mejores estudiantes en pocos colegios (tal como promueven hacer los Liceos Emblemáticos y los Liceos Bicentenario de Excelencia), si bien es cierto, se ha mostrado que produciría mejoras en términos académicos para este subconjunto de estudiantes, ello también implica profundizar aún más la segregación de un sistema escolar que ya es altamente segregado (Valenzuela, Villalobos y Gómez, 2013). Así, en este sentido, pareciera mejor promover la existencia de programas complementarios específicos enfocados a los alumnos al interior de cada establecimiento con mayor talento académico, tales como PENTA-UC, Delta-UCN, BETA de la PUCV, Talentos de la Universidad de Concepción, PROENTA de la Universidad de la Frontera y ALTA de la Universidad Austral de Valdivia, los cuales se orientan a fortalecer las capacidades de los estudiantes vulnerables en escuelas municipales, con actividades complementarias sin que éstos deban retirarse de sus colegios donde han estudiado históricamente (López et al., 2013), sin promover un estándar de exclusión en el sistema educacional, sino más bien uno de integración con oportunidades de diferenciación para aquellos más talentosos.

VII. Bibliografía

- Abadie, Alberto (2005). "Semiparametric Difference-in-Differences Estimators". *Review of Economic Studies* 72, 1–19
- Abadie, Alberto y Guido, Imbens (2008). "On the Failure of the Bootstrap for Matching Estimators." *Econometrica*, 76(6): 1537-57.
- Abdulkadiroglu, Atila; Angrist, Joshua; y Pathak, Parag (2014). "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools", *Econometrica*, Vol. 82, No 1. Enero 2014, 137-196.
- Alcalde, Pilar. (2013). "A Behavioral Model of Non-Linear Peer Effects in Cognitive Achievement". Document de trabajo.
- Altonji, Joseph.; Huang, Ching-I y Taber, Christopher (2014). "Estimating the cream skimming effect of school choice". Forthcoming, *Journal of Political Economy*. NBER Working Paper 16.579, Diciembre 2010.
- Angrist, Joshua (2014). "The perils of peer effects". *Labour Economics*. Volume 30. October 2014, Pages 98–108
- Athey, Susan y Imbens, Guido (2002). "Identification and Inference in Nonlinear Difference-In-Differences Models," NBER Technical Working Paper No. t0280.
- Athey, Susan y Imbens, Guido (2006). "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica*, 74(2): 431–97.
- Ashenfelter, Orley y Card, David (1985). "Using the Longitudinal Structure of Earnings to Estimate the Effects of Training Programs", *Review of Economics and Statistics*, vol. 67, 648-660.
- Ballou, Dale. y Liu, Keke (2009). "Magnet schools and peers: effects on student achievement". Borrador.
- Bellei, Cristian (2007). "Expansión de la educación privada y mejoramiento de la educación en Chile. Evaluación a partir de la evidencia". *Revista Pensamiento Educativo*, Vol. 40, No 1.
- Becker, Sascha y Ichino, Andrea (2002). "Estimation of average treatment effects based on propensity scores", *The Stata Journal* Vol.2, No.4, pp. 358-377.
- Bifulco, Robert; Ladd, Helen y Ross, Stephen (2009). "The Effects of Public School Choice on Those Left Behind: Evidence from Durham, North Carolina". *Peabody Journal of Education* 01/2009; 84(2):130-149.
- Bryson, Alex; Dorsett, Richard y Purdon, Sussan (2002). The use of propensity score matching in the evaluation of labour market policies. Working Paper No. 4, Department for Work and Pensions.
- Bui, Sa; Imberman, Scotty y Craig, Steven (2012). "Poor Results for High Achievers: Hard Evidence on the Impact of Gifted and Talented Programs," *Education Next*, 12(1): 70–76.
- Bucarey, A., Jorquera, M., Muñoz, P., and Urzúa, S. (2014). "El efecto del Instituto Nacional: Evidencia a partir de un diseño de regresión discontinua". *Estudios Públicos*, 133: 37-68.

- Burga, Cybele (2003). "Re-evaluando PROJoven: Propensity Score Matching y una evaluación paramétrica". CEDEP.
- Caliendo, Marco y Kopeinig, Sabine (2008). "Some Practical Guidance for the Implementation of Propensity Score Matching", *Journal of Economic Surveys*, 22(1), 31-72.
- Carrasco, Alejandro; Bogolasky, Francisca; Flores, Carolina; Gutiérrez, Gabriel; y San Martín, Ernesto (2014). "Selección de estudiantes y desigualdad educacional en Chile: ¿Qué tan coactiva es la regulación que la prohíbe?". Informe resumen proyecto FONIDE n° 711286: CEPPE, Pontificia Universidad Católica de Chile.
- Carrasco, Alejandro; Gutiérrez, Gabriel; Bogolasky, Francisca; Rivero, Rosario; y Zarhi, Magdalena. (2014b) "Informe Final. Análisis del Estado de Implementación del Programa: Liceos Bicentenario de Excelencia". Centro de Políticas Comparadas en Educación. Pontificia Universidad Católica de Chile.
- CEPPE (2010). "Liceos de Excelencia: Ideas y experiencias relevantes para su diseño e implementación", *Notas de Educación* No 3, junio 2010.
- Chay, Kenneth; McEwan, Patrick y Urquiola, Miguel (2003). "The central role of noise in evaluating interventions that use test scores to rank cholos". NBER, Working Paper 10,118.
- Clark, Damon (2008). "Selective Schools and Academic Achievement," Discussion Paper 3182, IZA.
- Clark, Damon (2010). "Selective Schools and Academic Achievement," *The B.E. Journal of Economic Analysis & Policy*: Vol. 10: Iss. 1 (Advances), Article 9.
- Coe, Robert; Jones, Karen; Searle, Jeff; Kokotsaki, Dimitra; Kosnin, Azlina y Skinner, Paul (2008). "Evidence on the effects of selective educational systems", CEM Centre, Durham University, a report for the Sutton Trust.
- Cochran, William G., and Donald B. Rubin. 1973. "Controlling Bias in Observational Studies: A Review." *Sankhya*, 35(4): 417-46.
- Cullen, Julie. y Jacob, Brian (2007). "Is gaining Access to selective elementary schools gaining ground?: Evidence from randomized lotteries". NBER Working Paper 13.443.
- Cullen, Julie. y Jacob, Brian (2009). "Is gaining access to a selective elementary school gaining ground? Evidence from randomized lotteries" en J. Gruber, ed., *An Economic Perspective on the Problems of Disadvantaged Youth* (Chicago, IL: University of Chicago Press).
- Cullen, Julie; Jacob, Brian y Levitt, Steven (2005). "The impact of school choice on student outcomes: an analysis of the Chicago public schools," *Journal of Public Economics*, 89/5-6), 729-760.
- Cullen, Julie; Jacob, Brian y Levitt, Steven (2006). "The effect of school choice on participants: evidence from randomized lotteries," *Econometrica* 74(5): 1191-1230.
- Dale, Stacy y Krueger, Alan (2011). "Estimating the return to college selectivity over the career using administrative earnings data". NBERG Working Paper 17159. <http://www.nber.org/papers/w17159>

