



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA

# **IMPACTO EN HABILIDADES COGNITIVAS DE LA EDUCACIÓN PREESCOLAR EN CHILE**

**MAGDALENA BENNETT COLOMER**

Tesis para optar al grado de  
Magister en Ciencias de la Ingeniería

Profesor Supervisor:  
**RICARDO PAREDES MOLINA**

Santiago de Chile, Noviembre, 2012

© 2012, Magdalena Bennett Colomer



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA

# **IMPACTO EN HABILIDADES COGNITIVAS DE LA EDUCACIÓN PREESCOLAR EN CHILE**

**MAGDALENA BENNETT COLOMER**

Tesis presentada a la Comisión integrada por los profesores:

**LUIS CIFUENTES LIRA**

**JORGE MANZI ASTUDILLO**

**MIGUEL NUSSBAUM VOEHL**

**RICARDO PAREDES MOLINA**

Para completar las exigencias del grado de  
Magister en Ciencias de la Ingeniería  
Santiago de Chile, Noviembre, 2012

*Esfuerzo y trabajo*

## **AGRADECIMIENTOS**

Agradezco a la Fundación Integra, especialmente a Natalia Sandoval, por la excelente disposición y entrega de datos para llevar a cabo esta investigación.

Agradezco a mis profesores, los que me ayudaron a formar mi camino. A Luis Cifuentes, por la confianza y el apoyo en la construcción de una Escuela y una Universidad más justa, y a Ricardo Paredes, mi profesor guía y mentor, el que me enseñó con dedicación y cariño lo que hoy es mi vocación.

Agradezco a mis amigos, dentro y fuera de la Universidad, por el ánimo constante, las largas conversaciones y las batallas emprendidas en conjunto. Esto no hubiera sido lo mismo sin ustedes.

A mi familia, especialmente a mis papás y hermanos, por el cariño y apoyo en esta tarea. A ellos, mi gratitud y eterna admiración.

## INDICE GENERAL

	Pág.
DEDICATORIA.....	ii
AGRADECIMIENTOS .....	iii
INDICE DE TABLAS .....	vi
INDICE DE FIGURAS .....	vii
RESUMEN.....	viii
ABSTRACT.....	ix
1. INTRODUCCIÓN.....	1
2. LA LITERATURA .....	7
3. DATOS Y METODOLOGÍA.....	12
3.1 Un Modelo Ad hoc a la disposición de datos.....	12
3.2 Identificación de efectos.....	15
3.3 Inscripción del niño en el sistema preescolar .....	19
3.4 Asistencia del niño al preescolar .....	22
3.4.1 Mínimos cuadrados ordinarios (MCO) controlando por estado de salud .....	24
3.4.2 Estimadores ponderados en base de probabilidad inversa ( <i>Inverse                 Probability-Weighted Estimators</i> ) .....	25
3.4.3 <i>Matching</i> por medio de <i>propensity score</i> (PS) .....	26
4. RESULTADOS .....	28
4.1 Modelos de Corto Plazo .....	29
4.2 Modelos de Mediano Plazo .....	32
5. CONCLUSIONES .....	36
BIBLIOGRAFÍA.....	39

A N E X O S.....	42
Anexo A: Resultados para estimadores con peso de probabilidad inversa .....	43

## INDICE DE TABLAS

	Pág.
Tabla 3-1 – Descripción de datos socioeconómicos y de rendimiento para alumnos Integra según tratamiento .....	18
Tabla 4-1 - Resultados para modelo de corto plazo según método de control por estado de salud, estimadores de peso con probabilidad inversa y <i>matching</i> con <i>propensity score</i> .	31
Tabla 4-2 - Resultados para modelo de mediano plazo según método de control por estado de salud, estimadores de peso con probabilidad inversa y <i>matching</i> con <i>propensity score</i> .....	34
Tabla A-1 - Resultados Primera Etapa Probit para estimadores con peso de probabilidad inversa .....	43

## INDICE DE FIGURAS

	Pág.
Figura 1-1: Ingreso autónomo promedio por nivel educacional para individuos entre 25 y 35 años de edad que no estudian (Fuente: Urzúa, 2012).....	2
Figura 1-2: Evolución de cobertura preescolar por quintil de ingreso socioeconómico (Fuente: Elaboración propia en base a datos CASEN 1990, 1996, 2003, 2006 y 2009) ...	4
Figura 4-1 – (a) Distribución Puntaje TEPSI Lenguaje según tratamiento (b) Distribución Puntaje SIMCE Lenguaje según tratamiento .....	28

## RESUMEN

Este trabajo aborda el impacto de la educación preescolar en las habilidades cognitivas de los estudiantes en el corto y mediano plazo, con el objetivo de cuantificar el efecto de programas preescolares extensivos enfocados a niños vulnerables. Utilizando un panel de estudiantes observados en dos momentos del tiempo, durante su etapa preescolar y luego en cuarto año de enseñanza básica, se analiza el efecto que tiene la exposición a programas de desarrollo temprano en pruebas estandarizadas. Se comparan grupos de alta exposición con estudiantes de baja exposición mediante tres métodos: mínimos cuadrados ordinarios, estimadores con pesos de probabilidad inversa y *matching* a través de *propensity score*, encontrando efectos sobre las pruebas estandarizadas en el corto plazo de 16%-65% de la desviación estándar y en el mediano plazo de 4.3%-11.6% de la desviación estándar.

Palabras Claves: Educación preescolar, desarrollo infantil temprano, habilidades cognitivas tempranas

## **ABSTRACT**

This investigation studies the impact of preschool education on cognitive abilities in the short and medium term. The objective is to quantify the effect of extensive preschool programs focused on vulnerable children. Using panel data that observes students in two different times, during preschool and in 4<sup>th</sup> year of primary school, we analyze the effect that exposition to early development programs have on standardized tests. We compare a high-exposition group with a low-exposition one through three methods: ordinary least squares, inverse probability-weighted estimators and matching using propensity score. We find positive short-term effects on standardized tests of 16%-65% of the standardized deviation and positive medium-term effects of 4.3%-11.6% of the standardized deviation.

Keywords: Preschool education, early childhood development, early cognitive abilities

## 1. INTRODUCCIÓN

Chile es un país con fuerte desigualdad de ingresos y altos índices de inequidad, lo que se traduce en un coeficiente Gini de 0.52 según el Banco Mundial (2009). Parte de esa desigualdad se debe a la diferencia en el acceso a oportunidades, como la salud, el empleo y la educación, y también a diferencias en la calidad de los servicios provistos. De hecho, gran parte de la desigualdad de ingresos económicos puede explicarse en base a diferencias en niveles de escolaridad (Lopez & Perry, 2007).

A pesar de ser un país muy desigual, Chile tiene una alta movilidad social con respecto a otros países de América Latina y la escolaridad parece ser, al menos en parte, responsable de dicho proceso (Andersen, 2001). En efecto, la educación suele concebirse como un motor de movilidad social, particularmente en países en vías de desarrollo, donde las tasas de retorno por años de escolaridad son más elevadas que en países desarrollados. Al respecto, una pregunta relevante es en qué etapa de la vida es más rentable la inversión en capital humano. En Chile, las mayores diferencias en ingreso se dan cuando una persona logra titularse de la universidad (Figura 1-1), pero la probabilidad de llegar a acceder varía enormemente entre estratos sociales (Comisión de Financiamiento Estudiantil para la Educación Superior, 2012). Las fuertes diferencias en los resultados de la Prueba de Selección Universitaria (PSU) según dependencia de establecimientos son un síntoma más de diferencias de oportunidades, ya que la desigualdad no comienza en el acceso a la educación superior. En las pruebas estandarizadas como el SIMCE (Sistema de Medición de la Calidad Educativa), ya se observan diferencias promedio de casi 50 puntos en estudiantes de 4to básico que

asisten a establecimientos públicos versus privados. La pregunta que se deriva entonces naturalmente es qué herramientas de política pública permitirían entregar iguales oportunidades de acceder a los más altos niveles educacionales y, así, además, acceder a aumentar sus ingresos de manera significativa.

