



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

¿POR QUÉ Y HASTA CUÁNTO OBLIGAR (O INDUCIR) LA ASISTENCIA A CLASES?

GABRIEL ALONSO UGARTE VERA

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

Profesor Supervisor:
RICARDO PAREDES MOLINA

Santiago de Chile, Junio, 2009

© 2009, Gabriel Ugarte



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

¿POR QUÉ Y HASTA CUÁNTO OBLIGAR (O INDUCIR) LA ASISTENCIA A CLASES?

GABRIEL ALONSO UGARTE VERA

Tesis (Proyecto) presentada(o) a la Comisión integrada por los profesores:

RICARDO PAREDES MOLINA

JAIME CASASSUS VARGAS

RÓMULO CHUMACERO

DAVID FULLER PADILLA

Para completar las exigencias del grado de
Magíster en Ciencias de la Ingeniería

Santiago de Chile, Junio, 2009

*A mi padre y madre que con su
esfuerzo me han dado la mejor
educación que puedo tener.*

AGRADECIMIENTOS

En primer lugar agradezco a Dios por todo lo que me ha dado ya que reconozco que todo ha sido gratuidad suya. Agradezco a mis padres por el esfuerzo realizado permitiéndome tener una educación de excelencia, a mis hermanos por su preocupación, a mis tíos que me han apoyado en los momentos difíciles y a todos mis otros familiares que me han manifestado sus buenos deseos y apoyo.

Agradezco con especial afecto a mi profesor supervisor Ricardo Paredes que me ha dedicado muchas horas de su tiempo, con quien he aprendido y recibido consejos, no sólo en lo académico, sino también en lo profesional y personal. También reconozco la ayuda del profesor David Fuller, de quien obtuve los datos, y de Rómulo Chumacero, que me guió en el modelo estadístico.

No puedo evitar referirme a mis compañeros de universidad y amigos que me dieron buenos consejos, retroalimentación y referencias muy útiles, entre los que están Luis Hernández, Valeska Véliz, Gonzalo Contreras y Carolina Osses. También agradezco a mis compañeros de trabajo en Valor Agregado UC, en especial a Ana Luisa Caviedes, Ailyn Silva y Pamela Acevedo.

INDICE GENERAL

	Pág.
DEDICATORIA	ii
AGRADECIMIENTOS	iii
INDICE DE TABLAS	vi
INDICE DE FIGURAS.....	vii
RESUMEN.....	viii
ABSTRACT	ix
1 Introducción.....	1
2 Revisión de la Literatura.....	4
2.1 Causas del ausentismo.....	4
2.2 Ausentismo y desempeño académico.....	5
2.2.1 Asistencia a clases en la universidad	7
2.3 Otros factores que impactan el desempeño académico.....	8
3 antecedentes descriptivos del caso chileno.....	10
4 Modelo y resultados.....	15
4.1 Modelo	15
4.2 Efectos de primer orden	16
4.3 Endogeneidad	19
4.4 Puntos de corte	26
5 Conclusiones.....	31
5.1 Implicaciones en políticas públicas.....	32
Bibliografía	34
Anexos ..	38

A.	Anexo A: Calendario escolar.....	39
B.	Anexo B: Porcentaje de inasistencias según día de la semana	40
C.	Anexo C: Modelo de media incondicional (Efectos aleatorios).....	41
D.	Anexo D: Evaluación de los supuestos del modelo de efectos aleatorios	44
	D.1 Inspección del error inter-grupo (within).....	44
	D.2 Supuesto: Normalidad	47
	D.3 Supuesto: Homocedasticidad	48
E.	Anexo E: Evaluación de los supuestos del modelo de efectos fijos	51
	E.1 Inspección del error inter-grupo (within) y normalidad	51
	E.2 Homocedasticidad.....	52
F.	Anexo F: Test de Hansen para detección de umbrales	54

INDICE DE TABLAS

	Pág.
Tabla 3-1: Número de escuelas que imparten cuarto básico en Chile.	10
Tabla 3-2: Media y desviación estándar de características seleccionadas (año 2005).....	11
Tabla 3-3: Porcentaje de alumnos ausentes para los distintos días de la semana y meses del año 2005.	12
Tabla 3-4: Promedios de días de inasistencias por alumno según sostenedor.	14
Tabla 4-1: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis.	17
Tabla 4-2: Test de exogeneidad débil para la variable asistencia y su cuadrado.	21
Tabla 4-3: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis, para la primera etapa, donde la variable dependiente es el número de inasistencias anuales.	23
Tabla 4-4: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis, para la segunda etapa.	24
Tabla 4-5: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis, para cada uno de los dos regímenes.	28
Tabla A-1: Feriados que cayeron en días escolares hábiles del 2005.	39
Tabla A-2: Días laborables año 2005.	39
Tabla B-1: Porcentajes de inasistencias según el día de la semana.	40
Tabla C-1: Resultados de modelo de media incondicional.	41

INDICE DE FIGURAS

	Pág.
Figura 4-1: Efecto de las inasistencias.....	18
Figura 4-2: Efecto de las inasistencias en el rendimiento.....	29
Figura C-1: Variación del intercepto para las distintas escuelas.....	43
Figura D-1: Boxplot de los residuos por escuela (efectos aleatorios).....	45
Figura D-2: Residuos estandarizados versus valores ajustados (efectos aleatorios).....	46
Figura D-3: Residuos estandarizados versus valores ajustados para 20 escuelas elegidas al azar (efectos aleatorios).....	46
Figura D-4: Examen del supuesto de normalidad para los dos niveles (efectos aleatorios).....	48
Figura D-5: Examen de heterocedasticidad para variables del nivel alumno (efectos aleatorios).....	49
Figura D-6: Examen de heterocedasticidad para variables del nivel escuela (efectos aleatorios).....	50
Figura E-1: Residuos estandarizados versus valores ajustados (efectos efectos fijos).	51
Figura E-2: Examen del supuesto de normalidad para los dos niveles (efectos fijos).....	52
Figura E-3: Examen de heterocedasticidad (efectos fijos).....	52
Figura F-1: Intervalo de confianza para el umbral con término cuadrático de la asistencia.	54
Figura F-2: Intervalo de confianza para el umbral sin término cuadrático de la asistencia.	55

RESUMEN

En un conjunto de políticas acerca de la calidad de la educación se asume que existe una clara y estable relación entre el éxito académico y la consistencia en la asistencia a clases. En el caso de Chile, el pago de la subvención estatal a la mayoría de las escuelas y el requisito de aprobación anual para el alumno dependen de la asistencia a clases. Sin embargo, pese a su importancia no existen mediciones del real impacto de esta variable en el rendimiento escolar. En este trabajo estudiamos el efecto de la asistencia a clases en el rendimiento académico, y evaluamos la existencia y pertinencia de umbrales de requerimientos de asistencia mínima. Hallamos que la asistencia tiene un impacto relevante y estadísticamente significativo en el rendimiento, junto con la existencia de un punto de corte, aunque contrariamente a lo esperado, no asociado con una caída del rendimiento, lo que cuestiona la existencia de un mínimo de asistencia.

Palabras Claves: Asistencia, Rendimiento Académico, Modelos Jerárquicos, Thresholds.

ABSTRACT

In a set of politics related to education, it is assumed that there is a stable relationship between academic success and consistency in attendance. In Chile, the voucher paid to subsidized schools depends on daily student attendance and there is a minimum attendance requirement to promotion. However, besides its importance, little is known about its impact on academic achievement, especially in Chile. In this study, we analyze the effect of class attendance on academic achievement, and evaluate the existence and importance of minimum attendance requirement thresholds. We found the attendance has a relevant and statistically significant impact on performance, together with the existence of a threshold, although contrary to the expected, not associated with a decrease on performance, which questions the existence of minimum attendance requirement.

Keywords: School Attendance, Academic Achievement, Hierarchical Linear Models, Thresholds

1 INTRODUCCIÓN

En un conjunto de políticas acerca de la calidad de la educación se asume que existe una clara y estable relación entre el éxito académico y la consistencia en la asistencia a clases. En el caso de Chile, el pago de la subvención estatal a la mayoría de las escuelas y el requisito de aprobación anual para el alumno dependen de la asistencia a clases. Sin embargo, pese a su importancia no existen mediciones del real impacto de esta variable en el rendimiento escolar.

El ausentismo está entre los más grandes problemas de la educación actual (DeKalb, 1999). En Inglaterra, la tasa de inasistencias diarias en escuelas públicas se estima en un 6,7%, lo que se traduce en poco menos de medio millón de estudiantes ausentes cada día (National Audit Office, 2005). En algunas ciudades de EE.UU. como Nueva York y Los Ángeles, las tasas diarias de ausentismo ascienden a un 15% y 10% respectivamente, pudiendo llegar hasta un 30% en otras ciudades (DeKalb, 1999). En Latinoamérica, el Banco Mundial (1995), en una revisión del sistema educacional de Honduras, identificó las bajas tasas de asistencia como uno de los dos factores más influyentes en las altas tasas de abandono de las escuelas (Bedi & Marshall, 2002). Para ilustrar la magnitud del efecto, si se considera un escuela de 400 alumnos con 5 horas académicas diarias y con un 90% de asistencia promedio, se estarían perdiendo más de 36.000 horas al año (Roby, 2004).

Existe consenso en el sentido que la asistencia ejerce un efecto positivo sobre la calidad de la educación. Ello es consistente con las diversas políticas que incentivan, e incluso obligan a los padres o escuelas, a contar con ciertos niveles de asistencia. Por ejemplo, en EE.UU., los estados de Carolina del Sur y Pennsylvania, cuentan con programas de recompensas y de reconocimiento a las escuelas que tienen bajos niveles de ausentismo (Ladd, 1996; Hoachlander, 2001); y en el estado de Ohio, se exigen tasas mínimas de asistencia para las escuelas (Roby, 2004). En Inglaterra, el *Ofsted* (Office for Standards in Education) inspecciona constantemente el nivel de ausentismo en las escuelas y gestiona el proceso de registro de las inasistencias (Reid, 2006), y el ausentismo injustificado puede ser castigado incluso con el encarcelamiento de los padres (Dfes,

1996). En el caso de Chile, el pago de la subvención estatal a las escuelas con voucher y que agrupan a cerca del 90% de los estudiantes, se hace en base a la asistencia individual, y es requisito para la promoción del alumno que asista como mínimo al 85% de las clases establecidas en el Calendario Escolar Anual (Mineduc, 1988)¹.