- Dale, Stacy y Krueger, Alan (2002). "Estimating the payoff to attending a more selective college: And application of selection on observables and unobservables". *The Quarterly Journal of Economics* 117(4):1491-1527.
- Dehejia, Rajeev y Wahba, Sadek (1999). "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs". *Journal of the American Statistical Association*, 94(448): 1053–62.
- Dehejia, Rajeev y Wahba, Sadek (2002). "Propensity Score-Matching Methods for Nonexperimental causal studies". *The Review of Economics and Statistics*, February 2002, 84(1): 151–161.
- Dobbie, Will y Fryer, Roland (2011). "Exam high schools and academic achievement: evidence from New York City". NBER Working Paper 17286.
- Dobbie, Will y Fryer, Roland (2014). "The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools". *American Economic Journal: Applied Economics* 2014, 6(3): 58–75. <http://dx.doi.org/10.1257/app.6.3.58>
- Duflo, Esther; Dupas, Pascaline. y Kremer, Michael (2010). "Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya", *American Economic Review*, 101(5):1739-74.
- Dupriez, Vincent (2010). "Methods of grouping learners at school", UNESCO.
- Dupriez, Vincent; Dumay, Xavier y Vause, Anne (2008). "How do school systems manage pupils' heterogeneity?. *Comparative Education Review*, 52(2): 245-274.
- Dustan, Andrew (2010). "Have Elite Schools Earned their Reputation?: High School Quality and Student Tracking in Mexico City". Borrador.
- Emsley, Richard; Lunt, Mark; Pickles, Andrew y Dunn, Graham (2008). "Implementing double-robust estimators of causal effects," *Stata Journal*, Vol. 8(3), pages 334-353.
- Epple, Dennis. y Romano, Richard (1998). "Competition between private and public schools, vouchers, and peer-group effects", *American Economic Review*, 88(1): 33-62.
- Epple, Dennis. y Romano, Richard (2002). "Educational vouchers and cream skimming", NBER Working Paper No 9354.
- Finn, Chester y Hockett, Jessica (2012a). "Exam Schools: Inside America's Most Selective Public High Schools" (Princeton University Press),
- Finn, Chester y Hockett, Jessica (2012b). "Exam Schools from the Inside: Racially diverse, subject to collective bargaining, fulfilling a need" *EducationNext Otoño 2012* , Vol. 12, N° 4
- Frangakis, Constantine y Rubin, Donald (2002). "Principal Stratification in Causal Inference" *Biometrics*, 58, 20–29.
- Galindo-Rueda, Fernando y Vignoles, Anna (2004). "The Heterogeneous Effect of Selection in Secondary schools: Understanding the Changing Role of Ability". IZA Discussion Paper Series, N°1245, Agosto 2004. <http://opus.zbw-kiel.de/voltexte/2005/2434/pdf/dp1245.pdf>
- Godoy, Felipe; Salazar, Felipe y Treviño, Ernesto (2014). "Prácticas de selección en el sistema escolar

- chileno: requisitos de postulación y vacíos legales” en Informes para la Política Educativa, N°1, Mayo 2014. Centro de políticas comparadas de educación. Universidad Diego Portales.
- Hanushek, Eric y Woessmann, Ludger (2010). “The economics of international differences in educational achievement”, NBER, working paper 15.949
- Heckman, James (1979). “Sample selection bias as a specification error”. *Econometrica*, 47, pp. 153 - 161.
- Heckman, James; LaLonde, Robert y Smith, Jeffrey (1998). “The Economics and Econometrics of Training Programs” in *Handbook of Labor Economics*, Volume III, ed. by O. Ashenfelter and D. Card. Amsterdam: North Holland.
- Heckman, James y Smith, Jeffrey (1995). “Assessing the Case for Social Experiments”, *Journal of Economic Perspectives*, vol. 9 (2), 85-100.
- Ho, Daniel; Imai, Kosuke; King, Gary y Stuart, Elizabeth (2007). “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference”. *Political Analysis* (2007) 15:199–236.
- Hoekstra, Mark (2009). “The effect of attending the flagship state university on earnings: A discontinuity-based approach”. *The Review of Economics and Statistics*, November 2009, 91(4): 717–724.
- Hox, Joop (2002). “Multilevel Analysis. Techniques and Applications”. Mahwah, NJ: Lawrence Erlbaum Associates.
- Illanes, Gastón (2014). “Efectos pares: ¿Qué sabemos realmente?”. Puntos de referencia 377, Centro de estudios públicos. www.cepchile.cl
- Imbens, Guido (2004). “Nonparametric estimation of average treatment effects under exogeneity: A review”. *Review of Economics and Statistics* 86 4–29.
- Imbens, Guido, y Lemieux, Thomas (2008). “Regression discontinuity designs: A guide to practice”. *Journal of Econometrics*, doi:10.1016/j.jeconom.2007.05.001
- Imbens, Guido y Wooldridge, Jeffrey (2009). “Recent Developments in the Econometrics of Program Evaluation”. *Journal of Economic Literature*, 47(1): 5–86.
- Jackson, C. Kirabo. (2010). “Do Students Benefit from Attending Better Schools? Evidence from Rule- Based Student Assignments in Trinidad and Tobago.” *Economic Journal* 120 (549): 1399–1429.
- Kane, Thomas; Staiger, Douglas y Geppert, Jeffrey (2002). “Randomly Accountable: Test Scores and Volatility.” *Education Next*, Spring 2002, 2(1), pp. 56-61.
- Kang, Joseph y Schafer, Joseph (2007). “Demystifying double robustness: A comparison of alternative strategies for estimating population means from incomplete data”. *Statistical Science*, 26, 523–539.
- Kalogridis, Demetra y Loeb, Susanna (2013). “Different Teachers, Different Peers- The Magnitude of Student Sorting Within Schools”. *Educational Researcher* vol. 42 no. 6304-316. August/September.

La Tercera (Junio, 05 del 2010).

Lara, Bernardo; Mizala, Alejandra y Repetto, Andrea (2011). "The Effectiveness of Private Voucher Education: Evidence from Structural School Switches". *Educational Evaluation and Policy Analysis*, Vol.33(2), 2011, 119-137.

Lee, Lung-Fei (1983). "Generalized econometric models with selectivity". *Econometrica*, Vol 51, pp. 507 - 512.

López, Verónica; Conejeros, Leonor; García, María; Gudenschwager, Helga y Proestakis, Alejandro (2013). "Gifted Education in Chile amidst Public Debate on Excellence without Equity in Education." In P. Sanchez-Escobedo (ed.) *Talent development around the world. A global perspective on gifted education* (pp.167-199). Germany: Lambert Academic Publishing.

Lunceford, Jared y Davidian, Marie (2004). "Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study". *Statistics in Medicine*. **23** 2937–2960.

Marsh, Herbert; Kuyper, Hans; Morin, Alexandre; Parker, Phillip y Seaton, Marjorie (2014). "Big-fish-little-pond social comparison and local dominance effects: Integrating new statistical models, methodology, design, theory and substantive implications". *Learning and Instruction* 33, 50-66

MINEDUC (2010). "Liceos Bicentenario: Una Promesa de Excelencia". *Revista de Educación*, Edición 344, Ministerio de Educación, Santiago, Chile.