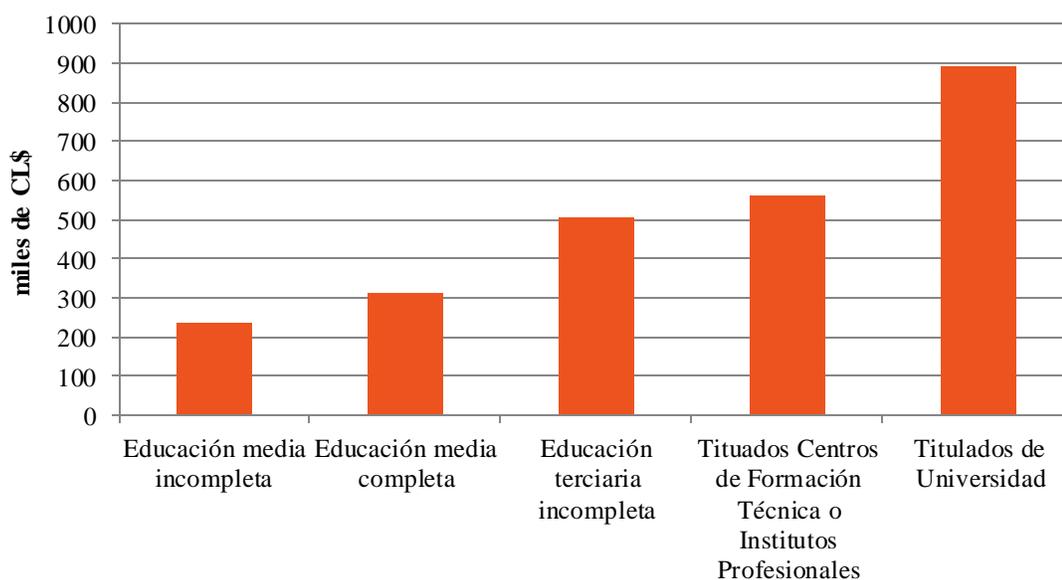


Figura 1-1: Ingreso autónomo promedio por nivel educacional para individuos entre 25 y 35 años de edad que no estudian (Fuente: Urzúa, 2012)

Una luz sobre la dirección de la política es la evidencia internacional que indica que gran parte de estas desigualdades se forjan antes de los 6 años de edad, profundizándose a través del tiempo y siendo casi prohibitivamente costoso revertirlas en años posteriores (Heckman, 1999; Heckman, 2006). Bajo este contexto, el desarrollo infantil temprano (DIT) cobra especial relevancia. Distintos análisis han concluido que la inversión en educación preescolar no es solo un asunto de promoción de justicia y mayor equidad, sino también un incentivo a la productividad económica de los países

(Heckman, 2006). Así, tanto en el desarrollo de habilidades cognitivas como no cognitivas, la intervención temprana ha demostrado tener efectos positivos a nivel individual y de sociedad, los que contribuyen de manera general al progreso de las naciones. Sin embargo, la diferencia en cobertura y calidad entre países es abismante.

En Chile, el año 1990 la cobertura neta<sup>1</sup> de educación preescolar para niños entre 0 y 5 años alcanzaba solo a un 15.9% del total, mientras que el 2003 era de 29.9% y el 2009 37.4% (CASEN 1990, 2003 y 2009). Aun así, la cobertura nacional está por debajo de otros países de la región y aún más lejos de la de países desarrollados. Según la UNESCO (2011), Chile contaba con una tasa bruta<sup>2</sup> de escolarización preescolar para niños entre 3-5 años de 56% el 2008, mientras que Brasil tenía una tasa de 65%, Argentina 74% y países como Suiza y Noruega alcanzaban una cobertura completa al final de dicho año. Existe bastante consenso político y creciente consenso conceptual, que la baja cobertura genera múltiples desventajas tanto para el niño como para la familia, ya que además de apoyar al desarrollo infantil, la cobertura preescolar facilita la inserción laboral femenina, donde Chile cuenta con una de las tasas más bajas de la región (Encina & Martínez, 2009).

Otro gran problema de la situación chilena, es la heterogeneidad de cobertura. Mientras las zonas más desarrolladas contaban el 2003 con una cobertura de casi el 37%, los sectores rurales alcanzaban solo el 24% (CASEN, 2003). Además, la heterogeneidad

---

<sup>1</sup> Número total de alumnos de entre 0 y 5 años de edad que asisten a preescolar sobre la población total de 0 a 5 años de edad.

<sup>2</sup> Número total de alumnos que asisten a preescolar (de cualquier edad) sobre la población total de 3 a 5 años de edad.

también se da en términos de ingreso, donde el quintil más rico supera en casi un 20% en cobertura al quintil más pobre, como se muestra en la Figura 1-2.

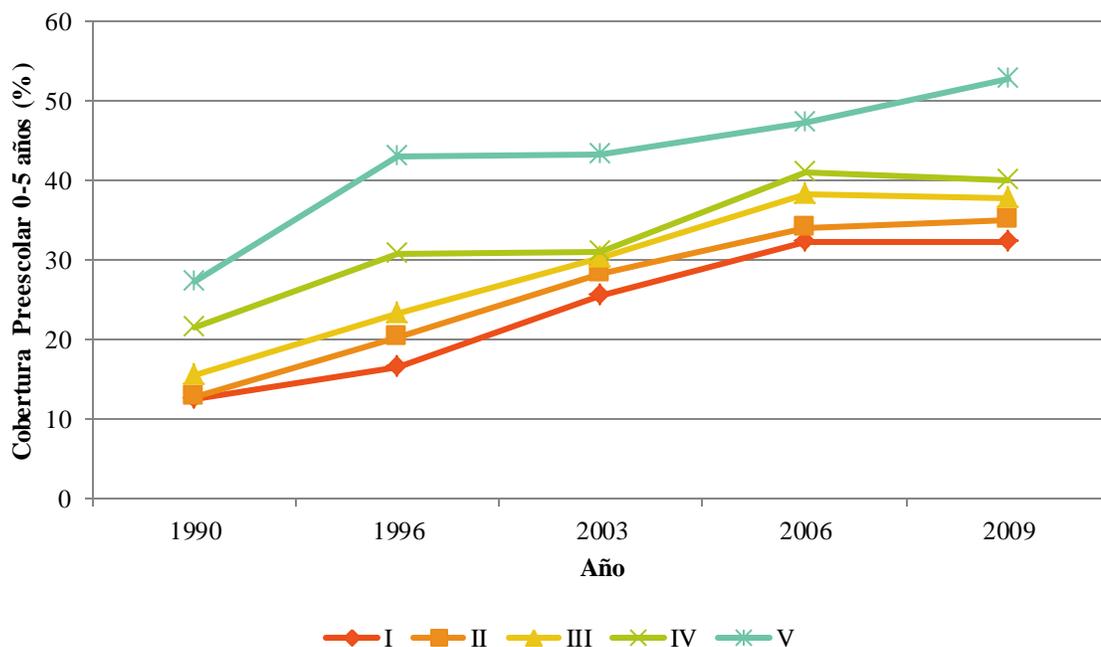


Figura 1-2: Evolución de cobertura preescolar por quintil de ingreso socioeconómico (Fuente: Elaboración propia en base a datos CASEN 1990, 1996, 2003, 2006 y 2009)

Para las familias más vulnerables de Chile, concentradas principalmente en el primer y segundo quintil de ingreso, las ofertas de preescolar están dadas principalmente por la Junta Nacional de Jardines Infantiles (JUNJI) y la Fundación Integra, ambos de carácter estatal, las que ofrecen servicios de sala cuna y jardín infantil de manera gratuita. La Fundación Integra recibe, al año 2012, a más de 70.000 niños, estando presente en 312 comunas con cerca de 1000 establecimientos en Chile. Además, se enfoca principalmente en la atención de familias que están dentro de los dos primeros

quintiles de ingreso socioeconómico, hogares monoparentales y madres adolescentes (Fundación Integra, 2011).

Si bien la importancia del desarrollo infantil temprano es transversal a todos los niveles socioeconómicos, un conjunto de estudios muestran que es aún más relevante para niños provenientes de estratos más vulnerables, debido a la falta de capital cultural provista en los hogares (Loeb, 2005; Heckman, 2006). De esta manera, la educación preescolar puede ser un factor clave para disminuir inequidades que comienzan en el seno de la familia, proveyendo estimulación a una edad en la cual el cerebro cuenta con una mayor disposición al aprendizaje y permitiendo que inversiones futuras sean más rentables (Heckman, 1999).

La mayoría de los estudios señalados que muestran que la estimulación temprana es relevante en el desarrollo de los niños, particularmente de los más vulnerables, estudian programas específicos, en la mayoría de los cuales hay una transferencia de recursos significativa y que poco tienen que ver con la mayoría de los programas relativamente masivos existentes en los países menos desarrollados, como es el caso de Chile. Siendo además claro que para un país con alta desigualdad y sustancial mayor escasez de recursos, donde la focalización de las políticas públicas es de la mayor importancia, es particularmente necesario contar estudios que evalúen el impacto de las políticas y en especial, de la educación preescolar. El objetivo de este trabajo es contribuir en esa dirección y detectar correlaciones que sugieren el sentido del efecto de la educación preescolar sobre las habilidades cognitivas de los niños en el tiempo y también sobre la capacidad de igualar condiciones que tiene la educación preescolar.