En la base de todos los ejemplos anteriores subyace la creencia de que la consistencia en la asistencia a la escuela es esencial para el éxito. Debido a que la asistencia es una variable que puede afectarse a través de políticas (Lamdin, 1996), el estudio del ausentismo es de gran relevancia al momento de enfrentar el tema de la calidad y equidad educacional. Sin embargo, no sabemos el real impacto de la asistencia en la calidad de la educación, particularmente en el caso chileno.

La variedad de políticas de cómo se incentiva la asistencia, y la falta de estudios sobre el efecto que tiene, sugieren que se trata de políticas débilmente fundamentadas. Así, por ejemplo, la restricción que exige en Chile un mínimo de asistencia del 85% por alumno (i.e., sobre 27 días de inasistencia reprueba), difiere sustancialmente del límite que impone el Estado de Indiana en EE.UU, que es de 11 ausencias, sin que existan estudios que muestren en qué punto se produce tal quiebre. En la misma dirección, en Chile la asistencia es la variable clave a través de la cual se ancla en voucher. El presupuesto para este efecto alcanza a los US\$ 3.633 millones (US\$1=CL\$600), lo que representa un 61% del presupuesto del Ministerio de Educación en el 2008 (véase en www.dipres.cl). Muchos de los dueños de escuelas y sostenedores se oponen a este sistema, argumentando que una subvención por matrícula sería sustancialmente menos engorrosa de fiscalizar y les entregaría aportes más estables. El Sindicato de Profesores ha sostenido que la subvención en base a la asistencia sólo tiene el efecto de castigar a los más pobres, incentivar los fraudes e inducir el dolo (Colegio de Profesores de Chile, 2008). Ellos proponen fórmulas más sencillas de fiscalizar, tales como el pago por matrícula o la subvención directa a la escuela. Algunos defensores del sistema sostienen

¹ Para un análisis descriptivo y orientado a política de subsidios, véase Paredes, Ugarte, y Volante (2009).

que pagar por asistencia incentiva a velar por ella, permitiendo a los padres elegir la escuela más adecuada (Libertad y Desarrollo, 2008). No obstante, también esta discusión ha carecido de evaluaciones empíricas del real efecto de la asistencia sobre el rendimiento.

En el último tiempo, se han producido una serie de vicios en el sistema de subvenciones vigente en Chile, asociados con la malversación y apropiación indebida de recursos por parte de los sostenedores, tanto particulares como municipales, debido a diversos factores, tales como la carencia de controles de los recursos entregados a los sostenedores, los mínimos requerimientos para ser sostenedor de un establecimiento y las aparentes debilidades que existen en la capacidad de gestión y fiscalización del mecanismo de subvención, entre otros. Por lo anterior, un diagnóstico apropiado permitirá saber la forma en que los incentivos a la educación deben ser puestos, de modo de fomentar acciones que mejoren el rendimiento académico. Esto es, aunque el financiamiento de las escuelas en base de la asistencia tiene bases pedagógicas, estimar el efecto de ella en el rendimiento, con sus eventuales no linealidades y umbrales de corte.

Este trabajo aporta al conocimiento del efecto empírico que tiene la asistencia sobre el rendimiento y la naturaleza de éste, donde un desafío de este estudio tiene que ver con que el estimador del efecto de las inasistencias sobre el rendimiento del alumno estaría sesgado debido a problemas de endogeneidad por variables relevantes omitidas y causalidad opuesta, lo que se tratará mediante técnicas apropiadas. Interesa especialmente conocer si este efecto muestra umbrales que justifiquen el financiamiento de vouchers basados en la asistencia y los requerimientos mínimos para la promoción en los niveles establecidos. El trabajo se estructura en cinco secciones, aparte de esta introducción. La segunda sección describe la literatura y en especial la relación entre asistencia y desempeño académico. La tercera describe los datos utilizados y entrega ciertos antecedentes descriptivos sobre el caso chileno. La cuarta sección presenta la metodología y los resultados, y la quinta sección concluye.

2 REVISIÓN DE LA LITERATURA

2.1 Causas del ausentismo

Dentro de los estudios sobre el ausentismo, se distingue el justificado, e.g. por razones de salud, muerte de un pariente, clima o ruralidad; de aquél que no lo es (en inglés truancy), entre cuyas causas, a nivel personal, se encuentran, el aburrimiento y la pérdida de interés, asignaturas irrelevantes, falta de monitoreo y de recalcar las consecuencias académicas de la ausencia a clases (Roderick, 1997), malas relaciones con los profesores y problemas con la familia o compañeros (DeKalb, 1999; Epstein & Sheldon, 2002; Roby, 2004), las vacaciones, el buen tiempo y las presiones de los pares (Kube & Ratigan, 1992), la etnia del alumno (Steward, 2002), el grado de participación en actividades estudiantiles (McCarthy, 2002) y la importancia que los estudiantes atribuyen a la asistencia (Gump, 2006; Roby, 2004).

Con respecto a las características familiares, no está claro el efecto de dichas variables sobre la asistencia. Epstein & Sheldon (2002) le atribuyen una mayor importancia a las prácticas familiares que a su estructura.

Entre los factores relacionados con la escuela destacan las suspensiones fuera de la escuela (Arcia, 2006), la calidad de la infraestructura de la escuela (Branham, 2004), el tamaño de la escuela (Finn & Voelkl, 1993), y el clima y ambiente estudiantil (Crone et al., 1993).

Epstein & Sheldon (2002) estudian la relación, a nivel de escuela, entre la asistencia y una serie de variables identificadas previamente en la literatura, en 12 escuelas primarias de EE.UU., y encuentra que la asistencia está negativamente relacionada con las escuelas que tienen una mayor cantidad de alumnos con almuerzo gratis o de precio reducido, y con aquéllos que no tienen casa. También hallan que la asistencia está positivamente relacionada con el porcentaje de alumnos que viven a una distancia mayor a una milla de la escuela. No encuentran una relación considerable con el tamaño de la escuela y con el porcentaje de familias que no hablan inglés como lengua materna.

Corville-Smith et al. (1998) tratan de distinguir los estudiantes que presentan mayores ausencias de aquellos que asisten regularmente, y concluyen que ambos perfiles difieren en la mayoría de las variables medidas, tanto en el ámbito escolar, como personal y familiar. Los alumnos ausentes tienen más baja autoestima, son menos competentes en sus relaciones sociales, perciben una menor cohesión en sus familias, una menor aceptación de sus padres y una disciplina inconsistente, e indican una menor satisfacción por las características de la escuela y su personal.

2.2 Ausentismo y desempeño académico

Las consecuencias del ausentismo se extienden más allá de la escuela, ya que afecta también al estudiante, a la familia y a la comunidad (DeKalb, 1999). Algunas de las consecuencias indeseadas de este problema son el bajo desempeño académico y su impacto en la autoestima del alumno (con la posible causalidad opuesta), la deserción escolar, las bajas posibilidades de encontrar un buen trabajo en el futuro y sus posibles consecuencias en la economía, la pérdida de recursos de escuelas que tienen ingresos por asistencia, los ajustes y tiempo necesario para reenseñar el material por parte del profesor, la delincuencia y drogadicción, entre otras (Strickland, 1999; DeKalb, 1999; Roby, 2004; Epstein & Sheldon, 2002).

En cuanto al desempeño académico, es de esperar que los alumnos que no están en clases tengan menos oportunidades para aprender el material necesario para terminar exitosamente la escuela, si es que antes no desertan (Epstein & Sheldon, 2002). Daugherty (2008), citando a Ding & Sherman (2006), indica que si los alumnos no están asistiendo a la escuela, no se produce la interacción necesaria para aprender, con el consiguiente efecto en su desempeño académico.

Son pocos los estudios que analizan este efecto con mayor profundidad, esgrimiéndose no tanto la disponibilidad de los datos, sino la posible correlación que existe entre la asistencia y los factores socioeconómicos, con la consiguiente dificultad en la identificación del efecto de cada variable por separado (colinealidad), además de la dificultad en encontrar una relación significativa entre los factores debido a la poca

variación de la asistencia, especialmente cuando se trabaja con estudios transversales (cross-sectional) y con datos agregados (Lamdin, 1996).

A pesar de la falta de investigación, varios autores indican que el impacto de la asistencia en el desempeño escolar puede ser mayor al que siempre se ha pensado (Lamdin, 1996; Johnston, 2000). Tan interesante como el impacto mismo de la asistencia en el desempeño académico, que casi no se cuestiona (Borland & Howsen, 1998), puede ser el análisis desde cuándo el impacto de esta variable se vuelve perceptible (Lamdin, 1998). Daugherty (2008) sostiene en sus resultados que después de quince ausencias el puntaje medio en la prueba de matemáticas está por debajo del nivel de dominio exigido por el Estado. Mientras que en Chile el nivel mínimo de asistencia para aprobar el año es de un 85% (que se traduce en un máximo de 27 inasistencias considerando un año escolar promedio de 180 días), el Estado de Indiana en EE.UU tiene un límite de 11 ausencias injustificadas (ídem).

Entre las investigaciones realizadas en este ámbito está un estudio (Strickland, 1999) llevado a cabo con una muestra reducida de alumnos de una escuela secundaria pública de Chicago, que concluye que existe una correlación positiva, de moderada a fuerte, entre la asistencia y el promedio de notas. Roby (2004) compara, a nivel de escuelas primarias y secundarias en el estado de Ohio, la relación entre el desempeño académico, medido a través del porcentaje de alumnos que pasaron todas las pruebas de dominio de Ohio, y la asistencia promedio anual, encontrando que existe una correlación de moderada a fuerte entre ambos factores.