MINEDUC (2013). "Estándares de aprendizaje".

Mizala, Alejandra y Torche, Florencia (2012). "Bringing the schools back in: the stratification of educational achievement in the Chilean voucher system". *International Journal of Educational Development* 32, 2012: 132-144.

Nechyba, Thomas (2003). "Introducing school choice into multidistrict public school system". *American Economic Review* 90(1): 130-146.

Neild, Ruth (2002). "The effects of magnet schools on neighborhood high schools; an examination of achievement among freshmen". Borrador, University of Pennsylvania.

Neild, Ruth y Balfanz, Robert (2001). "An extreme degree of difficulty; the educational demographic of the ninth grade in an urban school system", paper presented at the annual meeting of American Sociological Association, Anaheim, CA.

OECD (2010), "PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes (Volume II)". <http://dx.doi.org/10.1787/9789264091504-en>

Paternoster, Raymond; Brame, Robert; Mazerolle, Paul; y Piquero, Alex (1998). "Using the correct statistical test for the equality of regression coefficients". *Criminology*. Vol. 36 (4): 859-866.

Pischke, Jorn y Manning, Alan (2006). "Comprehensive versus selective schooling in England and Wales: what do we know?", NBER Working Paper No 12176.

- Planty, Michael; Provasnik, Stephen. y Daniel, Bruce (2007). "High school coursetaking: findings from the Condition of Education 2007", U.S. Department of Education. Washington, D.C., National Center for Education Statistics. NCES 2007-065.
- Pop-Eleches, Cristian y Urquiola, Miguel (2013). "Going to a Better School: Effects and Behavioral Responses". *American Economic Review* 2013, 103(4): 1289–1324.
- Rao, Gautam (2013). Familiarity does not breed contempt: Diversity, discrimination and generosity in Delhi schools.
- Robins, James; Rotnitzky, Andrea; y Zhao, Lue Ping (1994). "Estimation of regression coefficients when some regressors are not always observed". *Journal of the American Statistical Association*. Vol 89 846–866.
- Robins, James y Rotnitzky, Andrea (1995). "Semiparametric Efficiency in Multivariate Regression Models with Missing Data". *Journal of the American Statistical Association*, Vol. 90(429): 122–29.
- Robins, James; Rotnitzky, Andrea; y Zhao, Lue Ping (1995). "Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data". *Journal of the American Statistical Association*, Vol. 90(429): 106–21.
- Rosell, Christine (1992). "The carrot or the stick for school desegregation". Philadelphia: Temple University Press.
- Rosenbaum, Paul y Rubin, Donald (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika* 70(1), 41-55.
- Rosenbaum, Paul y Rubin, Donald (1984). "Reducing bias in observational studies using subclassification on the propensity score". *Journal of the American Statistical Association*, Vol 79, 516 – 524.
- Rosenbaum, Paul (1984). "The consequences of adjustment for a concomitant variable that has been affected by the treatment". *The Journal of the Royal Statistical Society A* 147, 656–666.
- Rosenbaum, Paul y Rubin, Donald (1985). "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score". *The American Statistician*. **39** 33–38.
- Rubin, Donald (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies", *Journal of Educational Psychology*, Vol. 66, No.5, (1974), pp. 689.
- Rubin, Donald (2006). "Matched Sampling for Causal Effects". Cambridge and New York: Cambridge University Press.
- Saporito, Salvatore (1998). "The structural consequences of strategies social action: Increasing racial segregation and socioeconomic stratification through school choice" (Doctoral dissertation, Temple University, 1998). *Dissertation Abstracts International*, 59, 3982.
- Schafer, Joseph y Kang, Joseph (2008). "Average causal effects from nonrandomized studies: A practical guide and simulated case study". *Psychological Methods* **13** 279–313.

- Schmidt, Rebecca (2011). "The sorting hat: processes of setting, streaming, tracking and ability grouping in the United States and England". Presentación en 55th Annual Conference of Comparative and International Education Society (CIES), Montreal, Canadá.
- Seaton, Marjorie; Marsh, Herbert, y Craven, Rhonda (2010). "Big-Fish-Little-Pond Effect: Generalizability and Moderation– Two Sides of the Same Coin". *American Educational Research Journal* June 2010, Vol. 47, No. 2, pp. 390–433
- Steiner, Peter y Cook, David (2013). "Matching and propensity score". En *The Oxford Handbook of Quantitative Methods*, Volume 1, 2013. Oxford University Press.
- Stuart, Elizabeth y Rubin, Donald (2007). "Best Practices in Quasi-Experimental Designs: Matching methods for causal inference". Chapter 11 (pp. 155-176) in *Best Practices in Quantitative Social Science*. J. Osborne (Ed.). Thousand Oaks, CA: Sage Publications.
- Stuart, Elizabeth (2010). "Matching Methods for Causal Inference: A review and a look forward". *Statistical Science* 25(1): 1-21.
- The Sutton Trust (2009). "Attainment gaps between the most deprived and advantaged schools".
- Valenzuela, Juan Pablo. y Allende, Claudio (2011). "Logros en Liceos Públicos de Excelencia en Chile: valor agregado o solo descreme de la elite?", Borrador CIAE-Universidad de Chile.
- Valenzuela, Juan Pablo; Allende, Claudio; Sevilla, Alejandro; Egaña, Pablo (2012). "La (ina)movilidad del desempeño educativo de los estudiantes chilenos: realidad, oportunidades y desafíos". Proyecto FONIDE N° 811333. En "Evidencias para políticas públicas en educación: Selección de Investigaciones Sexto Concurso FONIDE". Centro de Estudios MINEDUC. ISBN: 978-956-292-304-0. Chile. Santiago. 2012. Páginas 187-231.
- Valenzuela, Juan Pablo.; Bellei, Cristian y De los Ríos, Danae (2009). "Evolución de la segregación socioeconómica y su relación con el financiamiento compartido". En "Selección de Investigaciones Primer Concurso FONIDE: Evidencias para Políticas Públicas en Educación, MINEDUC y FONIDE", págs. 229-284.
- Valenzuela, Juan Pablo.; Villalobos, Cristóbal; y Gabriela Gómez (2013). "Segregación y polarización en el sistema escolar chileno y recientes tendencias: ¿Qué ha sucedido con los grupos medios?". Documento de Referencia N°3, Espacio Público, Agosto 2013.
- Walsh, Patrick (2009). "Effects of school choice on the margin: The cream is already skimmed", *Economics of Education Review* 28, 227-236.
- Wooldridge, Jeffrey (2007). "Inverse probability weighted estimation for general missing data problems". *Journal of Econometrics* 141:1281–1301.
- Wooldridge, Jeffrey (2010). "Econometric Analysis of Cross Section and Panel Data". 2nd ed. Cambridge, MA: MIT Press.
- Zell, Ethan y Alicke, Mark (2009). "Contextual neglect, self-evaluation, and the frog-pond effect". *Journal of Personality and Social Psychology*. Vol 97, No 3, Septiembre 2009, páginas 467-482.
- Zhang, Hongliang (2008). "Magnet Schools and Student Achievement: Evidence from a Randomized Natural Experiment in China". Job Market paper. Department of Economics, MIT.