Específicamente, deseamos determinar si niños especialmente vulnerables que pasan por el sistema preescolar en Chile logran mayores niveles de desarrollo cognitivo, y si ese sistema consigue reducir brechas iniciales en ese desarrollo.

Con ese propósito, analizamos la relación de la educación preescolar del programa Integra, orientado a niños especialmente vulnerables, con el resultado de sus pruebas estandarizadas en la educación escolar. Para ello construimos un panel que sigue a niños desde el preescolar hasta la educación básica y que considera su exposición efectiva al programa preescolar.

El trabajo se estructura en cuatro secciones, aparte de esta introducción. La segunda sección describe la literatura sobre desarrollo infantil temprano (DIT); la tercera sección describe los datos y la metodología; la cuarta presenta los resultados y, la quinta sección concluye con propuestas de política pública.

## 2. LA LITERATURA

Si bien se reconocen múltiples beneficios asociados al DIT, la literatura es escasa respecto del impacto que éste tiene sobre habilidades cognitivas y no cognitivas. Ello, en parte, se debe a la falta de información, imposibilidad de seguimiento y dificultades metodológicas.

Los estudios realizados de manera más exhaustiva corresponden a “experimentos naturales” en los Estados Unidos, enfocados a programas intensivos y de pequeña escala para niños de escasos recursos, donde se comparan grupos aleatorios de tratamiento y control. Dos de las experiencias más relevantes son el *Perry Preschool Program* y el *Abecedarian Project*, en base de las cuales se han realizado diversos estudios. En general, se puede señalar que en ambos programas se muestra un efecto positivo sobre las habilidades cognitivas y no cognitivas en el tiempo, aunque de manera diferente.

Schweinheart (2005), Currie (2001) y Carneiro y Heckman (2003) observan que los participantes del *Perry Preschool Program* tuvieron un aumento en las habilidades cognitivas, medidas por el CI, durante los primeros años de primaria en comparación al grupo control, pero que dicho efecto no se mantuvo en el largo plazo. Sin embargo, en términos de habilidades no cognitivas, los niños que participaron del programa mostraron mayor propensión a la permanencia escolar, menores tasas de embarazo adolescente y detenciones por crímenes, respecto a sus pares que no participaron del programa.

En el caso del *Abecedarian Project*, Campbell y Ramney (2002) muestran que los participantes sí presentaron un efecto permanente en sus habilidades cognitivas. Mediante el estudio de cuatro grupos (alumnos que asistieron a programa preescolar y reforzamiento escolar en primaria, alumnos que participaron en solo uno de esos programas y el grupo control, que no participó en ninguna etapa), se muestra que persisten diferencias significativas en términos de CI y pruebas académicas a los 40 años de edad. Además, los participantes eran más propensos a asistir a la universidad y contaban con menores tasas de embarazo adolescente.

Debido a problemas metodológicos asociados a los inherentes sesgos de selección, los estudios de impacto de la educación preescolar en programas extensivos o de acceso universal se deben abordar con mayor cautela, dado que raramente cuentan con grupos aleatorios de control y tratamiento. Sin embargo, muchos de ellos obtienen conclusiones similares a los observados anteriormente con experimentos naturales. Loeb et al (2005) utilizan una encuesta representativa a nivel nacional en los Estados Unidos realizada a estudiantes de Kinder para encontrar la exposición óptima a programas preescolares (en duración e intensidad). Usando mínimos cuadrados ordinarios y *propensity score*, encuentran un efecto positivo de la asistencia a estos programas en habilidades de lectura y matemática, aunque efectos negativos en habilidades sociales. Sugieren también que la edad óptima para ingresar al sistema preescolar es entre los 2 y 3 años, hallando también que niños pertenecientes a estratos socioeconómicos más vulnerables ven ganancias académicas mayores de una asistencia intensiva (más de 30 horas a la semana).

Garces, Thomas y Currie (2002) analizan el efecto en el largo plazo del *Head Start*, un programa extensivo enfocado en niños vulnerables implementado en Estados Unidos en los años 1960. Mediante regresiones que controlaban por un conjunto extensivo de características y efectos fijos de la familia, encuentran efectos positivos en las tasas de asistencia a la universidad y salarios iniciales a los 20 años para participantes blancos. Mientras, los participantes afroamericanos presentaban menores tasas de arresto.

La evidencia en América Latina respecto de un conjunto de programas también sugiere efectos positivos. Mediante estimadores generalizados y marginales de *matching*, Behrman, Cheng y Todd (2004) encuentran que el Programa Integral de Desarrollo Infantil (PIDI) en Bolivia, tiene un efecto positivo en los resultados de test cognitivos, aunque ellos son altamente dependientes de la exposición de los niños al programa, comparado con niños de la zona que no asistieron al PIDI. Usando métodos similares de *matching*, Bernal, Fernández y Flórez (2009) evalúan el efecto del programa de Hogares Comunitarios de Bienestar (HCB) en Colombia, comparando a los participantes con niños de la zona que no asisten a los HCB. Las autoras encuentran un efecto positivo sobre el desarrollo cognitivo de los participantes con una exposición mayor a los 15 meses, los cuales se mantienen en el mediano plazo, según las pruebas estandarizadas en 5to básico. Además, el estudio reporta mejoras significativas en el desarrollo psicosocial.

En el caso de Argentina, Berlinski, Galiani y Gertler (2009) evalúan el impacto de la expansión en cobertura de la educación preescolar universal en las pruebas estandarizadas de educación primaria. El trabajo sugiere que un aumento de un cupo en

preescolar equivale al aumento en una matrícula, lo que les permite utilizar los nuevos puestos creados durante la expansión como una variable *proxy* de inscripción en el jardín infantil. Controlando por efectos fijos de provincia y de cohortes, que un año más de educación preescolar aumenta los resultados en las pruebas estandarizadas en 23% de la desviación estándar, además de tener efectos positivos en el autocontrol de los estudiantes.

Para Chile, Noboa y Urzúa (2012) estiman el efecto de asistencia a centros preescolares públicos, permitiendo tratamiento heterogéneo según jardín infantil. Utilizando una encuesta longitudinal de la Junta Nacional de Jardines Infantiles (JUNJI), hacen una evaluación cuasi-experimental, mediante modelos de autoselección y la construcción de modelos contrafactuales para cada período. Reportan efectos positivos en el desarrollo cognitivo y motor de los participantes, los cuales son fuertemente dependientes de la edad del niño y su exposición al programa. En línea con estudios previos, se reportan efectos negativos en la interacción con adultos, lo que los autores sugieren podría deberse a problemas en la calidad de los programas.

Contreras, Herrera y Leyton (2008) estiman el efecto de la inscripción en el preescolar en el logro educacional de los estudiantes, medido como puntaje de pruebas estandarizadas SIMCE en segundo medio<sup>3</sup>. Para abordar problemas de sesgo de selección, los autores utilizan *propensity score matching* con dos especificaciones diferentes para mayor robustez, encontrando que existe un efecto positivo asociado a la

---

<sup>3</sup> Equivale al décimo año de enseñanza obligatoria (de 12 años), donde los alumnos deben tener alrededor de 16 años.

asistencia al preescolar sobre el puntaje de entre 14% y 28% de la desviación estándar para la prueba de lenguaje y entre 22% y 35% de la desviación estándar para la prueba de matemática, según la especificación y el método de *matching* utilizado.

En este trabajo complementamos el aporte previo analizando el efecto del preescolar en programas extensivos en Chile en el corto plazo y mediano plazo, controlando por habilidades iniciales del estudiante. Separar el efecto del sistema preescolar del impacto que tienen las habilidades del alumno en el rendimiento académico es fundamental para poder comprender de mejor manera la importancia de la educación temprana, y enfatizamos también la relevancia de disponer de una variable más precisa de exposición efectiva a programas preescolares.

### 3. DATOS Y METODOLOGÍA

Nos enfocamos en el impacto de la educación preescolar del Programa Integra, orientado a niños vulnerables, que son el foco más relevante desde la perspectiva de la política pública, particularmente en lo que se refiere al desafío de cerrar las brechas educativas que reportan diversos estudios en Chile (García & Paredes, 2010). Nuestro propósito es evaluar el efecto del desarrollo infantil temprano en las habilidades cognitivas en el corto plazo, esto es, mientras el niño aún asiste al preescolar, y también su efecto en el mediano plazo, 4 a 5 años después de recibir la educación preescolar cuando el alumno cursa cuarto básico.