Por su parte, Daugherty (2008) realiza una investigación más completa a nivel de alumnos de escuelas de un distrito del estado de Delaware, sobre un periodo de tres años (para los cursos octavo y décimo). El estudio, que controla por otras variables como el género, la raza y el factor socioeconómico, concluye que mayores tasas de ausentismo se traducen en un peor desempeño académico, medido a través de pruebas estandarizadas de matemáticas y lenguaje. Destaca que después de quince (dieciséis) ausencias el puntaje medio en la prueba de matemáticas (lenguaje) está por debajo del nivel de dominio exigido por el Estado.

En relación a los problemas que se deben enfrentar al hacer estimaciones en este ámbito, Lamdin (1996) pone especial énfasis en que uno debe ser cuidadoso al momento de analizar los efectos obtenidos de un análisis de regresión, debido a que la variable asistencia puede ser una proxy para otros factores latentes como la motivación innata del estudiante, la preocupación de los padres o la habilidad del profesor para estimular o motivar a los alumnos, por lo que la influencia de la asistencia se vería sobre estimada. Cabe tener en cuenta que este problema es inherente a este tipo de investigaciones y que no sólo afecta a la asistencia.

Asimismo, se debe poner especial atención en determinar cuidadosamente los efectos causales de la asistencia en el puntaje, dado que es previsible hayan causalidades opuestas. Arcia (2006) estudia el desempeño académico de alumnos suspendidos, controlando antes por factores como el status socioeconómico, el género y la raza, concluyendo que las suspensiones aumentan la brecha académica entre los estudiantes y sus escuelas, y que este correctivo es ocupado mayoritariamente en aquellos alumnos con el peor rendimiento. Aquellos alumnos que con mayor frecuencia están ausentes, son justo aquellos que no debieran permitirse faltar a la escuela (Murray, 2002). Daugherty (2008) enuncia que la falta de esfuerzo para apoyar académicamente a aquellos alumnos que se van quedando atrás en su nivel, promueve el ausentismo a medida que el estudiante va pasando de curso, debido a que éstos van perdiendo la esperanza y el deseo de luchar durante el día escolar (citando a Epstein & Sheldon, 2002).

Por último, debido a la naturaleza jerárquica de los datos, Lamdin (1996) sugiere el uso de un modelo multi-nivel que considere las influencias comunes que comparten alumnos de un mismo curso y escuela, lo que traería consigo nuevas luces para el tema.

2.2.1 Asistencia a clases en la universidad

Algunos de los estudios acerca del impacto de la asistencia en el desempeño académico se han realizado también en el ambiente universitario. Romer (1993) analiza las ausencias en algunas universidades y se pregunta por las consecuencias que tiene el hecho de que en un día típico alrededor de un tercio de los alumnos estén ausentes.

Concluye que a pesar de que existe la posibilidad de que la relación entre asistencia y desempeño refleje el impacto de factores omitidos – como puede ser el caso de la mayor motivación de los alumnos – pareciera que una parte importante de la relación refleja un efecto genuino de la asistencia. Cohn & Johnson (2006), por su parte, encuentran que una mayor asistencia tiene un impacto positivo en las notas, pero que puntajes bajos no resultan en mayores ausencias; mientras que Clump et al. (2003), además de encontrar un efecto positivo en la nota final del curso, subraya el impacto que tiene la ausencia en el corto plazo, medido mediante el rendimiento en pruebas parciales durante el semestre. Confirmando los resultados de la literatura anterior, Billington (2008) concluye que en la nota final de un curso de pregrado universitario, la asistencia sí es un factor significativo, junto al promedio ponderado del alumno hasta ese semestre. Destaca que las inasistencias afectan distintamente a los alumnos según su promedio ponderado (tiene un mayor impacto en alumnos con bajo promedio ponderado acumulado), y que el efecto de las inasistencias se hace notar después de un umbral.

2.3 Otros factores que impactan el desempeño académico

Según la literatura más reciente, son varios los factores que afectan el desempeño académico de los alumnos, entre los que destacan principalmente los factores socioeconómicos, asociados tanto al alumno como a su familia (Mizala & Romaguera, 2001; Sapelli & Vial, 2002; Sapelli, 2002; Gallego, 2002; Paredes & Lizama, 2006), que influyen principalmente a través de dos canales: de la selección de la escuela por parte de los padres (de ahí la importancia de corregir por el sesgo de selección) y del capital humano inicial del estudiante o la probabilidad de aprender (Sapelli & Vial, 2002).

Otros factores que han sido considerados en la literatura son los elementos institucionales, a nivel de escuela o municipio, tales como el tipo de escuela, el ratio alumnos-docentes, el estatuto docente y la competencia; las variables de gasto para el caso de establecimientos municipales, entre las que destaca el gasto por alumno, los recursos destinados a educación por parte de la municipalidad o los recursos extras con que cuentan algunas escuelas (restricción presupuestaria blanda); el efecto de los pares o *peer effect*, cuyo impacto todavía genera discusiones (Sapelli, 2002; Hsieh & Urquiola,

2006); y los factores de gestión (Paredes & Lizama, 2006; García & Paredes, 2006; Paredes & Paredes, 2006; Paredes, 2006), “que influirían directamente haciendo los procesos más eficientes, permitiendo una mayor cantidad de recursos destinados a las actividades docentes, una mejor gestión de los recursos humanos y una mayor motivación de los alumnos” (Paredes & Lizama, 2006).

No obstante, cabe advertir que el limitado poder predictivo de los estudios econométricos que no suele superar el 36%, sugiere la existencia de otros factores muy determinantes del desempeño educativo (ídem).

3 ANTECEDENTES DESCRIPTIVOS DEL CASO CHILENO

Los datos utilizados provienen del SIMCE de cuarto básico del año 2005, de donde se obtuvieron variables de rendimiento académico, sociales, económicas y culturales del estudiante, junto a información a nivel de curso y escuela; de SINEDUC, un sistema de información de apoyo a las escuelas, de donde se obtuvo la asistencia diaria a nivel de alumno del año 2005; y del sitio web del Ministerio de Educación de Chile (www.mineduc.cl), que contiene datos descriptivos de las escuelas de Chile.

La muestra que se escogió (ver Tabla 3-1) se obtuvo de la base de datos de SINEDUC, que contiene información completa de la asistencia diaria a nivel de alumno para cerca de un 10% de las escuelas municipales de Chile entre la quinta y octava región (287 escuelas, con un total de 17.262 alumnos), de los cuales un 43% corresponde a la Región Metropolitana (204 escuelas, con un total de 12.897 alumnos).

El hecho de considerar sólo a las escuelas municipales de Chile que se encuentren utilizando la plataforma de gestión SINEDUC, tiene la ventaja de que permite identificar el efecto propio de la asistencia para este tipo de establecimientos, que podría ser distinto para otros tipos de escuelas; y no será necesario tratar el problema del sesgo de selección que se produce si los padres de los mejores alumnos toman en cuenta el tipo de dependencia al momento de matricular a sus hijos. Entre las desventajas está la no representatividad de los datos a nivel de país, lo que sin embargo será tratado mediante la asignación de ponderadores en base al Índice de Vulnerabilidad Económica y la existencia de elementos de gestión en la escuela.

Tabla 3-1: Número de escuelas que imparten cuarto básico en Chile.

Región	Escuelas en SINEDUC	Escuelas Municipales	Escuelas Municipales con matrícula > 30 alumnos
Quinta	30	408	158
Metropolitana	204	652	484
Sexta	2	388	97
Séptima	25	592	121
Octava	26	959	242
Total	287	2.999	1.102

La Tabla 3-2 presenta estadísticas descriptivas de variables seleccionadas de los 17.262 estudiantes a los que se les conoce su asistencia diaria. No es sorpresa que el promedio tanto del SIMCE de Matemática como el de Lenguaje (pruebas estandarizadas) sea más bajo con respecto a la media de ese año ya que la muestra estudiada corresponde exclusivamente a escuelas municipales (véase, por ejemplo, García & Paredes, 2006; Paredes & Lizama, 2006; Paredes & Paredes, 2006).

Se aprecia que en promedio, los alumnos no asisten a clases cerca de 9 días en el año (entre abril y noviembre, a excepción de julio), aunque existe una gran variación en el número de ausencias entre los estudiantes. Entre todos, el 11% de los alumnos ha repetido algún curso entre primero y cuarto básico y el 70% de los alumnos ingresó en primero básico o antes a la escuela donde rindió la prueba SIMCE. El promedio del ingreso familiar es de aproximadamente U\$ 374. En relación a las escuelas, más del 95% de ellas se encuentran en un área urbana, y en promedio la matrícula es cercana a los 91 alumnos en cuarto básico por establecimiento.

Tabla 3-2: Media y desviación estándar de características seleccionadas (año 2005).