Anexo 1: Liceos Públicos Selectivos de Excelencia

RBD	Nombre del Colegio	Comuna	Certificado de notas (*)	Prueba de ingreso (**)	Promedio SIMCE 2008 Lenguaje	Promedio SIMCE 2008 Matemática	Género del colegio	Cursos que imparten	Tipo de enseñanza
2110	Liceo Municipal Óscar Castro Zúñiga	Rancagua	96.0%	93.9%	286.37	284.77	mixto	básica 7° y 8°; media	Científico Humanista
2973	Liceo Abate Molina	Talca	96.5%	95.5%	303.84	315.28	mixto	básica 7° y 8°; media	Científico Humanista
7325	Liceo Carmela Carvajal de Prat	Osorno	92.9%	91.9%	289.42	285.93	mixto	básica 7° y 8°; media	Científico Humanista
8002	Liceo Galvarino Riveros Cárdenas	Castro	88.1%	80.1%	271.30	281.26	mixto	básica 7° y 8°; media	Científico Humanista
8485	Liceo Instituto Nacional	Santiago	95.9%	95.3%	337.77	365.71	solo hombres	básica 7° y 8°; media	Científico Humanista
8487	Liceo Javiera Carrera	Santiago	96.4%	97.2%	320.52	335.34	solo mujeres	básica 7° y 8°; media	Científico Humanista
8489	Liceo Teresa Prat de Sarratea	Santiago	96.9%	90.7%	298.98	305.69	solo mujeres	básica 7° y 8°; media	Científico Humanista
8491	Liceo de Aplicación Rector Jorge E. Schneider	Santiago	94.2%	93.2%	291.22	306.34	solo hombres	básica 7° y 8°; media	Científico Humanista
8499	Internado Nacional Barros Arana	Santiago	99.3%	89.3%	294.03	314.12	solo hombres	básica 7° y 8°; media	Científico Humanista
8926	Liceo de niñas N°7 Luisa Saavedra de González	Providencia	95.9%	95.9%	322.58	344.28	solo mujeres	básica 7° y 8°; media	Científico Humanista
8927	Liceo Carmela Carvajal de Prat	Providencia	96.6%	99.4%	331.42	362.43	solo mujeres	básica 7° y 8°; media	Científico Humanista
9071	Liceo Augusto D'Halmar	Ñuñoa	98.6%	93.9%	319.25	344.23	mixto	básica 7° y 8°; media	Científico Humanista
9073	Liceo Carmela Silva Donoso	Ñuñoa	98.0%	92.6%	319.27	345.02	solo mujeres	básica 7° y 8°; media	Científico Humanista
9074	Liceo José Toribio medina	Ñuñoa	95.7%	73.6%	286.86	280.67	mixto	básica 7° y 8°; media	Científico Humanista
9406	Liceo Andrés Bello	San Miguel	93.9%	92.1%	295.26	304.01	solo hombres	básica 7° y 8°; media	Científico Humanista
25770	Liceo Nacional de Maipú	Maipú	95.9%	97.8%	313.89	338.88	solo hombres	básica 7° y 8°; media	Científico Humanista
25855	Colegio municipal instituto Chacabuco	Colina	97.9%	87.6%	293.90	306.64	mixto	básica 7° y 8°; media	Científico Humanista

(*) Porcentaje de padres que informó que se les solicitó entregar el certificado de notas del colegio anterior del alumno postulante.

(**) Porcentaje de padres que informó que los alumnos tuvieron que realizar una prueba para ingresar al colegio.

Anexo 2: Mejora en el balance de las variables medidas antes del tratamiento

Distintas muestras de la cohorte 2002-2008 (Utilizando soporte común para Lenguaje en bases de ex compañeros)

	Diferencias Normalizadas Tratados - Controles		
	Base completa	Ex-compañeros	Emparejados
Promedio SIMCE Lenguaje (estandarizado)	0.61	0.28	0.27
Promedio SIMCE Matemática (estandarizado)	0.61	0.32	0.33
Capital Social, Cultural y Económico			
Libros menor a 10	-0.28	-0.08	-0.07
Libros mayor que 11 y menor a 50	0.01	-0.01	-0.03
Libros mayor que 51	0.23	0.07	0.07
Educación madre	0.31	0.08	0.06
Educación padre	0.30	0.07	0.05
Ingreso del Hogar (estandarizado)	0.05	-0.06	-0.01
Características Individuales			
Sexo (0 hombre; 1 Mujer)	-0.04	-0.06	-0.07
Repite antes de 4to básico (0 no; 1 si)	-0.15	-0.02	0.05
Asistió a educación pre escolar (0 no; 1 si)	0.14	0.01	-0.06
Promedio notas (5to y 6to básico)	0.76	0.61	0.63
Características del establecimiento			
Número de Alumnos en el Curso	0.22	-0.04	-0.03
Matricula educación básica del colegio (estandarizada)	0.17	-0.03	-0.04
Tasa de repitencia educación básica del colegio	-0.19	0.04	0.05
Ruralidad (0 Urbano; 1 Rural)	-0.17	0.01	0.02
Dependencia público	-0.15	0.16	0.18
Dependencia particular Subvencionado	0.21	-0.10	-0.12
Dependencia particular pagado	-0.11	-0.12	-0.11
Efecto par educación madre	0.24	-0.11	-0.12
Efecto par educación padre	0.23	-0.11	-0.12
Efecto par Ingreso	0.00	-0.15	-0.14
Efecto par habilidad individual lenguaje	0.01	-0.06	-0.03
Efecto par habilidad individual matemática	-0.01	-0.08	-0.06

<i>Razones de ingreso al establecimiento</i>			
Razón de ingreso: Cercanía del establecimiento	-0.41	-0.44	-0.47
Razón de ingreso: Otros miembros de la familia estudian o estudiaron en el establecimiento	-0.20	-0.18	-0.17
Razón de ingreso: Prestigio del establecimiento	0.71	0.60	0.61
Razón de ingreso: Buenos resultados en PSU	0.84	0.85	0.88
<i>Expectativas de los padres sobre educación futura</i>			
Expectativa padres: Básica, media incompleta o técnico	-0.42	-0.18	-0.18
Expectativa padres: media CH, IP o CFT	-0.13	-0.10	-0.11
Expectativa padres: Universidad o postgrado	0.46	0.21	0.22
N° Observaciones	168775	16685	10323
Tratados	4591	2759	1856
Controles	164184	13926	8467

Anexo 3: Mejora en el balance de las variables medidas antes del tratamiento

Distintas muestras de la cohorte 2006-2012 (Utilizando soporte común para Lenguaje en bases de ex compañeros)