#### 3.1 Un Modelo Ad hoc a la disposición de datos

La ecuación (1) muestra una forma reducida de la relación a estimar entre habilidades cognitivas en el período  $t$  ( $hc_t$ ), y las variables que la determinan; ambientales ( $Amb$ ), familiares ( $Fam$ ) y factores individuales ( $Ind$ ). Las variables individuales incluyen características físicas del niño, habilidades innatas y si es que recibió algún tipo de apoyo especial en su desarrollo temprano, como podría ser la asistencia a un centro preescolar.

$$hc_t = f_t(Amb, Fam, Ind) \quad (1)$$

Como métrica de las habilidades cognitivas, disponemos de los resultados de pruebas reportadas durante la educación preescolar, y el rendimiento en pruebas estandarizadas de lenguaje en la educación escolar. Las variables dependientes utilizadas

corresponden a puntajes en pruebas estandarizadas en dos momentos del tiempo. A nivel preescolar, usamos el Test de Desarrollo Psicomotor (TEPSI) de Lenguaje (TL), el cual se realiza a niños entre 2 y 5 años de edad. Esta prueba, ajustada por edad con niveles de confiabilidad y validez probados, se realiza a una muestra aleatoria de alumnos de la Fundación Integra. En el caso de la educación primaria utilizamos el SIMCE de lenguaje (SL), que es la prueba que mayor coherencia tiene respecto del área que se midió en el nivel preescolar a través del test señalado.

Usamos las bases de datos de la Fundación Integra (2001 y 2003) y datos de los cuestionarios de la prueba SIMCE de cuarto básico (2007 al 2010). Con información de identificación de los estudiantes, hacemos un *match* por niño desde la educación preescolar entre 2001 y 2003, a la escolar años después. Los controles ambientales y familiares corresponden a las características reportadas en el cuestionario de padres del SIMCE y proveniente de información acerca de los niños en su etapa preescolar, como es el caso del peso, talla y exposición al programa. Solo se utilizaron para la muestra los alumnos que estaban dentro del rango normal de edad al momento de rendir el SIMCE de cuarto básico (9 o 10 años).

El impacto que tiene la educación preescolar en las habilidades cognitivas en el corto plazo, o dentro del mismo sistema, se especifica en la ecuación (2). En este modelo reducido, se verifica el efecto sobre el puntaje TEPSI del niño, que tiene la asistencia al preescolar ( $TT$ ), condicional a un set de factores socioeconómicos  $X_I$

$$TL = f_{CP}(TT, X_I) \quad (2)$$

El efecto en el mediano plazo de la educación preescolar y que se manifiesta en el puntaje SIMCE de lenguaje en cuarto básico es considerablemente más complejo de interpretar y distinguir. Para ello, separamos el efecto del tratamiento de las habilidades iniciales del estudiante, controlando por el puntaje TEPSI (TL) obtenido, donde centrándonos en el efecto del tratamiento, se procura controlar por factores socioeconómicos  $X_2$  y por las habilidades cognitivas totales ( $hc_T$ ) que el niño tiene antes de ingresar al sistema escolar primario, medidas como su puntaje en el TEPSI de Lenguaje (TL).

$$SL = f_{MP}(TT, TL, X_2) \quad (3)$$

Como desarrollamos, la interpretación del efecto del tratamiento (TT) en dicha ecuación tiene relación con el efecto de habilidades cognitivas no medidas, en vez de una interpretación directa del efecto de la exposición al preescolar. Así, podemos separar conceptualmente las habilidades cognitivas iniciales de un estudiante de la siguiente manera:

$$hc_T = hc_f + hc_a + hc_g + hc_{pe} = hc_o + hc_{pe} \quad (4)$$

donde  $hc_T$  corresponde a las habilidades iniciales totales de un estudiante antes de ingresar al sistema escolar,  $hc_f$  corresponde a las habilidades iniciales entregadas por la familia,  $hc_a$  mide las habilidades entregadas por factores ambientales,  $hc_g$  a factores genéticos y, finalmente,  $hc_{pe}$  corresponde a las habilidades iniciales adquiridas en el preescolar. Consideremos que las primeras tres habilidades iniciales pueden agruparse en  $hc_o$ . Dado el modelo de corto plazo, sabemos que la asistencia a preescolar tiene un

efecto en  $hc_T$ , medidas como el puntaje TEPSI (TL). Sin embargo, no tenemos medidas observables o *proxys* para separar  $hc_o$  y  $hc_{pe}$ .

Dado que el impacto del tratamiento considera iguales habilidades cognitivas totales (TL) y las mismas características socioeconómicas ( $X_2$ ), el efecto del tratamiento en el modelo de mediano plazo lo que captura realmente es un mayor nivel de habilidades cognitivas no entregadas por el preescolar  $hc_o$  para alumnos que no asisten a preescolar versus los que tienen una mayor exposición a este sistema. Consecuentemente, si dos alumnos alcanzaron igual nivel de  $hc_T$ , el primero con alta exposición a preescolar (lo que afecta positivamente  $hc_{pe}$  según la literatura) y el otro con baja exposición a educación temprana, ello puede interpretarse como un indicador que el segundo tiene  $hc_o$  mayor que el primero, lo que afecta positivamente su desempeño en pruebas estandarizadas posteriores. Luego, el efecto del tratamiento (TT) estaría capturando el impacto de otras habilidades cognitivas, mientras que el efecto de la exposición estaría impactando al SIMCE de Lenguaje (SL) a través de las habilidades cognitivas totales (TL).

### **3.2 Identificación de efectos**

Para poder determinar el efecto que tiene la asistencia al preescolar en el desarrollo cognitivo, necesitamos contar con dos grupos de niños, idealmente iguales, que sólo se diferencien por el hecho de haber o no asistido al jardín infantil o similar y otros. Como no contamos con un experimento natural (dos grupos aleatorios), ello nos lleva a construir un grupo de pseudo control.

La dificultad es mayor debido a que tampoco contamos con los resultados del TEPSI para alumnos que no se inscribieron en algún establecimiento de la Fundación Integra. Por ello, y siguiendo a Behrman (2004) y Bernal (2009), usamos como grupo control y de tratamiento a niños con mayor y menor grado de exposición efectiva al programa preescolar. Así, buscamos identificar a niños que, estando matriculados en Integra, hayan tenido una exposición baja al programa, de modo que sus habilidades cognitivas no hayan sido alteradas por la asistencia al preescolar. Mientras, el grupo de tratamiento lo definimos compuesto por niños con una alta exposición al programa, de manera que sus habilidades cognitivas sí pudieran haber sido afectadas por el tratamiento.

Utilizando el tiempo desde que el niño se matricula en la Fundación Integra y la fecha en que rinde el TEPSI, calculamos el tiempo de exposición bruta que tuvo al programa, esto es, bajo un supuesto de asistencia del 100%. Sin embargo, dadas las altas tasas de inasistencia y la varianza en ella, es razonable pensar que la exposición bruta al programa no refleja cuan expuesto a la educación estuvo el niño. Por ello, multiplicamos la exposición bruta por la asistencia promedio mensual del niño, lo que permite disponer de un indicador más adecuado de la exposición efectiva que ha tenido el niño al jardín infantil.

Consideramos la cola del 25% inferior de la distribución de exposición efectiva como grupo control y la del 25% superior como grupo de tratamiento. Esto se traduce en que los alumnos del grupo control tienen una exposición menor a 6 meses y los del grupo tratamiento mayor a 15 meses. La exposición así obtenida para el grupo bajo

tratamiento es coherente con la literatura respecto del tiempo mínimo necesario para que los programas preescolares tengan efecto.

En la Tabla 3-1 se muestra la descripción de características socioeconómicas y de rendimiento para el grupo control (no tratado) y el grupo de tratamiento. Observamos que no existen diferencias estadísticamente significativas en sus características socioeconómicas, pero sí hay diferencias significativas en características individuales y de ruralidad. Tales diferencias son relevantes de considerar al momento de realizar inferencias, dado que es necesario contar con dos grupos comparables o realizar ajustes econométricos para que así sea. Concretamente, el procedimiento para tratar el eventual sesgo que estas diferencias pueden producir en la evaluación del impacto consiste en abordar la fuente de endogeneidad mediante tres métodos econométricos: control de variable omitida, reponderación de la muestra y *matching* a través de *propensity score*, las que son abordadas en la sección 3.4.1, 3.4.2 y 3.4.3 respectivamente.