Variable	Promedio	D. Estándar
ALUMNO		
Puntaje SIMCE Matemática ^a	236,3	53,8
Puntaje SIMCE Lenguaje ^a	243,8	52,4
Inasistencias (días en el año) ^b	8,9	9,1
Hombre	0,5	-
Ha repetido algún curso anterior ^c	0,1	-
Ha estado en la escuela desde primero ^d	0,7	-
Escolaridad del padre (años)	10,3	-
Escolaridad de la madre (años)	10,2	-
Ingreso familiar (/ \$10.000)	18,7	17,1
Número personas en el hogar	5,2	1,9
Cursó Kinder	1	-
ESCUELA		
Urbano	1	-
Número de alumnos ^e	838,4	410,2
IVE ^f	31,6	13,6
Media de ingresos de la escuela (/10.000)	18,7	8,3
Media de escolaridad de la madre (años)	10,2	1,5
Cantidad de alumnos	17.262	
Número de escuelas	287	

Notas:

a. El promedio SIMCE se fijó en 250 puntos a nivel nacional con una desviación estándar de 50 puntos el año 1998. El promedio nacional de los puntajes de Matemática y Lenguaje para el año 2005 fue de 247,6 y 255,5 respectivamente. Se puede apreciar que el promedio tanto en matemáticas como lenguaje es bastante menor al promedio nacional, lo que se debe principalmente a que estamos trabajando con escuelas municipales. El promedio SIMCE del año 2005 para matemática y lenguaje de todas las escuelas

municipales de Chile es 234,6 y 242,9 respectivamente.

b. Este variable indica el número de inasistencias del alumno entre abril y noviembre, a excepción del mes de julio.

c. Esta variable indica si el alumno ha repetido uno o más cursos entre primero y cuarto básico.

d. Esta variable indica si el alumno entró al establecimiento y se ha mantenido desde primero básico.

e. Esta variable indica el número de alumnos en cuarto básico.

f. Índice de Vulnerabilidad Académica de la escuela, especificado por el Ministerio de Educación.

A diferencia de otros estudios que sólo cuentan con datos agregados de asistencia por alumnos, en este trabajo se dispone de información día a día de las inasistencias de cada uno de los estudiantes, lo que permite construir una medida más depurada de las ausencias.

Tabla 3-3: Porcentaje de alumnos ausentes para los distintos días de la semana y meses del año 2005.

Mes	Porcentaje	Día de la semana	Porcentaje
Marzo	3,9	Lunes	6,6
Abril	4,5	Martes	5,9
Mayo	6,6	Miércoles	6
Junio	9,3	Jueves	5,7
Julio	8	Viernes	6,7
Agosto	6,6		
Septiembre	5,7		
Octubre	5,9		
Noviembre	5,4		
Diciembre	7,1		
	6,2	Total	6,2

Nota: El porcentaje que corresponde a los días lunes y viernes es significativamente distinto al de los otros días ($F < 0,001$)

Una primera vista de los datos diarios de asistencia sugiere un patrón de inasistencias según el día de la semana, tal como se puede ver en la Tabla 3-3. En ella se aprecia que el porcentaje de alumnos ausentes los días lunes (6,56%) y viernes (6,67%) es significativamente mayor al de los otros días (ver Anexo B). Ello sugiere la existencia de un factor ajeno al sistema escolar, donde, por ejemplo, algunos padres no estarían llevando a sus hijos a la escuela, alargando los fines de semana.

Si estos datos fueran representativos de lo que sucede a nivel nacional, en promedio cerca de 15.700 alumnos de cuarto básico estarían ausentes cada día. Si se extrapola este porcentaje a la totalidad de alumnos en escuelas subvencionadas de Chile, se obtendría que poco más de 175.000 estudiantes están ausentes cada día, lo que equivale aproximadamente a U\$610.000 (U\$1=CL\$600) que no estarían percibiendo las escuelas diariamente.

Una segunda vista a los porcentajes de inasistencias según el mes del año (Tabla 3-3), muestra que los porcentajes de inasistencias varían de forma considerable. En el primer semestre, durante los dos primeros meses (marzo y abril) los porcentajes de inasistencias son los más pequeños del año, pero éstos se disparan en los meses siguientes, donde destaca especialmente junio, con más de un 9%, posiblemente por el efecto conjunto de las bajas temperaturas y enfermedades con el del relajamiento escolar. En el segundo semestre el porcentaje se mantiene estable en cerca de un 5,5% a excepción de agosto que muestra un 6,64%. Los menores porcentajes que se ven a partir de septiembre se deberían, en parte, a la mayor importancia que se le empieza a dar a la asistencia en las escuelas, tanto en los consejos de profesores como en las reuniones con apoderados, con miras a no sobrepasar el 15% de inasistencias permitido. El mayor porcentaje en diciembre estaría en parte justificado ya que las escuelas tienden a cerrar el año con anterioridad al término de clases.

En tercer lugar, y más sugerente en relación con propósitos de política pública, la Tabla 3-4 muestra una gran variabilidad de asistencias según el municipio. Ello sugiere que las políticas municipales, tendientes a aumentar los ingresos por la vía de captar más financiamiento del voucher, y a mejorar los resultados, difieren enormemente. Estas políticas estarían teniendo un efecto importante y serían relevantes al momento de estudiar los determinantes de la asistencia. En efecto, el promedio general de inasistencias es de 8,9 días en el año, pero hay municipios que observan promedios de 13,3 días como la Municipalidad 1 (norte), o tan bajos como 4,8 días en la Municipalidad 11 (Región Metropolitana). Otras con promedios muy inferiores a la media son las Municipalidades 4, 5, 6, 13 y 15; mientras que entre las comunas con promedios muy superiores a la media se puede encontrar a la Municipalidad 3, 8, 10, 17, 18 y 19. También se puede apreciar que en el sur los alumnos faltan menos que en las comunas de la Región Metropolitana, y éstas a su vez, tienen una mayor asistencia diaria, en promedio, que los municipios del norte.

Tabla 3-4: Promedios de días de inasistencias por alumno según sostenedor.

Municipalidad	N° escuelas	Días inasistencia	D. Estándar	Mínimo	Máximo
NORTE					
Municipalidad 1	370	9,2	8,3	0	50
Municipalidad 2	696	13,3	8,5	0	59
Municipalidad 3	586	10,9	9,1	0	72
<i>Subtotal norte</i>	<i>1.652</i>	<i>11,5</i>	<i>8,8</i>	<i>0</i>	<i>72</i>
SUR					
Municipalidad 4	1.320	5,2	6,5	0	69
Municipalidad 5	1.400	7,8	7,8	0	67
<i>Subtotal sur</i>	<i>2.773</i>	<i>6,4</i>	<i>7,2</i>	<i>0</i>	<i>69</i>
REGION METROPOLITANA					
Municipalidad 6	1.011	6,9	7,9	0	53
Municipalidad 7	1.629	9,3	9,8	0	100
Municipalidad 8	929	11,8	10,1	0	76
Municipalidad 9	731	9,5	9,6	0	59
Municipalidad 10	236	11,2	8,5	0	54
Municipalidad 11	482	4,8	6,9	0	48
Municipalidad 12	1.931	10,3	10,3	0	93
Municipalidad 13	636	6	8,1	0	75
Municipalidad 14	540	9,2	8,9	0	66
Municipalidad 15	2.135	7,2	8	0	62
Municipalidad 16	347	8,5	8,6	0	56
Municipalidad 17	293	11	9,6	0	45
Municipalidad 18	609	11	10,5	0	68
Municipalidad 19	1.132	11,5	9,3	0	60
Municipalidad 20	194	10,4	7,5	0	41
Vitacura	<i>12.837</i>	<i>9,1</i>	<i>9,4</i>	<i>0</i>	<i>100</i>
<i>Total XIII Región</i>	<i>12837</i>	<i>9,1</i>	<i>9,4</i>	<i>0</i>	<i>100</i>
TOTAL	17262	8,903	9,098	0	100

Nota: Sólo aparecen sostenedores con más de 100 alumnos.

4 MODELO Y RESULTADOS

4.1 Modelo

El modelo empleado sigue la literatura en esta área, en el que el rendimiento es explicado en base a variables del alumno, de la escuela y del entorno (véase Mizala y Romaguera, 2000; Gallego, 2002; Sapelli y Vial, 2002). En nuestro caso, explicamos el rendimiento académico en matemáticas medido por la prueba estandarizada SIMCE, en base a un conjunto de variables consideradas exógenas por la literatura (variables del alumno, de su familia y de la escuela) junto con la inclusión de una variable que indica la asistencia a clases del estudiante. Concretamente, estimamos la regresión usando un modelo multinivel, de modo que contemple las influencias comunes que comparten los alumnos de una misma escuela, ya que las observaciones no serían independientes (véase, Steenbergen & Bradford, 2002)². En este modelo, el logro académico, representado por Y_{ij} , depende de un conjunto de factores y de la asistencia, lo que planteamos en la ecuación (4-1) que especifica el nivel 1 del modelo de efectos mixtos:

$$Y_{ij} = \beta_{0j} + \beta_1 A_{ij} + \beta_2 A_{ij}^2 + \beta_3 S_{ij} + \varepsilon_{ij} \quad (4-1)$$

Para $i=1, \dots, n_j$ alumnos en la escuela j , con $j=1, \dots, 287$. El predictor de interés, A_{ij} , representa el número de días que el alumno estuvo ausente durante el año (entre los meses de abril a noviembre, con excepción de julio). El vector S_{ij} contiene una serie de características del alumno (sexo, la escolaridad de la madre, el ingreso familiar y la cantidad de personas que viven en el hogar). El error ε_{ij} se asume independiente e idénticamente distribuido $N(0, \sigma_e^2)$.

El nivel 2 (escuela) queda representado por la ecuación (4-2):

$$\beta_{0j} = \gamma_{00} + \gamma_{01} C_j + \eta_{0j} \quad (4-2)$$

Para $j=1,\dots,287$. El vector C_j contiene características de la escuela (ruralidad, Índice de Vulnerabilidad Económica, promedio del ingreso de los padres de los alumnos y de la escolaridad de la madre, la experiencia de los profesores y la matrícula de la escuela). El error η_{0j} , que sigue una distribución $N(0,\sigma_0^2)$, representa aquella porción del intercepto que se mantiene sin explicar por los predictores a nivel de escuela y se supone que es independiente de los predictores a nivel de alumno.