	Diferencias Normalizadas Tratados - Controles		
	Base completa	Ex-compañeros	Emparejados
Promedio SIMCE Lenguaje estandarizado	0.62	0.27	0.27
Promedio SIMCE Matemática estandarizado	0.67	0.31	0.29
Capital Social, Cultural y Económico			
Libros menor a 10	-0.25	-0.04	-0.03
Libros mayor que 11 y menor a 50	0.12	-0.01	-0.02
Libros mayor que 51	0.15	0.05	0.06
Educación madre	0.32	0.02	0.02
Educación padre	0.31	0.04	0.04
Ingreso del Hogar (divido en 100000)	0.00	-0.12	-0.08
Características Individuales			
Sexo (0 hombre; 1 Mujer)	-0.05	-0.12	-0.12
Repite antes de 4to básico (0 no; 1 si)	-0.15	-0.03	0.08
Asistió a educación pre escolar (0 no; 1 si)	0.12	0.02	-0.05
Promedio notas (5to y 6to básico)	0.76	0.60	0.58
Características del establecimiento			
Número de Alumnos en el Curso	0.30	-0.05	-0.08
Matricula educación básica del colegio	0.24	-0.02	-0.03
Tasa de Repitencia colegio	-0.15	0.08	0.08
Ruralidad (0 Urbano; 1 Rural)	-0.16	0.06	0.06
Dependencia público	-0.16	0.19	0.19
Dependencia particular Subvencionado	0.24	-0.14	-0.15
Dependencia particular pagado	-0.17	-0.10	-0.07
Efecto par educación madre	0.22	-0.18	-0.17
Efecto par educación padre	0.21	-0.17	-0.16
Efecto par Ingreso	-0.04	-0.17	-0.14
Efecto par habilidad individual lenguaje	0.00	-0.06	-0.06
Efecto par habilidad individual matemática	-0.02	-0.10	-0.10

<i>Razones de ingreso al establecimiento</i>			
Razón de ingreso: Cercanía del establecimiento	-0.45	-0.48	-0.47
Razón de ingreso: Excelencia académica y resultados PSU/SIMCE	1.08	0.91	0.89
Razón de ingreso: Asisten niños del mismo ambiente y realidad familiar	-0.12	-0.13	-0.13
Razón de ingreso: Hermanos estaban en establecimiento	-0.12	-0.11	-0.11
<i>Expectativas de los padres sobre educación futura</i>			
Expectativa padres: Básica, media incompleta o técnico	-0.36	-0.10	-0.09
Expectativa padres: media CH, IP o CFT	-0.16	-0.09	-0.09
Expectativa padres: Universidad o postgrado	0.41	0.14	0.13
N° Observaciones	168775	11425	10578
Tratados	3236	1919	1918
Controles	161288	9506	8660

Anexo 4: Estadísticas descriptivas panel 2002-2008

Número de Observaciones	Tratados (T)		Controles (C)		Diferencia T - C		Test-t	T - C
	(4591)		(164184)		2002	2008	2002	2008
Promedio SIMCE Lenguaje	301.89 (36.53)	307.96 (40.00)	265.13 (48.53)	260.16 (49.98)	36.76	47.80	49.34	64.17
Promedio SIMCE Matemática	299.67 (40.11)	323.29 (49.22)	260.67 (49.97)	257.32 (61.57)	39.00	65.97	50.78	71.82
Promedio SIMCE Lenguaje estandarizado	0.74 (0.75)	0.92 (0.79)	-0.02 (1.00)	-0.03 (0.99)	0.76	0.95		
Promedio SIMCE Matemática estandarizado	0.76 (0.80)	1.03 (0.79)	-0.02 (1.00)	-0.03 (0.99)	0.78	1.06		
Capital Social, Cultural y Económico								
Libros menor a 10	0.15 (0.35)	0.11 (0.32)	0.31 (0.46)	0.25 (0.43)	-0.16	-0.14	-23.15	-20.31
Libros mayor que 11 y menor a 50	0.43 (0.50)	0.43 (0.49)	0.42 (0.49)	0.47 (0.50)	0.01	-0.04	1.07	-4.67
Libros mayor que 51	0.42 (0.49)	0.46 (0.50)	0.27 (0.44)	0.28 (0.45)	0.15	0.18	22.87	24.54
Educación madre	12.48 (2.97)	12.82 (2.95)	11.01 (3.73)	11.33 (3.70)	1.47	1.49	26.04	25.75
Educación padre	12.56 (3.19)	12.92 (3.19)	11.05 (3.86)	11.39 (3.86)	1.51	1.53	25.64	25.45
Ingreso del Hogar (estandarizado)	0.16 (0.76)	0.10 (0.79)	0.09 (1.09)	0.04 (1.04)	0.07	0.06	4.17	3.32
Nivel socioeconómico(NSE)	0.51 (0.78)	0.43 (0.76)	0.14 (0.99)	0.06 (1.00)	0.37	0.37	19.80	23.56
Quintil de NSE								
Quintil I de NSE	0.06 (0.24)	0.06 (0.24)	0.21 (0.41)	0.21 (0.40)	-0.15	-0.15	-19.67	-23.08
Quintil II de NSE	0.13 (0.33)	0.12 (0.33)	0.20 (0.40)	0.21 (0.40)	-0.07	-0.09	-8.96	-13.14
Quintil III de NSE	0.24 (0.43)	0.25 (0.43)	0.25 (0.43)	0.22 (0.42)	-0.01	0.03	-1.17	3.39
Quintil IV de NSE	0.27 (0.45)	0.29 (0.45)	0.14 (0.35)	0.17 (0.37)	0.13	0.12	19.42	21.06
Quintil V de NSE	0.30 (0.46)	0.28 (0.45)	0.20 (0.40)	0.20 (0.40)	0.10	0.08	12.95	12.95
Características Individuales								
Sexo (0 hombre; 1 Mujer)	0.49 (0.50)	0.49 (0.50)	0.52 (0.50)	0.52 (0.50)	-0.03	-0.03	-4.15	-4.38
Repite antes de 4to básico (0 no; 1 si)	0.01 (0.09)		0.04 (0.20)		-0.03		-11.10	
Asistió a educación pre escolar (0 no; 1 si)	0.99 (0.11)		0.95 (0.21)		0.04		10.61	
Promedio notas (5to y 6to básico)	6.36 (0.39)		5.87 (0.52)		0.49		63.60	