Tabla 3-1 – Descripción de datos socioeconómicos y de rendimiento para alumnos

Integra según tratamiento

	No Tratado		Tratado		Total	
	Media	DS	Media	DS	Media	DS
<i>Educación Madre (años)</i>	10.7	3.3	10.7	3.3	10.7	3.3
<i>Educación Padre (años)*</i>	10.6	3.4	10.3	3.5	10.5	3.4
<i>Ingreso Familiar (CL\$)</i>	\$250,446	\$242,837	\$252,462	\$260,645	\$251,315	\$250,614
<i>Edad Niño Ingreso Integra (meses)**</i>	40.3	7.5	26.4	7.9	34.4	10.3
<i>Edad Niño TEPSI (meses)**</i>	46.5	7.4	51.7	5.6	48.7	7.2
<i>Peso Niño (kg)**</i>	16.9	2.7	17.8	2.7	17.3	2.7
<i>Talla Niño (cm)**</i>	100.5	6.8	103.0	6.0	101.6	6.6
<i>Exposición efectiva a preescolar (meses)**</i>	3.8	1.7	22.9	6.9	12.0	10.5
<i>Establecimiento particular (%)</i>	46.5%	49.9%	44.8%	49.8%	45.6%	49.8%
<i>Establecimiento rural (%)**</i>	10.0%	30.3%	7.5%	26.3%	8.9%	28.5%
<i>Región Metropolitana (%)**</i>	9.8%	29.8%	14.8%	35.5%	12.0%	32.5%
<i>TEPSI Lenguaje (TL)**</i>	48.77	9.69	52.58	8.13	50.41	9.24
<i>SIMCE Lenguaje (SL)</i>	251.83	53.39	250.62	52.08	251.31	52.82
<b>N Total</b>	<b>1232</b>		<b>934</b>		<b>2166</b>	

\*: Diferencia significativa al 5% entre T y NT \*\*: Diferencia significativa al 1% entre T y NT

Un segundo problema metodológico surge de la representatividad de la muestra a nivel nacional, dado que solo contamos con niños inscritos en el sistema preescolar. Dada la aún baja cobertura preescolar, especialmente en sectores más vulnerables, la

muestra puede no representar a la población. En efecto, para que los participantes estuvieran dentro del grupo de tratamiento, los padres o cuidadores tuvieron que tomar dos decisiones: i) Inscribir al niño en la educación preescolar, y ii) que el niño tenga una exposición efectiva al programa  $\mu > \bar{\mu}$

Dado que ninguna de las dos decisiones parece ser aleatoria, la ausencia de un tratamiento a la endogeneidad podría generar resultados sesgados. Esto es, podríamos atribuirle un efecto al tratamiento en circunstancias que el efecto está, por ejemplo, asociado al hecho que el niño tenga padres más motivados que lo hacen asistir, o que tenga una salud deteriorada que lo hace asistir menos, ambos factores posiblemente relacionados con el desarrollo cognitivo. Por ello, es relevante tratar el problema de endogeneidad que podría llevar a resultados sesgados para poder solucionarlo y obtener conclusiones robustas. La primera decisión de inscribir al niño en el preescolar es determinante para la segunda, pero como se observa en los datos, dadas las altas tasas de inasistencia, matricular al niño en el jardín infantil no es equivalente a que tenga una exposición efectiva relevante.

### **3.3 Inscripción del niño en el sistema preescolar**

El problema asociado a que los niños tratados y el grupo control no sean exógenamente determinados puede generar un sesgo de selección. La decisión de inscribir a un niño en la educación preescolar está determinada por variables familiares y ambientales, como puede ser la vulnerabilidad y el acceso a centros preescolares, algunas no observables. Particularmente para familias de estratos socioeconómicos más vulnerables, la elección

de enviar o no a sus hijos a un establecimiento preescolar gratuito (JUNJI o Integra, principalmente) pudiera estar relacionada con la existencia de redes, interés de los padres, acceso a estimulación y, por ende, correlacionar con medidas de habilidad futura.

Para abordar este problema, seguimos a Heckman (1979), quien sugiere tratar el problema como uno de omisión de variable relevante. Así, mediante un modelo *probit*, estimamos la probabilidad de inscribirse en el preescolar en función de variables observables, y de ella, derivamos el inverso de la razón de Mill, o regresor *Heckit*

$\lambda = \frac{\phi(pe_i)}{\Phi(pe_i)}$ , donde  $pe_i$  representa la probabilidad estimada que un niño  $i$  se inscriba en el preescolar.

Específicamente a partir de la base de datos del SIMCE de cuarto básico podemos identificar a alumnos que asistieron al preescolar entre los 2 y los 5 años y los que no. Dentro de los alumnos que asistieron al preescolar, pueden haber asistido a un jardín JUNJI, Integra u a otro establecimiento. Considerando solo a aquellos niños provenientes de estratos socioeconómicos vulnerables<sup>4</sup>, el foco de este trabajo, podemos generar una variable binaria *preescolar* que tome el valor 1 si es que el niño asistió a algún establecimiento preescolar (Integra, JUNJI u otro) entre los 2 y los 5 años y 0 en caso contrario.

Suponemos que la decisión de llevar o no a un niño al preescolar (variable *preescolar*) depende de la oferta de jardines infantiles cercana al hogar y de la

---

<sup>4</sup> Familias de ingreso familiar  $I \leq \bar{I}$ , donde  $\bar{I}$  equivale a CL\$450000 mensuales

disponibilidad de un cuidador alternativo en el hogar (ya sea padre, madre u otro), la que está correlacionada con el trabajo de los padres y características socioeconómicas. Para medir la oferta de jardines infantiles, utilizamos como variable *proxy* la ruralidad de la comuna donde vive el niño y el ingreso familiar para controlar por factores socioeconómicos que pueden influir en la decisión, lo que deriva en (5)

$$Pr(\text{preescolar} = 1) = f(\text{ruralidad}, \text{ingreso}) \quad (5)$$

Mediante el modelo *probit*, aplicado sobre todas las observaciones de niños que rindieron el SIMCE de cuarto básico con ingresos familiares menores a \$450.000 mensuales<sup>5</sup>, obtenemos la probabilidad  $pe_i$  de asistir al preescolar.

Si bien queremos obtener el efecto del preescolar para los niños vulnerables en Chile, solo tenemos una muestra representativa de los niños inscritos en la Fundación Integra y no contamos con una muestra de todas las instituciones preescolares. Sin embargo, cabe recordar que la oferta fundamental de jardines infantiles para familias vulnerables se concentra en los establecimientos JUNJI e Integra, es decir, de dependencia estatal. Si no hay diferencias entre los establecimientos JUNJI e Integra, la familia será indiferente frente a la decisión de inscribir a los niños en uno u otro jardín infantil y los resultados obtenidos aplicando el ajuste del método de Heckman serán extrapolables a todos los niños vulnerables que asistan a alguno de estos establecimientos.

---

<sup>5</sup> Se utiliza este corte debido a que el 99% de las familias pertenecientes a la Fundación Integra, presentan ingresos familiares menores a este monto.

### **3.4 Asistencia del niño al preescolar**

Un segundo problema asociado a la medición de variables y también con endogeneidad dice relación con la exposición efectiva de los niños a los programas prescolares después de adoptar la decisión de inscribir al estudiante en dichos programas. Hay alumnos que estando inscritos en el jardín infantil presentan altas tasas de inasistencia, por lo que matricularse en el preescolar no es equivalente a asistir. Si las inasistencias se deben a factores que se correlacionan con la habilidad, como enfermedad o falta de interés de los padres, entonces el efecto del tratamiento estará sesgado. En el caso que nos compete, consideramos tres factores que pueden aumentar la inasistencia: edad del niño, estado de salud y distancia de traslado al jardín infantil. La edad de ingreso al jardín infantil puede ser también una variable relevante en la asistencia, dado que niños pequeños son más propensos a faltar, como muestran los datos de asistencia mensual de la Fundación Integra para los años 2008 a 2010. Por esta razón también es relevante utilizar esta variable como control.

Según datos de la Fundación Integra (Fundación Integra, 2011), cerca del 70% de las inasistencias corresponden a problemas de salud. Además, según la descripción estadística de la Tabla 3-1, los grupos considerados se diferencian en peso de los niños y talla. Es razonable suponer que niños con menor peso son más propensos a enfermedades. Por esto es esperable que el grupo control tenga, en promedio y controlando por edad, niños de menor peso y podría considerarse esta variable como una medida del estado de salud.

La distancia recorrida hasta el jardín infantil puede ser una razón que aumente la inasistencia del niño al preescolar. Dado que en zonas rurales y en regiones la oferta de jardines infantiles es menor, ello sugiere un mayor tiempo de traslado entre el hogar y el establecimiento preescolar, lo que fomentaría más ausentismo. Para abordar este problema de endogeneidad en la exposición efectiva al programa, utilizamos tres métodos, de modo de chequear consistencia y robustez de los resultados al método usado, los que se exponen en las secciones siguientes.