4.2 Efectos de primer orden

Los resultados del modelo de rendimiento académico medido por el puntaje SIMCE en la prueba de Matemáticas al incluir la variable de asistencia a clases de los alumnos se muestran en la Tabla 4-1. Al realizar el test de Hausman robusto a heterocedasticidad, se rechaza la hipótesis nula de que la diferencia de los coeficientes no es sistemática, lo que indica que el modelo de efectos aleatorios no estaría arrojando estimadores consistentes y el modelo a usar es uno de efectos fijos ($p<0,05$). A pesar de este resultado, en la Tabla 4-1 se muestran los coeficientes estimados por ambos modelos, aunque en las estimaciones posteriores siempre se ocuparán efectos fijos (por lo que no se incluyen variables a nivel de escuela)³.

² El modelo de efectos aleatorios incondicional indica que un 12,5% de la variabilidad de los datos se encuentra a nivel de escuela. Esto sugiere que un análisis mediante MCO llevaría a resultados equivocados (ver Anexo C)

³ Véase Anexo D y Anexo E para el análisis de supuestos del modelo de efectos aleatorios y fijos respectivamente.

Tabla 4-1: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis.

Predictor	Coefficiente	Coefficiente	Coef.
	Ef. aleatorios	Efectos Fijos ^a	Estandarizados ^b
ALUMNO			
Inasistencia anual	-1,101*** (0,113)	-1,15*** (,115)	-0,186***
(Inasistencia anual) ²	0,011*** (0,003)	,012*** (,003)	0,076***
Hombre	5,384*** (0,885)	5,526*** (,894)	0,052***
Escolaridad de la madre	2,454*** (0,154)	2,458*** (,155)	0,149***
Ingreso familiar (/10.000)	0,407*** (0,058)	,395*** (,058)	0,126***
(Ingreso familiar) ² (/10.000)	-0,002*** (0,000)	-0,002*** (0,000)	-0,076***
Nº personas en el hogar	-1,166*** (0,249)	-1,165*** (,251)	-0,039***
ESCUELA			
IVE	-0,083 (0,109)		
Urbano	-8,816 (4,670)		
Media ingresos (/10.000)	0,684* (0,294)		
Media escolaridad de la madre	2,381 (1,278)		
Experiencia profesor	0,178** (0,052)		
Matrícula escuela	0,006* (0,003)		
Constante	177,064*** (15,547)	218,881*** (2,328)	
Número de alumnos	12.725	12.725	
Número de escuelas	285	285	
Varianza entre-escuelas (intercepto)	124,2 ^c	333	
Varianza intra-escuela	2.351,7	2.353,6	
Correlación intraclase condicional ^d	0,05	0,124	
AJUSTE			
Pseudo-R ² _e (R ² entre escuelas)	0,656	0,345	
Pseudo-R ² _e (R ² intra escuelas)	0,053	0,053	
Pseudo-R ²	0,13	0,103	
Deviance	135.210		
AIC	135.279		
BIC	135.540		
Método – Grados de Libertad	ML – 35		

Notas:

*** p<0,001; **p<0,01; *p<0,05

a. Coeficientes del modelo de efectos fijos con errores robustos. El test de Hausman entre el modelo de efectos fijos y aleatorios que es robusto a heterocedasticidad indica que debemos quedarnos con el modelo de efectos fijos (p<0.05)

b. Se calculan los coeficientes estandarizados para facilitar la interpretación al momento de comparar los efectos de los distintos coeficientes. Éstos se calculan ocupando los coeficientes no estandarizados y aplicando la fórmula de estandarización.

c. Todavía queda varianza a nivel de escuela sin explicar a nivel de escuela.

d. Se trata de la correlación intraclase condicional que mide el grado de dependencia entre observaciones intra-escuela que tienen las mismas características (medidas en el modelo)

Si se observan las magnitudes de los coeficientes estandarizados de la Tabla 4-1, que se incluyen para facilitar la interpretación al momento de compararlos, las variables con mayor impacto son, en orden, las inasistencias, la escolaridad de la madre y el ingreso familiar.

El número de inasistencias anuales aparece altamente significativo y negativo, por lo que un alumno que haya faltado 9 veces en el año (promedio de la variable inasistencia), manteniendo constante todas las otras variables, tendrá 9 puntos menos en la prueba SIMCE de Matemáticas, lo que representa un 18% de la desviación estándar. Visto desde este punto de vista, el Ministerio de Educación exige como máximo un 15% de inasistencias para aprobar el año, umbral que el año 2005 estaba en los 28 días (ver Anexo A). Un alumno que haya faltado todos estos días tendrá 22,3 puntos menos, manteniendo todo lo demás constante, lo que representa casi un 50% de la desviación estándar.

Se aprecia que el coeficiente del cuadrado de las inasistencias es positivo y significativo, lo que estaría indicando que a medida que el número de inasistencias es mayor el efecto negativo en el puntaje SIMCE de matemáticas va disminuyendo, tal como se puede apreciar en la Figura 4-1.

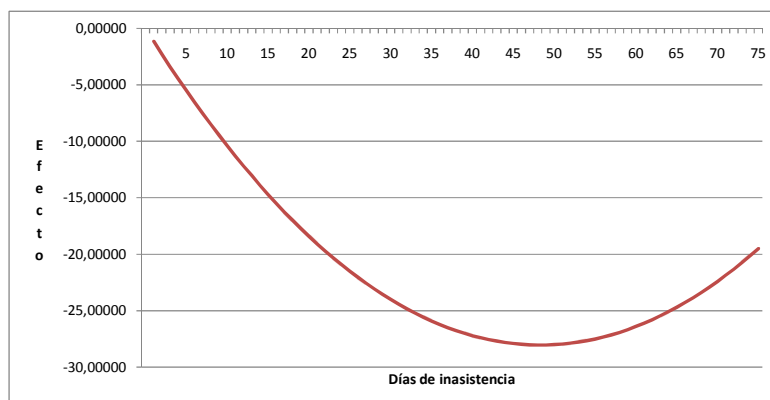


Figura 4-1: Efecto de las inasistencias.

En relación a las otras variables y acorde a la literatura, existe un sesgo de género en la prueba de matemática, donde sólo por ser hombre, un alumno obtiene 5 puntos más. Por otra parte, el número de personas que viven en el hogar afecta negativamente los resultados en el SIMCE de matemáticas, observándose que por cada persona extra en la casa el puntaje disminuye en aproximadamente un punto.

La escolaridad de la madre es preponderante y su efecto es uno de los más grandes. Cada año de escolaridad se traduce en 1,9 puntos más; es decir la diferencia entre una madre que sólo llegó hasta educación básica versus una que terminó sus estudios universitarios es de casi 19 puntos. El ingreso familiar sigue siendo importante en el estudio.

Para el caso en que sólo se consideran los alumnos en escuelas de la Región Metropolitana, los estimadores son consistentes y tienen los mismos signos y similares magnitudes.

4.3 Endogeneidad

Un problema que eventualmente tiene un estudio de la asistencia es la endogeneidad. Los estimadores β_1 y β_2 pueden estar sesgados hacia arriba debido a que las tasas de inasistencias de los alumnos pueden estar correlacionadas con características no observadas de los padres, como el esfuerzo, la motivación y la preocupación; y de los profesores, como su habilidad para estimular o motivar a los alumnos y el avance que tienen con ellos (Lamdin, 1996). Por otra parte, si se asume que los padres determinan el patrón de asistencia de sus hijos de acuerdo a las ganancias esperadas, principalmente en capital humano, y a los costos asociados de ir a la escuela (Bedi y Marshall, 2001), entonces existe un problema de endogeneidad por causalidad opuesta⁴.

En una primera etapa se testeará la exogeneidad de la variable asistencia mediante una prueba de exogeneidad débil, lo que busca testear que el rendimiento académico en el SIMCE es determinado por la asistencia a clases pero no viceversa. En una segunda

⁴ También deberían considerarse como otro factor de endogeneidad las características no observadas de los alumnos como la propia motivación, pero que no serían relevantes en este estudio, ya que sólo trabaja con alumnos de cuarto básico.

etapa se utilizará una estrategia de variables instrumentales (VI) para encontrar estimadores consistentes de la asistencia a clases.

El test de exogeneidad débil testea que el rendimiento académico en el SIMCE es determinado por la asistencia a clases pero no viceversa. Para ello, consideramos la siguiente ecuación básica:

$$Y_{ij} = \beta_0 + \beta_1 A_{ij} + \beta_2 A_{ij}^2 + \beta_3 S_{ij} + \beta_4 E_j + \varepsilon_{1ij} \quad (4-3)$$

Donde se sospecha que la asistencia a clases del alumno (y su cuadrado) pueden depender del rendimiento. Por ello, se plantean las siguientes dos ecuaciones:

$$A_{ij} = \gamma_0 + \gamma_1 S_{ij} + \gamma_2 E_j + \varepsilon_{2ij} \quad (4-4)$$

$$A_{ij}^2 = \gamma_3 + \gamma_4 S_{ij} + \gamma_5 E_j + \varepsilon_{3ij} \quad (4-5)$$

A fin de probar la exogeneidad débil de la asistencia a clases y de su cuadrado, se debe verificar que los errores de las ecuaciones (4-4) y (4-5) no estén correlacionados. Esto es, en las siguientes regresiones:

$$\varepsilon_{1ij} = \delta_0 + \delta_1 * \varepsilon_{2ij} + \varepsilon_4 \quad (4-6)$$

$$\varepsilon_{1ij} = \delta_2 + \delta_3 * \varepsilon_{3ij} + \varepsilon_5 \quad (4-7)$$

Si δ_1 y δ_3 no son significativos, querrá decir que los errores no están relacionados y por tanto que la asistencia a clases del alumno (y su cuadrado) no dependen del rendimiento. La Tabla 4-2 muestra que efectivamente, los estimadores de δ_1 y δ_3 no son significativos al 5%. Este resultado indica que no se rechaza la hipótesis de que tanto la asistencia a clases como su cuadrado son variables exógenas débiles en relación al rendimiento

académico medido por el SIMCE, lo que indica que el rendimiento académico en el SIMCE es determinado por la asistencia a clases pero no viceversa, y nos permite concluir que la estimación del modelo de efectos fijos genera estimadores eficientes.