<i>Características del establecimiento</i>								
Número de Alumnos en el Curso	35.74 (8.42)	40.11 (3.96)	32.96 (9.13)	33.65 (7.13)	2.78	6.46	20.43	61.04
Matricula educación básica del colegio	840.84 (693.77)	1649.12 (670.02)	696.66 (501.77)	743.60 (545.13)	144.18	905.52	18.15	110.24
Tasa de Repitencia colegio	0.02 (0.02)	0.04 (0.05)	0.03 (0.03)	0.08 (0.05)	-0.01	-0.04	-15.55	-48.09
Ruralidad (0 Urbano; 1 Rural)	0.05 (0.21)	0.00 (0.00)	0.11 (0.31)	0.03 (0.16)	-0.06	-0.03	-13.64	-11.39
Dependencia público	0.38 (0.49)	1.00 (0.00)	0.49 (0.50)	0.36 (0.48)	-0.11	0.64	-14.38	91.01
Dependencia particular Subvencionado	0.56 (0.50)		0.42 (0.49)	0.56 (0.50)	0.14		19.67	-76.25
Dependencia particular pagado	0.05 (0.22)		0.09 (0.28)	0.08 (0.28)	-0.04		-8.92	-20.60
Efecto par educación madre	11.57 (1.86)	12.82 (1.01)	10.83 (2.51)	11.32 (2.42)	0.74	1.50	19.36	40.10
Efecto par educación padre	11.63 (1.97)	12.93 (1.14)	10.89 (2.61)	11.39 (2.55)	0.74	1.54	18.65	39.00
Efecto par Ingreso	0.07 (0.53)	0.09 (0.30)	0.08 (0.88)	0.05 (0.87)	-0.01	0.04	-0.37	3.43
Efecto par habilidad individual lenguaje	0.00 (0.23)	0.60 (0.29)	-0.00 (0.26)	-0.02 (0.38)	0.00	0.62	0.62	94.43
Efecto par habilidad individual matemática	-0.00 (0.24)	0.66 (0.32)	0.00 (0.27)	-0.02 (0.39)	0.00	0.68	-0.56	100.57
<i>Razones de ingreso al establecimiento</i>								
Razón de ingreso: Cercanía del establecimiento		0.12 (0.32)		0.35 (0.48)			-0.23	-31.39
Razón de ingreso: Otros miembros de la familia estudian o estudiaron en el establecimiento		0.19 (0.39)		0.31 (0.46)			-0.12	-16.64
Razón de ingreso: Prestigio del establecimiento		0.90 (0.30)		0.48 (0.50)		0.42		53.51
Razón de ingreso: Buenos resultados en PSU		0.59 (0.49)		0.11 (0.31)		0.48		98.25
<i>Expectativas de los padres sobre educación futura</i>								
Expectativa padres: Básica, media incompleta o técnico	0.09 (0.28)		0.31 (0.46)		-0.22			-32.42
Expectativa padres: media CH, IP o CFT	0.11 (0.31)		0.17 (0.38)		-0.06			-10.68
Expectativa padres: Universidad o postgrado	0.81 (0.39)		0.52 (0.50)		0.29		38.03	

Anexo 5: Estadísticas descriptivas panel 2006-2012

	Tratados (T)		Controles ©		Diferencia T - C		Test-t	T - C
	(3236)		(161288)					
	2006	2012	2006	2012	2006	2012		
Promedio SIMCE Lenguaje	304.07 (37.30)	311.62 (38.57)	265.48 (50.05)	265.54 (52.60)	38.59	46.08	43.54	48.20
Promedio SIMCE Matemática	303.48 (36.56)	329.07 (44.23)	260.97 (51.35)	272.79 (61.43)	42.51	56.28	46.73	50.40
Promedio SIMCE Lenguaje estandarizado	0.75 (0.74)	0.86 (0.73)	-0.02 (1.00)	-0.02 (1.00)	0.77	0.88		
Promedio SIMCE Matemática estandarizado	0.81 (0.71)	0.89 (0.72)	-0.02 (1.00)	-0.02 (1.00)	0.83	0.91		
Capital Social, Cultural y Económico								
Libros menor a 10	0.30 (0.46)	0.10 (0.30)	0.46 (0.50)	0.21 (0.41)	-0.16	-0.11	-18.23	-13.40
Libros mayor que 11 y menor a 50	0.46 (0.50)	0.44 (0.50)	0.38 (0.48)	0.48 (0.50)	0.08	-0.04	9.43	-3.91
Libros mayor que 51	0.24 (0.43)	0.46 (0.50)	0.16 (0.37)	0.31 (0.46)	0.08	0.15	12.28	16.04
Educación madre	13.10 (2.71)	12.99 (2.93)	11.68 (3.57)	11.82 (3.65)	1.42	1.17	22.00	16.07
Educación padre	13.26 (3.00)	13.04 (3.17)	11.79 (3.77)	11.81 (3.83)	1.47	1.23	21.57	16.07
Ingreso del Hogar (estandarizado)	4.79 (3.92)	5.07 (4.04)	4.77 (5.73)	5.05 (5.20)	0.02	0.02	0.17	0.16
Nivel socioeconómico (NSE)	0.49 (0.73)	0.34 (0.77)	0.15 (1.01)	0.06 (1.01)	0.34	0.28	18.83	13.73
Quintil de NSE								
Quintil I de NSE	0.06 (0.24)	0.07 (0.25)	0.20 (0.40)	0.21 (0.40)	-0.14	-0.14	-19.18	-16.96
Quintil II de NSE	0.12 (0.33)	0.14 (0.35)	0.20 (0.40)	0.20 (0.40)	-0.08	-0.06	-11.21	-7.32
Quintil III de NSE	0.21 (0.41)	0.25 (0.43)	0.20 (0.40)	0.21 (0.41)	0.01	0.04	1.22	5.41
Quintil IV de NSE	0.34 (0.47)	0.32 (0.47)	0.19 (0.39)	0.19 (0.39)	0.15	0.13	19.99	16.23
Quintil V de NSE	0.27 (0.44)	0.22 (0.41)	0.20 (0.40)	0.20 (0.40)	0.07	0.02	9.35	2.83
Características Individuales								
Sexo (0 hombre; 1 Mujer)	0.49 (0.50)		0.52 (0.50)		-0.03		-3.95	
Repite antes de 4to básico (0 no; 1 si)	0.01 (0.09)		0.04 (0.20)		-0.03		-9.37	
Asistió a educación pre escolar (0 no; 1 si)	0.99 (0.07)		0.97 (0.16)		0.02		7.16	
Promedio notas (5to y 6to básico)	6.30 (0.40)		5.81 (0.51)		0.49		53.92	