A pesar de que se busca resolver los numerosos problemas metodológicos que surgen al no contar con experimentos naturales, existen limitaciones de estas soluciones. Principalmente, debido a la falta de información socioeconómica de los niños durante su etapa preescolar, no se pueden utilizar datos contemporáneos para el modelo de corto plazo. Así, se asume que las características socioeconómicas de los estudiantes no han cambiado drásticamente en un periodo de 5 a 7 años o, si es que lo han hecho, ha sido igual para todos. Otra limitación es la incapacidad de identificar si existe un cuidador alternativo o si los niños pertenecen a hogares monoparentales en la época preescolar, lo que haría más robusto el modelo de sesgo de selección por ausentismo. A los tres métodos presentados a continuación se incorpora el método de Heckman abordado en la sección 3.3, para resolver el sesgo de inscripción en preescolar.

### 3.4.1 Mínimos cuadrados ordinarios (MCO) controlando por estado de salud

Para cuantificar el efecto de la educación preescolar sobre las habilidades cognitivas ( $hc_t$ ), además de controles ambientales, familiares e individuales ( $X_t$ ), introducimos una variable binaria en el modelo que toma valor 1 si es que el niño es tratado y 0 en caso contrario, de modo que su coeficiente capture el efecto de asistir a programas preescolares (TT).

Sin embargo, como la asistencia a preescolar está relacionada con la edad de ingreso del alumno y su estado de salud y éstas, a su vez, están correlacionadas con la habilidad del niño, también se debe controlar por dichas variables, como en (6):

$$hc_t = \beta_0 + \beta_1 X_t + \beta_2 preescolar + \beta_3 salud + \beta_4 EdadIngreso + \beta_5 \lambda + \varepsilon \quad (6)$$

Concretamente, para nuestros modelos empíricos usamos como variable *proxy* del estado de salud el peso del alumno en la edad preescolar (*peso*), como se muestra en la ecuación (7) y (8). Los controles  $X_t$  del modelo de corto plazo de la ecuación (7) incluyen los años de escolaridad del padre y la madre, el ingreso familiar y la edad del niño al momento de rendir el examen. En el caso del modelo de mediano plazo (ecuación (8)), los controles  $X_2$  incluyen solo la escolaridad de los padres y el ingreso familiar. Ambos modelos también incluyen el regresor *Heckit*  $\lambda$  para resolver el sesgo de selección de inscripción al preescolar. En ambos modelos la variable *preescolar* equivale a la variable binaria de tratamiento (TT) y en el caso del modelo de mediano

plazo (ecuación (8)), controlamos también por las habilidades iniciales del estudiante, medidas como su puntaje TEPSI en el preescolar (TL).

$$TL = \alpha_0 + \alpha_1 TT + \alpha_2 X_1 + \alpha_3 peso + \alpha_4 EdadIngreso + \alpha_5 \lambda + \mu \quad (7)$$

$$SL = \beta_0 + \beta_1 TT + \beta_2 TL + \beta_3 X_2 + \beta_4 peso + \beta_5 EdadIngreso + \beta_5 \lambda + \eta \quad (8)$$

### 3.4.2 Estimadores ponderados en base de probabilidad inversa (*Inverse Probability-Weighted Estimators*)

Para aislar el efecto de las diferencias en características observadas de la muestra, seguimos a Curtis et al. (2007). Este método estima una regresión que asigna diferentes ponderaciones a cada observación, de modo que las medias ponderadas de las diferentes características sean similares para el grupo tratamiento y el grupo control, como el caso de ruralidad, peso y edad de ingreso del niño (ver Tabla 3-1).

En una primera etapa calculamos la probabilidad  $p_i$  de pertenecer al grupo de tratamiento ( $TT_i=1$ ) versus al grupo de control ( $TT_i=0$ ) mediante un modelo *probit* que controla por peso y talla del niño, además de su edad (en meses) al momento de ingresar al preescolar, como en (9).

$$TT_i = f(Peso_i, Talla_i, EdadIngreso_i) \quad (9)$$

Luego, se asigna una ponderación equivalente al inverso de la probabilidad  $\frac{1}{p_i}$  para los niños que pertenecen al grupo de tratamiento y  $\frac{1}{1-p_i}$  para los del grupo control. De esta manera, la observación que corresponde a un alumno que no participó del

tratamiento (i.e. tuvo baja exposición) pero que tenía una alta probabilidad de ser tratado debido a sus características de peso, talla y edad de ingreso a la fundación, recibe un mayor peso debido a su similitud con las observaciones del grupo de tratamiento. Lo mismo aplica para el caso inverso.

Como segunda etapa, se realizan las regresiones mediante MCO, utilizando los pesos de probabilidad inversa obtenidos anteriormente, como se ve en las ecuaciones (10) y (11). Los controles  $X_1$  y  $X_2$  son los mismos utilizados anteriormente.

$$TL = \alpha_0 + \alpha_1 TT + \alpha_2 X_1 + \alpha_3 \lambda + \mu \quad (10)$$

$$SL = \beta_0 + \beta_1 TT + \beta_2 TL + \beta_3 X_2 + \beta_7 \lambda + \eta \quad (11)$$

### 3.4.3 *Matching por medio de propensity score (PS)*

Siguiendo a Behrman et al. (2004) y a Noboa & Urzúa (2012), este método para observaciones del grupo de tratamiento con observaciones del grupo control que tengan características socioeconómicas e individuales similares, dado que existen diferencias entre ambos grupos en características individuales de los niños (ver Tabla 3-1). Uno de los supuestos clave de este método es que la selección se realiza sobre variables observables, de modo que si se controla por la probabilidad de tratamiento, los niños del grupo control se comportan de la misma forma que los del grupo de tratamiento.

Para ello, calculamos en primer lugar el *propensity score* (PS) de cada observación, que corresponde a la probabilidad de ser tratado controlando por factores socioeconómicos e individuales del niño, como muestra la ecuación (12):

$$\Pr(\text{preescolar} = 1) = f(\text{Fam}, \text{Ind}) \quad (12)$$

Los factores familiares (*Fam*) e individuales (*Ind*) corresponden a la educación de los padres, ingreso familiar, peso y talla del niño, además de la edad que tenía al momento de matricularse en la Fundación Integra.

En base a dicho PS, se define un vecindario para cada observación del grupo de tratamiento, definiendo como “vecinos cercanos” a aquellas observaciones del grupo control que estén suficientemente cerca en términos del valor de PS. De esta manera, se obtienen las diferencias en el TL y SL entre un grupo control y de tratamiento que tienen probabilidades similares de ser tratados.

#### 4. RESULTADOS

Los efectos que nos interesa constatar se manifiestan en eventuales diferencias en puntaje TEPSI y SIMCE entre el grupo control y de tratamiento. La Figura 4-1 muestra las distribuciones no condicionales de ambas pruebas. Si se observa la distribución de densidad de TL, existe una diferencia positiva entre los niños que obtuvieron mayor exposición a educación preescolar versus los que tuvieron menos (Figura 4-1 (a)). Sin embargo, dicha diferencia no se mantiene en SL, donde ambas distribuciones son similares (Figura 4-1 (b)).

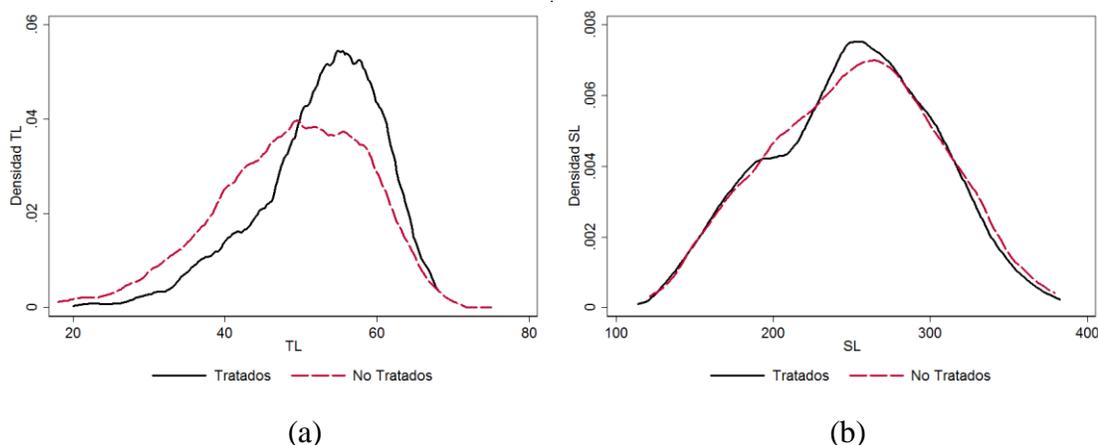


Figura 4-1 – (a) Distribución Puntaje TEPSI Lenguaje según tratamiento (b)

#### Distribución Puntaje SIMCE Lenguaje según tratamiento

Las figuras previas muestran distribuciones no condicionadas, que sugieren un efecto positivo de la asistencia al preescolar en el corto plazo pero no en el mediano plazo. Para identificar el efecto del tratamiento, sin embargo, se requiere aislar las características ambientales, familiares e individuales.