Tabla 4-2: Test de exogeneidad débil para la variable asistencia y su cuadrado.

Predictor	Coefficiente
Modelo $\epsilon_{1ij} = \delta_0 + \delta_1 * \epsilon_{2ij} + \epsilon_4$	
Residuo Asistencia	-2,87e-9 (0,055)
Constante	4,16e-09 (0,425)
Modelo $\epsilon_{1ij} = \delta_2 + \delta_3 * \epsilon_{3ij} + \epsilon_5$	
Residuo Asistencia ²	-2,12e-11 (0,001)
Constante	4,17e-09 (0,425)

Notas: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$

Una segunda aproximación al posible problema de endogeneidad presente en los datos es mediante la implementación de una estrategia de variables instrumentales. La principal dificultad de esta estrategia radica en encontrar instrumentos que estén relacionados con las inasistencias de los alumnos y que no lo estén con la variable dependiente (puntaje de la prueba SIMCE), excepto a través de su relación con la asistencia.

La variable instrumental considerada es la cantidad de días con altas precipitaciones (más de 10 mm.) en el año en el área donde vive el alumno. Esta variable, que sí varía a nivel de alumno, presenta la ventaja en la justificación teórica de la no correlación con el error. La principal desventaja es la reducida variabilidad debido a la cantidad restringida de estaciones de medición a lo largo de Chile (se consideraron 9 estaciones de medición).

Para usar esta estrategia se utilizará Mínimos Cuadrados en Dos Etapas, estimando la ecuación estructural, donde A_{ij} es el predictor endógeno de interés; junto con la ecuación (4-8), donde A_{ij} , variable que indica el número de inasistencias anuales, es la variable dependiente.

$$A_{ij} = \gamma_0 + \gamma_1 X + \gamma_2 V_{ij} + \eta_{ij} \quad (4-8)$$

En este caso X es un vector que contiene todos los predictores que aparecen en (4-1), a excepción de la asistencia del alumno y su cuadrado, V_{ij} es un vector que incluye los instrumentos y η_{ij} es el término de error.

En la primera etapa, requerimos testear, a través de un test de Fisher, que los instrumentos de exclusión están correlacionados significativamente con la asistencia. Siguiendo a Staiger y Stock (2007), verificaremos además, si el estadístico F es mayor a diez⁵. Para determinar si la matriz es de rango completo, utilizaremos el test propuesto por Kleibergen & Paap (2006). Una vez comprobado que los instrumentos cumplen con las condiciones necesarias para desarrollar el método de variables instrumentales, podemos testear si la asistencia a clases es realmente endógena.

La Tabla 4-3 muestra los estimadores de los parámetros de (4-8), para la estrategia que utiliza como instrumento de exclusión el número de días con más de 10 milímetros de precipitaciones. Además, para incluir en la estimación el término cuadrático de la asistencia, se incluye también el cuadrado del instrumento elegido.

⁵ Si el instrumento se correlaciona débilmente, falla la convergencia asintótica convencional, a pesar que se pueda contar con intervalos de confianza estándares (5% o 1%). Por esta razón, Staiger y Stock (2007) recomiendan ver el estadístico F .

Tabla 4-3: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis, para la primera etapa, donde la variable dependiente es el número de inasistencias anuales.

1° ETAPA: ASISTENCIA		Instrumentos exclusión: Clima	Instrumentos exclusión: Clima + Clima ²
ALUMNO			
Días con más de 10mm de precipitación		0,218* (0,110)	-0,768~ (0,427)
(Días con más de 10mm de precipitación) ²			0,031* (0,013)
Hombre		0,091 (0,145)	0,09 (0,145)
Escolaridad de la madre		-0,059* (0,024)	-0,06* (0,025)
Ingreso familiar (/10.000)		- 0,054*** (0,009)	-0,053*** (0,009)
(Ingreso familiar) ² (/10.000)		0,0003*** (0,000)	0,0003*** (0,000)
Nº personas en el hogar		0,307*** (0,044)	0,308*** (0,044)
Número de alumnos		12.723	12.723
Número de escuelas		283	283
TESTS PRIMERA ETAPA			
Test F instrumento de exclusión ^a	Ho: Coeficiente instrumento=0	3,92*	4,43*
Kleibergen-Paap rk Estadístico LM ^b	Ho=Modelo underidentified	3,852*	0,974

Nota:

*** p<0,001; **p<0,01; *p<0,05; ~p<0,1

a. Se busca que el estadístico F sea mayor a 10 para testear si el instrumento es débil (Staiger y Stock, 2007)

b. Test de rango. La hipótesis nula es que los regresores endógenos no estarían identificados por los instrumentos. La hipótesis alternativa es que los instrumentos excluidos son relevantes, esto es correlacionados con la asistencia.

Tabla 4-4: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis, para la segunda etapa.

PUNTAJE SIMCE	Sin instrumento	Instrumentos exclusión: Clima	Instrumentos exclusión: Clima + Clima ²
ALUMNO			
Inasistencia anual	-1,15*** (,115)	-0,692 (2,206)	3,568 (11,161)
(Inasistencia anual) ²	,012*** (,003)	-	-0,109 (0,311)
Hombre	5,526*** (,894)	5,447*** (0,921)	4,72* (2,196)
Escolaridad de la madre	2,458*** (,155)	2,456*** (0,204)	2,435*** (0,251)
Ingreso familiar (/10.000)	,395*** (,058)	0,4** (0,131)	0,441** (0,46)
(Ingreso familiar) ² (/10.000)	-,002*** (0,000)	-0,002* (0,001)	-0,002* (0,001)
Nº personas en el hogar	-1,165*** (,251)	-1,183 (0,726)	-1,378~ (0,774)
Constante	218,881*** (2,328)		
Número de alumnos	12.725	12.723	12.723
Número de escuelas	285	283	283
Varianza entre-escuelas (intercepto)	311		
Varianza intra-escuela	2.341		
Correlación intraclase condicional	0,117		
AJUSTE			
R ² entre escuelas	0,385		
R ² intra escuelas	0,058		
Pseudo-R ²	0,114		
TESTS SEGUNDA ETAPA			
Sargan-Hansen ^a	Ho = los instrumentos excluidos son válidos	Identificada	Identificada
Endogeneidad ^b	Ho = Exogeneidad del regresor	0,000	0,022

Nota: *** p<0,001; **p<0,01; *p<0,05; ~p<0,1

a. Test para restricciones de sobreidentificación. Testea que los instrumentos no estén correlacionados con el error y que estén correctamente excluidos de la ecuación a estimar.

b. Testea si la variable a instrumentalizar es endógena (robusto)

Se puede apreciar, de la Tabla 4-3, que el test de Fischer rechaza la hipótesis nula de que los instrumentos de exclusión no estén correlacionados significativamente con la asistencia condicional a las correlaciones con los otros instrumentos. Por otro lado, el test de Kleibergen rechaza la hipótesis nula de que la matriz no sea de rango completo, para el caso en que sólo se considera el clima como instrumento. Con el rechazo de la hipótesis nula de los tests de Fischer y de Kleibergen se cumplen los requisitos exigidos a la matriz de instrumentos en la primera etapa. Debido a que las dos estimaciones están exactamente identificadas no es posible realizar el test de sobre identificación de Sargan y Hansen para verificar que los instrumentos no estén correlacionados con el error.

Verificado lo anterior, realizamos el test de exogeneidad. Los resultados indican que no se rechaza la hipótesis nula de que la asistencia a clases es exógena. Ello sugiere que la estimación reportada en la Tabla 4-1 es consistente. No obstante lo anterior, cabe remarcar algunos de los resultados encontrados.

Varios de los resultados mostrados en la Tabla 4-3 son consistentes con estudios pasados. La escolaridad de la madre sigue teniendo un impacto significativo en el patrón de asistencia de su hijo, donde los hijos de madres que completaron estudios universitarios faltan casi un día menos que los hijos de madres que sólo completaron la educación básica, manteniendo todo lo demás constante. El ingreso familiar también es significativo y con signo negativo, i.e. a mayor ingreso familiar disminuye el número de inasistencias del alumno. Este resultado es intrigante ya que la población en análisis sólo pertenece a establecimientos municipales y se podría esperar que la relación entre ingreso y asistencia no sea tan clara, ya que algunos alumnos almuerzan o desayunan en la escuela, lo que favorece particularmente a las familias de menores recursos. De todas formas, este resultado debería llevar a una investigación más profunda para averiguar cuáles son los factores que están relacionados con éste fenómeno, como puede ser el costo del transporte o la necesidad de algunas familias que necesitan del trabajo de sus hijos.

Así mismo, el número de personas en el hogar aparece con signo positivo (i.e. está directamente relacionado con el número de inasistencias anuales) y de gran magnitud.

Manteniendo todo constante, un alumno con tres familiares más en la casa asiste cerca de un día menos a clases, en promedio.

En relación a los resultados de la segunda etapa de la estrategia, se puede observar en la Tabla 4-4 que el coeficiente de las inasistencias no es significativo.

En relación al efecto de las otras variables, los coeficientes son muy similares entre OLS y cuando se utiliza la estrategia de variables instrumentales, a excepción del número de personas en el hogar que deja de ser significativo.

4.4 Puntos de corte

Una pregunta que surge naturalmente y que se asocia a la política de fijar un máximo de inasistencias, dice relación con la existencia de quiebres en determinados niveles de inasistencia. Interesa saber, entonces, si existen quiebres en el rendimiento a partir de ciertos niveles de inasistencia que permitan, por un lado, analizar la consistencia del umbral determinado por el Ministerio de Educación; y por otro, una adecuada especificación del modelo, ya que no sería lineal.