Características del establecimiento								
Número de Alumnos en el Curso	35.20 (7.97)	41.20 (3.64)	31.46 (9.70)	36.39 (9.64)	3.74	4.81	21.78	28.32
Matricula educación básica del colegio	810.33 (671.77)	1635.05 (799.12)	616.79 (454.78)	582.67 (422.67)	193.54	1052.38	23.56	136.80
Tasa de Repitencia colegio	0.03 (0.03)	0.05 (0.05)	0.04 (0.03)	0.07 (0.05)	-0.01	-0.02	-10.50	-28.30
Ruralidad (0 Urbano; 1 Rural)	0.05 (0.22)	0.00 (0.00)	0.11 (0.32)	0.03 (0.18)	-0.06	-0.03	-11.24	-10.31
Dependencia público	0.33 (0.47)	1.00 (0.00)	0.44 (0.50)	0.31 (0.46)	-0.11	0.69	-12.33	84.18
Dependencia particular Subvencionado	0.64 (0.48)		0.47 (0.50)	0.59 (0.49)	0.17		18.67	-68.29
Dependencia particular pagado	0.03 (0.17)		0.09 (0.28)	0.10 (0.29)	-0.06		-11.51	-18.55
Efecto par educación madre	12.18 (1.85)	12.97 (1.02)	11.50 (2.48)	11.81 (2.40)	0.68	1.16	14.95	24.39
Efecto par educación padre	12.33 (1.92)	13.06 (1.19)	11.64 (2.62)	11.81 (2.56)	0.69	1.25	14.57	24.49
Efecto par Ingreso	4.49 (2.85)	5.06 (1.61)	4.69 (4.87)	5.06 (4.26)	-0.20		-2.31	0.01
Efecto par habilidad individual lenguaje	0.00 (0.24)	0.66 (0.26)	0.00 (0.25)	-0.02 (0.35)	0.00	0.68	0.22	97.06
Efecto par habilidad individual matemática	-0.01 (0.26)	0.70 (0.27)	0.00 (0.26)	-0.02 (0.37)	-0.01	0.72	-1.38	96.84
Razones de ingreso al establecimiento								
Razón de ingreso: Cercanía del establecimiento		0.11 (0.31)		0.37 (0.48)			-0.26	-26.92
Razón de ingreso: Excelencia académica y resultados PSU/SIMCE		0.96 (0.19)		0.40 (0.49)		0.56		57.96
Razón de ingreso: Asisten niños del mismo ambiente y realidad familiar		0.12 (0.33)		0.18 (0.38)		-0.06		-7.75
Razón de ingreso: Hermanos estaban en establecimiento		0.11 (0.31)		0.17 (0.37)		-0.06		-8.03
Expectativas de los padres sobre educación futura								
Expectativa padres: Básica, media incompleta o técnico	0.06 (0.24)		0.23 (0.42)		-0.17		-22.56	
Expectativa padres: media CH, IP o CFT	0.09 (0.29)		0.17 (0.37)		-0.08		-10.94	
Expectativa padres: Universidad o postgrado	0.85 (0.36)		0.60 (0.49)		0.25		27.75	

Anexo 6: Distribución de los grupos de control y tratamiento en la región de soporte común

Base ex compañeros para ambas cohortes.

Figura A6.1: Área de soporte común grupo de control y tratamiento sector lenguaje cohorte 2002-2008

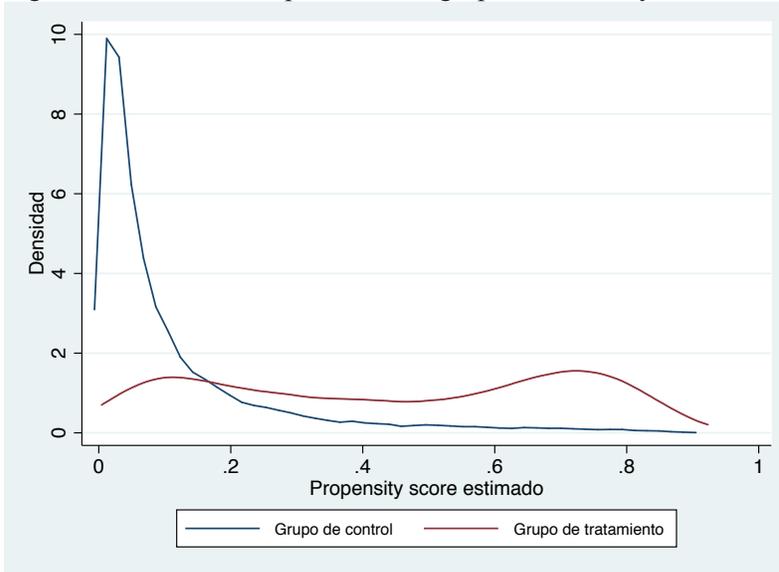


Figura A6.2: Área de soporte común grupo de control y tratamiento sector matemática cohorte 2002-2008

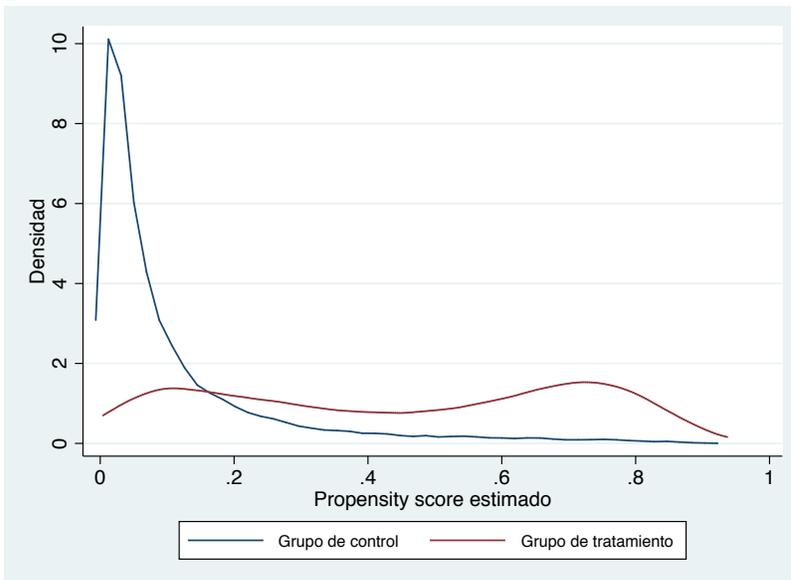


Figura A6.3: Área de soporte común grupo de control y tratamiento sector lenguaje cohorte 2006-2012

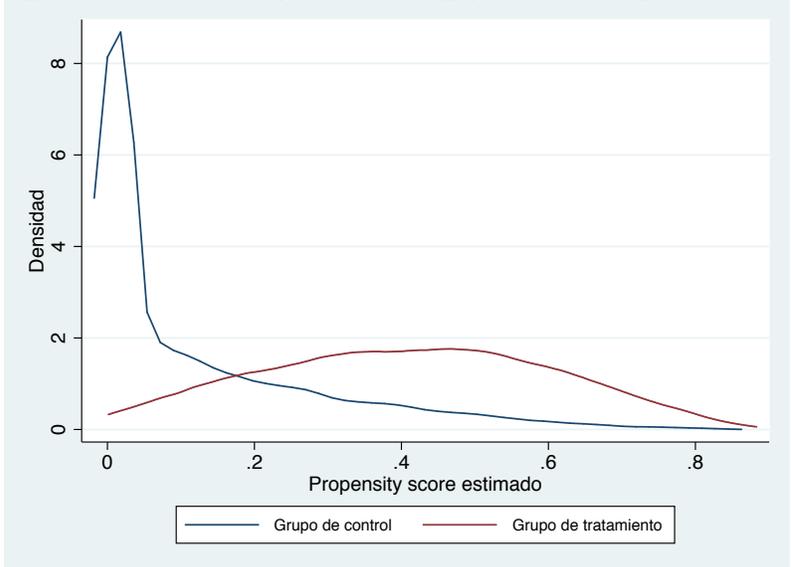
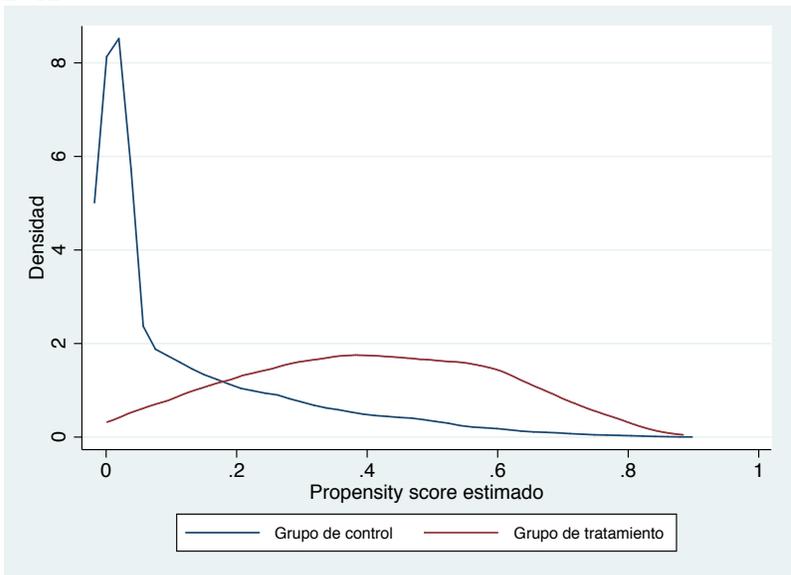