#### 4.1 Modelos de Corto Plazo

La Tabla 4-1 muestra los resultados de la aplicación de los tres métodos para determinar el efecto del tratamiento sobre las habilidades cognitivas en el mismo sistema preescolar y que denominamos efecto de corto plazo. El efecto del preescolar según los diferentes métodos varía entre 1.5 puntos del TEPSI Lenguaje y 6.5 puntos, donde el efecto es significativo al 5% de significancia en todos los modelos (la desviación estándar del TEPSI (TL) para la muestra es de 9.23 puntos). Como coeficientes estandarizados, la diferencia de magnitudes está entre 8% y 36% de la varianza explicada. Todos los métodos utilizados fueron sometidos a testeos de covariables para verificar la robustez de los resultados frente a cambios de las estructuras de los modelos, obteniendo resultados similares para todas las pruebas.

El método que utiliza el peso del niño como control del estado de salud, muestra un efecto positivo y estadísticamente significativo de una mayor exposición al jardín infantil. Los demás regresores tienen el signo esperado y son estadísticamente significativos al 5%, a excepción del ingreso familiar que no es estadísticamente diferente a cero.

En el caso de los estimadores obtenidos con pesos de probabilidad inversa, la primera etapa con la que se obtiene la probabilidad de ser tratado, muestra la dirección de efectos esperada, como muestra la Tabla A-1. Frente a un mayor peso y talla, la probabilidad de pertenecer al grupo de tratamiento aumenta, mientras que a menor edad de ingreso a Integra, la probabilidad de ser tratado disminuye. En el caso del modelo a

corto plazo, se observa un efecto positivo y estadísticamente significativo de una mayor exposición al programa preescolar de alrededor de 4.4 puntos, es decir, cerca de 0.5 desviaciones estándar. La magnitud del efecto es mayor que en el caso anterior, probablemente debido a que en el modelo que utiliza la variable peso como *proxy* del estado de salud hay una subestimación del efecto, ya que las muestras del grupo control y tratamiento no están balanceadas.

Finalmente, el método de *matching* mediante *propensity score* (PS) sugiere resultados similares. Para balancear las diferentes muestras en el corto y mediano plazo, se utilizan dos PS diferentes. El primero, para el modelo de corto plazo, utiliza como variables independientes el peso y la talla del niño, además de la edad de ingreso a Integra. Se consideran también variables socioeconómicas como la educación de los padres y el ingreso familiar. Para realizar el *matching*, se utiliza el método de kernel con un emparejamiento con remplazo de los vecinos más cercanos en un radio de 0.01. Mediante este método se puede calcular el efecto promedio del tratamiento en los tratados (ATT) para las observaciones pareadas, donde el grupo de tratamiento presenta 6.48 puntos más en el TL que el grupo control, siendo una diferencia estadísticamente significativa.

Si bien los tres métodos se diferencian en magnitud, hay coherencia en ellos respecto del efecto de corto plazo. En los tres modelos se encuentra un efecto significativo y positivo en el desarrollo infantil temprano, medido como puntaje del TEPSI Lenguaje. En base a la literatura y el tratamiento de los problemas de endogeneidad, el modelo más robusto parece ser el de *propensity score*, resultado que

indica un efecto del preescolar sobre las habilidades cognitivas del niño en el corto plazo que no solo estadísticamente significativo, sino también sustantivamente relevante en magnitud.

Tabla 4-1 - Resultados para modelo de corto plazo según método de control por estado de salud, estimadores de peso con probabilidad inversa y *matching* con *propensity score*

TL	Control por Estados de Salud		Peso con Prob. Inv.		PS	
	Coef.	Err. Est.	Coef.	Err. Est. Robusto	Dif.	Err. Est.
TT	1.643*	0.815	4.354**	0.608	6.48**	2.27
EdPadre	0.251**	0.068	0.249 <sup>†</sup>	0.131		
EdMadre	0.228**	0.071	0.42**	0.124		
Ingreso Familiar (en \$10.00 del 2011)	-0.016	0.011	-0.072**	0.018		
Peso	0.200*	0.081				
EdadIngreso	-0.090*	0.038				
Edad	0.156**	0.046	0.076	0.065		
lambda	-23.142*	9.358	-64.760**	18.427		
constante	46.042**	4.327	63.984**	9.460		
<b>R<sup>2</sup> Ajustado</b>	<b>0.0828</b>		<b>0.2529</b>			
<b>Observaciones</b>	<b>2130</b>		<b>2130</b>			

<sup>†</sup>: significancia estadística al 10% \*: significancia estadística al 5% \*\*: significancia estadística al 1%

## 4.2 Modelos de Mediano Plazo

En la Tabla 4-2 se muestran los resultados de la aplicación de los modelos para la medición del efecto de la asistencia al preescolar sobre las habilidades cognitivas del estudiante en el mediano plazo. Las habilidades cognitivas en el mediano plazo son medidas como el puntaje en la prueba SIMCE de 4to básico. En este caso, utilizamos como control dentro de los tres modelos las habilidades iniciales del alumno antes de ingresar al sistema escolar, medidas como su puntaje obtenido en la prueba TEPSI de lenguaje.

En los tres métodos utilizados los resultados son coherentes entre sí, donde se observa un efecto positivo y estadísticamente significativo de las habilidades iniciales del estudiante. El efecto negativo y significativo asociado a la variable de tratamiento debe ser interpretado como señalamos previamente, es decir, como el efecto de las habilidades previas al sistema preescolar, o no entregadas en él ( $hc_o$ ), y que podemos asociar a las habilidades previas entregadas por la familia, el ambiente y los factores genéticos.

En el caso del método que controla directamente por estado de salud, se encuentra un efecto positivo asociado a las habilidades iniciales, medidas como puntaje del TEPSI Lenguaje (TL), donde el efecto alcanza una magnitud de 1.3 puntos (0.223 desviaciones estándar como coeficiente estandarizado) (Tabla 4-2).

Para la estimación con pesos de probabilidad inversa, observamos un efecto positivo y significativo de las habilidades iniciales sobre el SIMCE de Lenguaje (SL) de

magnitud similar a la encontrada en el primer modelo, de 1.3 puntos (0.227 desviaciones estándar como coeficiente estandarizado).

Para el método de *propensity score*, se utiliza el mismo método, solo que el PS se calcula sobre el peso y talla del niño, además de la educación de los padres, el ingreso familiar, la edad al momento de matricularse a Integra y las habilidades iniciales medidas como puntaje del TL. Al igual que en los casos anteriores, se encuentra una diferencia negativa entre el grupo de tratamiento y el de control que es estadísticamente significativa, y de magnitud -16.15 puntos del SL.

Tabla 4-2 - Resultados para modelo de mediano plazo según método de control por estado de salud, estimadores de peso con probabilidad inversa y *matching* con *propensity score*

SL	Control por Estados de Salud		Peso con Prob. Inv.		PS	
	Coef.	Err. Est.	Coef.	Err. Est. Robusto	Dif.	Err. Est.
TT	-5.860**	2.221	-25.093**	7.973	-16.150**	-1.260
TL	1.304**	0.121	1.331**	0.220		
EdPadre	1.255**	0.382	2.157*	0.870		
EdMadre	1.761**	0.394	1.999*	0.782		
Ingreso Familiar (en CL\$10.000 del 2011)	0.158**	0.064	0.271**	0.093		
Lambda	15.627	52.464	279.134**	84.275		
constante	146.040**	23.343	31.595	50.923		
<b>R<sup>2</sup> ajustado</b>	<b>0.1078</b>		<b>0.1769</b>			
<b>Observaciones</b>	<b>2177</b>		<b>2110</b>			

†: significancia estadística al 10% \*: significancia estadística al 5% \*\*: significancia estadística al 1%

Si bien las magnitudes encontradas a través de los diferentes métodos difieren, las conclusiones son coherentes entre sí. Hay un efecto positivo asociado a una mayor exposición al sistema preescolar en cuanto a las habilidades iniciales totales, y también un efecto positivo de las habilidades cognitivas no observadas. Luego, el efecto de la exposición al preescolar en el mediano plazo se da a través de un aumento de las

habilidades cognitivas del alumno en el preescolar, donde el puntaje SIMCE del alumno (SL) depende del puntaje TEPSI (TL).

$$\frac{\partial SL}{\partial TT} = \frac{\partial SL}{\partial TL} \cdot \frac{\partial TL}{\partial TT} \quad (13)$$

De esta forma, el efecto del preescolar sobre el SIMCE varía entre 2.14 puntos y 5.80 puntos, que equivale a 4.3% y 11.6% de la desviación estándar del SIMCE Lenguaje<sup>6</sup>.