Para determinar la existencia de puntos de corte, seguimos el método de regresión de umbrales de Hansen (2000), aplicado a las ecuaciones (4-9) y (4-10). Éste permite identificar múltiples puntos de corte a través de la obtención de parámetros de regresión distintos dependiendo de la cantidad de días de inasistencia del alumno. Concretamente se definen dos regímenes:

$$Y_{ij} = \beta_0 + \beta_3 S_{ij} + \beta_4 C_j + \varepsilon_{ij} \quad \text{si } A_{ij} \leq \gamma \quad (4-9)$$

$$Y'_{ij} = \beta_0' + \beta_3' S_{ij} + \beta_4' C_j + \varepsilon_{ij}' \quad \text{si } A_{ij} > \gamma \quad (4-10)$$

Donde A_{ij} es la asistencia a clases, γ es el valor crítico de la asistencia que divide la muestra en dos grupos distintos y que se desconoce a priori, S_{ij} es el vector con las

características de los alumnos, C_j el vector con las características de las escuelas y ε_{ij} es el error de la regresión.

Como desconocemos el punto de corte, también desconocemos la distribución de errores, no pudiéndose identificar el quiebre (Hansen, 1996) y consecuentemente, tampoco se puede hacer inferencia. Sin embargo, en base de la teoría de distribución asintótica es posible construir intervalos de confianza a través de simulaciones de Monte Carlo (Hong et al., 2005).

El método de Hansen (2000), usado comúnmente en análisis de corte transversal, lo podemos extender a datos de paneles. Ello requiere restar cada variable con el promedio a nivel de escuela, eliminando los efectos fijos, lo que es válido siempre y cuando no exista endogeneidad, lo que ya fue descartado. Así, restando a cada variable el promedio de la escuela, creamos una variable binaria $d(\gamma) = \{A_{ij} \leq \gamma\}$ y definiendo $s_{ij}(\gamma) = s_{ij} * d(\gamma)$, las ecuaciones (4-9) y (4-10) se pueden escribir como:

$$y_{ij} = \beta_0' + \beta_3' s_{ij} \delta_n * (1 + s_{ij}(\gamma)) + e_{ij} \quad (4-11)$$

Donde las minúsculas expresan que se está utilizando la resta de cada variable con el promedio de la escuela (*demeaned variables*), y el subíndice n recorre todas las observaciones. Los parámetros de la regresión son β_0' , β_3' , δ_n y γ ; cuyos estimadores $\hat{\beta}_0(\gamma)$, $\hat{\beta}_3(\gamma)$, $\hat{\delta}(\gamma)$, condicionales en un valor γ , se obtienen mediante el método de mínimos cuadrados ordinarios, minimizando la suma de cuadrados residuales. Luego, $\hat{\gamma}$ es el valor que minimiza $S_n(\gamma) = S_n(\hat{\beta}_0(\gamma), \hat{\beta}_3(\gamma), \hat{\delta}(\gamma), \gamma)$, que es la suma de cuadrados residual concentrada (Ahmed & Iqbal, 2007). Siguiendo a Hansen (2000), usamos el *Likelihood Ratio* test para testear la hipótesis nula $\gamma = \gamma_0$. Los intervalos de confianza robustos a heterocedasticidad y asintóticamente correctos para el test LR se obtienen mediante replicaciones de bootstrap.

Para testear la existencia de umbrales, evaluamos si los coeficientes estimados sobre un grupo de la muestra de alumnos (los que faltan más de γ días) son iguales a los

coeficientes estimados sobre otro grupo (los que faltan más de γ días). La hipótesis es que a partir de γ días hay un quiebre reflejado en los datos.

Ya que la asistencia no es endógena, es posible aplicar el método de Hansen (2000) para testear la existencia de umbrales.

Aplicando el método de Hansen (2000), hallamos un punto de corte en torno a los 13 días de inasistencia⁶. La Tabla 4-5, que presenta los estimadores de efectos fijos para cada uno de los dos regímenes encontrados, permite observar que un alumno con características observables en el promedio en el grupo de los que faltan a los más 13 días en el año obtiene 11 puntos más que un alumno del otro grupo.

Se observa también que el estimador de las inasistencias es significativo en ambos regímenes, aunque mayor en el primero, lo que estaría indicando que el efecto de las inasistencia es mayor en el grupo de los que faltan a lo más 13 días.

Tabla 4-5: Estimadores de los parámetros seleccionados y error estándar robusto entre paréntesis, para cada uno de los dos regímenes.

Predictor	Coeficientes estandarizados			
	Inasistencia \leq 13	Inasistencia $>$ 13	Inasistencia \leq 13	Inasistencia $>$ 13
Inasistencia anual	-1,651 (0,197)	-1,318 (0,343)	-0,13	-0,222
(Inasistencia anual) ²	0,026 (0,008)	0,012 (0,005)	0,05	0,134
Hombre	6,646 (0,008)	0,848 (1,998)	0,065	0,008
Escolaridad de la madre	2,521 (0,172)	2,206 (0,342)	0,148	0,132
Ingreso familiar (/10.000)	0,427 (0,063)	0,234 (0,135)	0,128	0,068
(Ingreso familiar) ² (/10.000)	-0,002 (0)	0 (0,001)	-0,087	-0,017
Nº personas en el hogar	-1,095 (0,281)	-1,412 (0,527)	-0,038	-0,054
Constante	218,503 (2,624)	232,714 (6,898)		
Número de Observaciones	10.316	2.409		
Promedio Puntaje Matemáticas	241	230		

⁶ Valor-p<0,05. Véase Anexo E para más detalles..

En la Figura 4-2 se muestra la disminución que experimenta un alumno con características observables en el promedio (para cada uno de los dos grupos) en el puntaje SIMCE de matemáticas a medida que aumenta el número de inasistencias durante el año. Para analizar la consistencia con el umbral determinado por el Ministerio de Educación, es necesario comparar el punto de corte encontrado con los niveles de aprendizajes mínimos identificados por el SIMCE⁷, cuyo límite entre el nivel de logro intermedio e inicial también se encuentra graficado en la Figura 4-2.

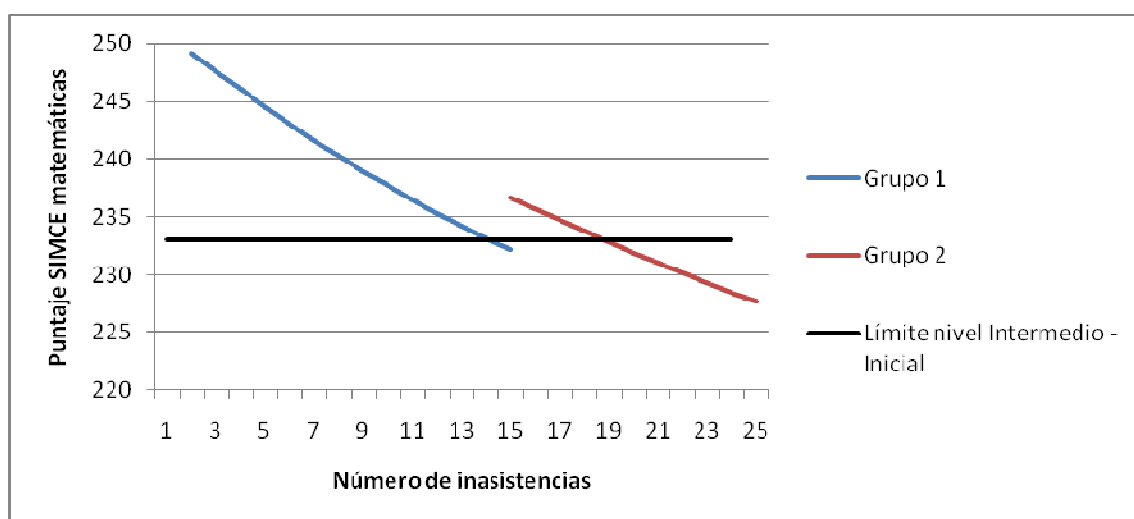


Figura 4-2: Efecto de las inasistencias en el rendimiento.

El quiebre que se aprecia tiene una interpretación opuesta a la que esperábamos. En efecto, la población de alumnos que falta más de 13 días (Grupo 2), tiene características no observables que son especialmente positivas para el rendimiento. De hecho, la Figura 4-2 sugiere que, en promedio, los estudiantes que faltan entre 14 y 17 días tienen mayores puntajes que los alumnos que faltan hasta 12 días. Ello si bien es sorprendente, le cabe una interpretación coherente, en términos que refleja que entre los alumnos que

⁷ Los posibles niveles son: Avanzado, Intermedio o Inicial, donde la categoría en que se encuentre un alumno dependerá de su puntaje, y se obtiene de considerar un puntaje mínimo que debería obtener un alumno para ser clasificado como intermedio o avanzado. Un grupo de expertos definen las preguntas que debería contestar como mínimo un alumno en el nivel intermedio obteniéndose así un puntaje de corte para este nivel. Lo mismo se realiza para el nivel avanzado. Véase, www.simce.cl

faltan mucho, aquellos que tienen capacidades especialmente altas están sobre representados, siendo una posible causa de su inasistencia la desmotivación por la calidad de la enseñanza que reciben.

Desde luego, el resultado previo sugiere que la existencia de quiebres o umbrales como requisito para que los alumnos puedan aprobar los cursos no tiene sentido, al menos desde la perspectiva de la existencia de un quiebre económicamente relevante. La disminución en el rendimiento es paulatina en el promedio de la población y por ende, la existencia de un quiebre no tiene lógica como tal.

Sin duda que a medida que aumentan las inasistencias el rendimiento medio disminuye, y que después de cierto punto el nivel de logro de los aprendizajes, en promedio, cae en el nivel inicial. De hecho, después de 17 inasistencias, un alumno con características observables en el promedio del Grupo 2, pasaría del nivel intermedio al inicial, mientras que ello ocurriría sobre 12 inasistencias para los alumnos del Grupo 1. Sin embargo, la inexistencia de un quiebre que refleje una caída drástica, indica que los umbrales no diferencian adecuadamente entre las poblaciones. Más aún, como sugieren nuestros resultados, las diferencias en el aprendizaje en torno a los umbrales detectados, de los alumnos que faltan más y menos, es contraria a la imposición de un umbral mínimo de asistencia.