Figura A6.4: Área de soporte común grupo de control y tratamiento sector matemática cohorte 2006-2012



Anexo 7: Probabilidad de ingresar a un LSPE con grupo de control ex compañeros emparejados

VARIABLES	2002-2008		2006-2012	
	Lenguaje	Matemática	Lenguaje	Matemática
Promedio SIMCE Lenguaje (estandarizado)	0.161*** (0.059)		0.212*** (0.054)	
Promedio SIMCE Matemática estandarizado		0.301*** (0.056)		0.241*** (0.055)
<i>Capital Social, Cultural y Económico</i>				
Educación madre	-0.003 (0.016)	-0.005 (0.016)	-0.001 (0.015)	-0.001 (0.015)
Educación padre	-0.000 (0.015)	0.001 (0.015)	0.038*** (0.014)	0.037*** (0.014)
Ingreso del Hogar (estandarizado)	-0.092* (0.048)	-0.107** (0.049)	-0.254*** (0.047)	-0.259*** (0.047)
Libros mayor que 11 y menor a 50	0.118 (0.103)	0.109 (0.103)	0.050 (0.073)	0.046 (0.073)
Libros mayor que 51	0.203* (0.109)	0.198* (0.109)	0.162* (0.088)	0.164* (0.088)
<i>Características Individuales</i>				
Repitente antes de 4to básico (0 no; 1 si)	2.761*** (0.741)	2.768*** (0.749)	3.431*** (0.418)	3.466*** (0.415)
Asistió a educación pre escolar (0 no; 1 si)	-1.910*** (0.389)	-1.841*** (0.399)		
Promedio notas (5to y 6to básico)	2.464 (3.381)	3.229 (3.390)	2.455 (3.180)	2.715 (3.174)
Promedio notas al cuadrado (5to y 6to básico)	-0.045 (0.269)	-0.112 (0.270)	-0.022 (0.254)	-0.043 (0.254)
<i>Razones de ingreso al establecimiento</i>				
Razón de ingreso: Cercanía del establecimiento	-0.920*** (0.093)	-0.909*** (0.093)	-1.126*** (0.089)	-1.116*** (0.089)
Razón de ingreso: Otros miembros de la familia estudian o estudiaron en el establecimiento	-0.088 (0.077)	-0.090 (0.077)		
Razón de ingreso: Prestigio del establecimiento	1.304*** (0.094)	1.301*** (0.094)		
Razón de ingreso: Buenos resultados en PSU	2.004*** (0.068)	1.999*** (0.068)		
Razón de ingreso: Excelencia académica y resultados PSU/SIMCE			3.208*** (0.143)	3.202*** (0.143)
Razón de ingreso: Asisten niños del mismo ambiente y realidad familiar			-0.331*** (0.088)	-0.326*** (0.088)
Razón de ingreso: Hermanos estaban en establecimiento			-0.109 (0.091)	-0.111 (0.092)
<i>Expectativas de los padres sobre educación futura</i>				
Expectativa padres: media CH, IP o CFT	0.331** (0.142)	0.329** (0.143)	0.272* (0.159)	0.254 (0.158)
Expectativa padres: Universidad o postgrado	0.627*** (0.117)	0.622*** (0.118)	0.535*** (0.135)	0.517*** (0.135)
<i>Características del establecimiento</i>				
Dependencia público	0.522*** (0.073)	0.537*** (0.073)	0.583*** (0.073)	0.583*** (0.074)
Dependencia particular pagado	-1.147*** (0.170)	-1.179*** (0.171)		
Matrícula educación básica del colegio (estandarizada)	0.026 (0.017)	0.025 (0.017)	0.021 (0.015)	0.020 (0.015)
Tasa de repitencia educación básica del colegio	5.369*** (1.733)	5.156*** (1.735)	7.681*** (1.329)	7.767*** (1.327)
Ruralidad (0 Urbano; 1 Rural)	-0.074 (0.200)	-0.032 (0.199)	0.305* (0.186)	0.298 (0.185)

Constante	-15.798 (10.627)	-18.088* (10.654)	-19.908** (9.923)	-20.695** (9.905)
Número de Observaciones	10,415	10,404	10,939	10,935
R-cuadrado	0.342	0.344	0.321	0.321

Errores estándar robustos en paréntesis

*** Significancia al 1%; **Significancia al 5%; * Significancia al 10%

Anexo 8: Estimación de ATT solo para la base reducida con grupo de control ex compañeros emparejados

Tabla A8.1: Efecto promedio del tratamiento sobre los tratados sector Lenguaje (ATT), modelos basados en propensity score

	Panel SIMCE		Panel SIMCE	
	2002-2008	2006-2012	2002-2008	2006-2012
	Efecto estimado (ATT)		Test de validez	
Estratificación	0.252*** (0.072)	0.168*** (0.048)	0.025 (0.023)	0.022 (0.019)
ANCOVA ajustado	0.246*** (0.071)	0.177*** (0.050)	0.031 (0.022)	0.028 (0.019)
Wooldridge's double-robust	0.236*** (0.061)	0.167*** (0.050)	0.019 (0.013)	0.009 (0.012)
Nº de Observaciones	10145	10337	10178	10417
Tratados	1843	1899	1846	1905
Controles	8302	8438	8332	8512

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%

Tabla A8.2: Efecto promedio del tratamiento sobre los tratados sector Matemática (ATT), modelos basados en propensity score

	Panel SIMCE		Panel SIMCE	
	2002-2008	2006-2012	2002-2008	2006-2012
	Efecto estimado (ATT)		Test de validez	
Estratificación	0.250*** (0.085)	0.123** (0.055)	0.022 (0.025)	0.000 (0.019)
ANCOVA ajustado	0.271*** (0.084)	0.130** (0.056)	0.037 (0.027)	0.002 (0.019)
Wooldridge's double-robust	0.251*** (0.078)	0.109* (0.057)	0.021 (0.016)	-0.012 (0.010)
N	10131	10235	10166	10316
Tratados	1838	1890	1845	1904
Controles	8293	8345	8321	8412

Errores estándar en paréntesis

*** Significancia al 1%; ** Significancia al 5%; * Significancia al 10%