---

<sup>6</sup> La desviación estándar del SIMCE Lenguaje a nivel nacional es de 50 puntos.

## 5. CONCLUSIONES

Utilizando un panel que recopila información detallada de características individuales del niño durante el preescolar, además de su asistencia promedio, medimos el efecto que tiene el preescolar sobre las habilidades cognitivas de niños de alta vulnerabilidad en Chile. Los resultados indican que la educación preescolar tiene un efecto positivo en dichas habilidades, medidas como resultados de pruebas estandarizadas de lenguaje. En el corto plazo, hay un impacto positivo relacionado directamente con la asistencia a preescolar. En el mediano plazo el efecto del preescolar está asociado a un aumento de las habilidades cognitivas con que el alumno ingresa al sistema escolar, las que están positivamente correlacionadas con los resultados en pruebas estandarizadas posteriores y, dado el efecto positivo en el corto plazo de la asistencia al preescolar, repercute en un efecto positivo en el mediano plazo de la exposición a la educación inicial.

La especificación realizada en este trabajo permitió identificar el efecto de las habilidades obtenidas previo al ingreso al preescolar, como el efecto de la exposición a programas de desarrollo temprano. Si es que, bajo las mismas condiciones socioeconómicas, dos estudiantes alcanzaron el mismo nivel de habilidades iniciales, uno asistiendo al preescolar y otro sin asistir, y sabiendo que el preescolar contribuye a la mayor generación de habilidades cognitivas, ello sugiere que el niño que asistió menos tendría habilidades innatas mayores. De este modo, el coeficiente negativo asociado al tratamiento está capturando realmente el efecto positivo de las habilidades previas no observadas, mostrando la importancia de la estimulación de la familia y los

factores genéticos. De esta forma, los resultados sugieren que el jardín infantil tiene un impacto positivo en las habilidades cognitivas tanto en el corto como en el mediano plazo, y que las habilidades iniciales exógenas también juegan un rol importante dentro del desempeño de los alumnos.

Dado que no contábamos con grupos aleatorios y de control, hay problemas de endogeneidad que intentamos resolver mediante diferentes métodos econométricos para obtener resultados robustos. Sin embargo, la falta de datos e información limitaron también estas soluciones, ya que no pudimos controlar por todos los efectos deseados. Esto denota la importancia de la generación de datos pertinentes para los estudios de evaluación de impacto, de manera de poder orientar políticas públicas tan relevantes como lo son las políticas educativas de la forma más eficiente.

Tanto los resultados de este trabajo como la evidencia nacional muestran la importancia del desarrollo infantil temprano. A pesar de que utilizamos datos de un programa extensivo y enfocado en niños vulnerables, de igual forma se encuentran efectos positivos de la exposición a preescolar en las habilidades cognitivas, lo que muestra que a pesar de las deficiencias del sistema actual en cuanto a recursos, hay un impacto relevante en los estudiantes tanto en el corto como en el mediano plazo.

Luego, es necesario avanzar no solo en cobertura, sino también en una mejora de la calidad del servicio entregado, especialmente para los niños de menores recursos, de modo de poder potenciar y aumentar el efecto positivo de dichos programas, como fue realizado en programas intensivos en países desarrollados (Heckman et al, 2010; Campbell et al, 2002). Si es que efectivamente los logros educativos están

estrechamente relacionados con la inversión en el desarrollo infantil, como muestra la evidencia, necesitamos crear iguales oportunidades para todos los niños para que, en un futuro, puedan competir en igualdad de condiciones.

## BIBLIOGRAFÍA

- Andersen, L. (2001). *Social Mobility in Latin America: Links with Adolescent Schooling*. Washington, DC: Inter-American Development Bank.
- Behrman, J., Cheng, Y., & Todd, P. (2004). Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach. *The Review of Economic and Statistics MIT Press*, Vol. 86, No. 1, Pages 108-132.
- Berlinski, S., Galiani, S., & Gertler, P. (2009). The effect of Pre-Primary Education on Primary School Performance. *Journal of Public Economics*, Vol. 93(1-2), Pages 219-234.
- Bernal, R., Fernandez, C., & Flórez, C. (2009). Evaluación de Impacto del Programa Hogares Comunitarios de Bienestar del ICBF. *Documentos CEDE*, No. 2009-16, Páginas 1-32.
- Campbell, F., Ramey, C., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes from the Abecedarian Project. *Applied Developmental Science*, Vol. 6, Issue 1, Pages 42-57.
- Comisión de Financiamiento Estudiantil para la Educación Superior. (2012). *Análisis y Recomendaciones para el Sistema de Financiamiento Estudiantil*. Santiago, Chile.
- Contreras, D., Herrera, R., & Leyton, G. (2008). Impacto de la Educación Preescolar sobre el Logro Educativo. Evidencia para Chile. *Documento de Trabajo*.
- Curtis, L., Hammill, B., Eisenstein, E., Kramer, J., & Anstrom, K. (2007). Using Inverse Probability-Weighted Estimators in Comparative Effectiveness Analyses with Observational Databases. *Medical Care*, Vol. 45, No. 10, Suppl. 2, S103-S107.
- Encina, J., & Martínez, C. (2009). Efecto de una mayor cobertura de salas cuna en la participación laboral femenina: Evidencia de Chile. *Documento de Trabajo, Departamento de Economía Universidad de Chile*, SDT 303.

- Fundación Integra. (2011). *Indicadores de Matrícula Mensual Marzo-Julio*. Santiago, Chile.
- Fundación Integra. (2011). *Reporte*. Santiago, Chile.
- Garces, E., Thomas, D., & Currie, J. (2002). Longer-Term Effects of Head Start. *The American Economic Review*, Vol. 92, No. 4, Pages 999-1012.
- García, C., & Paredes, R. (2010). Reducing the Educational Gap: Good Results in Vulnerable Groups. *Journal of Development Studies*, Vol. 46, No.3, 535-555.
- Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, Vol. 47, No. 1, Pages 153-161.
- Heckman, J. (1999). Policies to Foster Human Capital. *NBER Working Paper #7288*, 1-50.
- Heckman, J. (2006, January 10). Catch 'em young. *Wall Street Journal*, p. A14.
- Heckman, J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, 1901-1903.
- Heckman, J., Hyeok Moon, S., Pinto, R., Savelyev, P., & Yavitz, A. (2010). The Rate of Return to the High Scope Perry Preschool Program. *Journal of Public Economics*, Vol. 94, No. 1-2, Pages 114-128.
- Loeb, S., Bridges, M., Basole, D., Fuller, B., & Rumberguer, R. (2007). How much is it too much? The Influence of Preschool Centers on Children's Social and Cognitive Development. *Economics of Education Review*, Vol. 26, No. 1, Pages 52-66.
- Lopez, J. H., & Perry, G. (2007). Inequality in Latin America: Determinants and Consequences. *Paradigma y Opciones de Desarrollo en América Latina*. Santiago, Chile: The World Bank.
- Ministerio de Planificación. (1990-2009). *Encuesta de Caracterización Socioeconómica Nacional (CASEN) - Educación*. Santiago, Chile.

Noboa, G., & Urzúa, S. (2012). The effect of participation in Public Childcare Centers: Evidence from Chile. *Journal of Human Capital*, Vol. 6, No. 1, Pages 1-34.

Reyes, L., & Urzúa, S. (2012). La demanda y oferta de la educación temprana en Chile. *Estudios Públicos*, No. 126, 45-86.

Schady, N. (2006). Early Childhood Development in Latin America and the Caribbean. *World Bank Policy Research Working Paper Series*, No. 3869.

UNESCO. (2011). *Informe de Seguimiento de la EPT en el Mundo*.

**ANEXOS**

**ANEXO A: RESULTADOS PARA ESTIMADORES CON PESO DE  
PROBABILIDAD INVERSA**

Tabla A-1 - Resultados Primera Etapa Probit para estimadores con peso de probabilidad  
inversa

<b>TF</b>	<b>Coef.</b>	<b>Err. Est.</b>	<b>Z</b>	<b>P&gt;z</b>
peso	.0557057	.0222768	2.50	0.012
talla	.1697647	.0107729	15.76	0.000
EdadIngreso	-.2578558	.0101148	-25.49	0.000
constante	-9.560654	.7487039	-12.77	0.000
<b>Pseudo R2</b>	<b>0.6536</b>			
<b>Obs</b>	<b>2412</b>			