5 CONCLUSIONES

Este estudio se enmarca dentro del debate en Chile acerca del esquema de incentivos para las escuelas, donde la asistencia juega un rol protagónico. Por otro lado, también se inserta dentro del debate sobre la calidad de la educación, por cuanto es una variable poco estudiada en su relación con el desempeño académico y porque existen políticas que las regulan, y algunas de ellas no tienen sustento empírico.

El hecho que la asistencia sea consensuadamente reconocida como un factor que impacta el rendimiento, ha llevado a la imposición de un conjunto de políticas públicas que la fomentan y obligan a ciertos mínimos. Tales políticas, sin embargo, requieren de evidencia concreta sobre los impactos efectivos, la que por las dificultades de estimación y disposición de datos, no existe.

En este trabajo entregamos evidencia que permite evaluar el impacto de la asistencia sobre el rendimiento, que aborde el posible problema de endogeneidad, y determine la existencia de puntos de corte o umbrales que permitan, por un lado, una mejor especificación del modelo, y por otra, la determinación de puntos críticos que ameriten políticas de asistencia mínima.

Hallamos que efectivamente la asistencia tiene efectos significativos y económicamente importantes sobre el rendimiento educacional. Concretamente, que 9 días de inasistencia anual (promedio de las inasistencias) reduce el rendimiento en al menos un 23% de la desviación estándar del puntaje en la prueba de matemáticas.

En cuanto a la existencia de saltos cualitativos, encontramos la existencia de un punto de corte estadísticamente significativo en las 13 inasistencias, pero contrariamente a lo esperado, ésta no es una discontinuidad que implique que sobre ese umbral el rendimiento cae. La ausencia de un quiebre en el sentido que esperábamos cuestiona la existencia de un mínimo de asistencia porque, al menos el quiebre, no se asocia con una caída del rendimiento.

La existencia de un mínimo de asistencia en Chile de cerca de 28 días no tiene relación ni con el quiebre encontrado ni con los niveles de logros de los aprendizajes de los alumnos. De hecho para una población con las características observables en el

promedio de los que faltan más de 13 días, la ausencia de 28 días implicaría que el 59% de esa población tendría un nivel de logro inicial, porcentaje no tan definitorio desde la perspectiva de política pública.

Una pregunta que queda sin responder en esta investigación y que puede servir para futuros estudios es si el efecto encontrado es igual en escuelas de distintas dependencias. ¿El efecto de no asistir a clases es el mismo si se trata de una escuela subvencionada que una particular pagada, o entre una pública y otra particular? Por otra parte, también queda sin estudiar la relación que existe entre las inasistencias de los alumnos y las de los profesores.

5.1 Implicaciones en políticas públicas

Debido al importante efecto de las inasistencias sobre el desempeño académico, es valioso revisar el efecto de las políticas a nivel de Estado, comuna y de escuela sobre la distribución de las inasistencias de los alumnos. A nivel de Estado, destaca especialmente el requisito de asistencia mínimo para no repetir el año. A nivel de sostenedor y de escuela, se deben estudiar los patrones de asistencia propios de los establecimientos para generar políticas acordes a la realidad.

Algunas medidas que se han utilizado exitosamente son: Informar a la comunidad, a los padres y a los propios alumnos acerca del impacto que tiene el faltar a clases, de los niveles de asistencia vigentes y de los niveles que se quiere alcanzar; el uso de llamadas por teléfono automáticas a la casa del alumno cuando falta a clases; asignar tutores a alumnos con inasistencia crónica; planear eventos especiales los lunes y viernes; asignar trabajos especiales a los alumnos que motiven su asistencia a clases; provisión de servicios de ayuda a los menores que puedan tener problemas en sus casas; proveer de una persona que esté en contacto entre la escuela y la familia del alumnos; generar los incentivos correctos, por ejemplo, premiando los alumnos con mejores niveles de asistencia; entre otras (National Association of Elementary School Principals, 2001; McCray, 2006).

Epstein y Sheldon (2002), sugieren el uso de la siguiente estrategia: Estudiar profundamente la asistencia con actividades que involucren a los estudiantes, las

familias y la comunidad; utilizar actividades más positivas de involucramiento que de castigo; y mantener el foco en la consistencia de la asistencia en el tiempo.

En relación al tema financiero, el impacto encontrado de las inasistencias en el rendimiento académico confirma la relevancia de generar los incentivos correctos que promuevan la asistencia a clases de los alumnos. En Chile, la subvención estatal por alumno se entrega a la escuela según la asistencia diaria del alumno, lo que en un principio, estaría indicando que los incentivos están bien alineados.

Pero un análisis más profundo muestra que no todas las escuelas están expuestas a este esquema de incentivos, en particular las escuelas municipales, donde existe una restricción presupuestaria blanda y los ingresos son fijos y no dependen de la matrícula o asistencia, anulándose el mecanismo de incentivo (Gallegos, 2004). Sucede también que existe una grande debilidad en la capacidad de gestión y fiscalización del mecanismo de subvención, lo que debilitaría el esquema de incentivos.

Por otro lado también es cierto que se premia el total de asistencia de la escuela, no importando si las inasistencias están distribuidas de forma similar entre todos los alumnos o si son unos pocos los que estarían faltando una cantidad considerable de días. La importancia del efecto de las inasistencias indica que el incentivo no debería estar sólo en premiar la asistencia en general, sino que llevaría a idear una forma de premiar la variabilidad que muestra el número de inasistencias de los alumnos, premiándose la menor posible.

BIBLIOGRAFÍA

Ahmed, M. & Iqbal, K. (2007). "Is There Any Threshold in Mother's Education and Child Health Relationship? Evidence from Nigeria". *Carleton College Department of Economics. Working Series*.

Arcia, E. (2006). "Achievement and enrollment status of suspended students". *Education and Urban Society, Vol. 38, 3, 359*.

Bedi, A. & Marshall, J. (2002). "Primary School Attendance in Honduras". *Journal of Development Economics, 69, 1, 129*.

Billington P. (2008). "Impact of Student Attendance on Course Grades". *Journal of American Academy of Business, 12, 2, 256*.

Borland, M., & Howsen, R. (1998). "Effect of Student Attendance on Performance: Comment on Lamdin". *Journal of Educational Research, 91, 4, 195*.

Branham D. (2004). "The Wise Man Builds His House Upon the Rock: The Effects of Inadequate School Building Infrastructure on Student Attendance". *Social science quarterly, 2004, 85, 5, 111*.

Clump, M., Bauer, H., and Whiteleather, A. (2003). "To attend or not to attend: is that a good question?" *Journal of Instructional Psychology, 3, 3, 200*.

Cohn, E. & Johnson, E. (2006). "Class Attendance and Performance in Principles Economics". *Education Economics, 14, 2, 211*.

Corville-Smith et al. (1998). "Distinguishing Absentee Students from Regular Attenders: the Combined Influence of Personal, Family and School Factors". *Journal of Youth and Adolescence, 27, 5, 629*.

Crone, L., Franklin, B., Glascock, C., & Kockan, S. (1993). "An Examination of Attendance in Louisiana Schools". *Paper was presented at the Annual Meeting of the Mid-South Educational research Association (22nd, New Orleans, LA, on November 10-12, 1993), 1-20*.

Daugherty, M. (2008). "Attendance And Other Factors That Influence Student Achievement in A Delaware Public School District". *Dissertation submitted by the faculty of Willmington University*.

DeKalb, J. (1999). "Student Truancy". *Eric Digest Number 125. ERIC Clearinghouse on Educational Management*

Dfes, U.K. Department for Education and Skills (1996). Education Act.

Ding C. & Sherman, H. (2006). "Teaching Effectiveness and Student Achievement: Examining the Relationship". *Educational Research Quarterly*, 29, 4, 39.

Epstein, J. & Sheldon, S. (2002). "Present and accounted for: Improving student attendance through family and community involvement". *The Journal of Educational Research*, 95, 5, 308.

Finn J. & Voelkl, K. (1993). "School Characteristics Related to Student Engagement". *The Journal of Negro Education*, 62, 3, 249.

Gallego, F. (2002), "Competencia y Resultados Educativos: Teoría y Evidencia para Chile", *Cuadernos de Economía, Año 39, N° 118, pp. 309-352.*

Gallego, F. (2004), "School Choice, Incentives, and Academic Outcomes: Evidence for Chile". *MIT, mimeo.*

García, C., Paredes, R. (2006). "Reducing the Educational Gap in Chile: Good Results in Vulnerable Groups". *Documento de Trabajo, Departamento de Ingeniería Industrial, Universidad Católica.*

Gump, S. (2006). "Guess Who's (Not) Coming to Class: Students Attitudes as Indicators of Attendance". *Educational Studies*, 32, 1, 39.

Hansen, B. E. (1996) "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis". *Econometric Society, the University of Chicago.*

Hansen, B. E. (2000) "Sample Splitting and Threshold Estimation". *Econometric Society, the University of Chicago* 68, 3, 575.

Hoachlander, G., et al. (2001). "Attending to Attendance". *Education Week*, 20, 36, 40.

Hong Y. et al. (2005). "Identifying Threshold Effects and Typologies in Economic Growth: A Panel Approach". *Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, July 24-27, 2005*

Hsieh C. & Urquiola M. (2006). "The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program". *Journal of Public Economics*, 90, 8-9, 1477.

Johnston, R. (2000). "As Studies Stress Link to Scores, Districts Get Tough on Attendance". *Education week*, 2000, 20, 7, 1.

Kleibergen, F., & Paap R. (2006): "Generalized reduced rank tests using the singular-value decomposition". *Journal of Econometrics*, Vol. 133, N° 1, pp. 97-126.

Kube B. & Ratigan G. (1992). "Putting the Attendance policy to the Test". *Clearing House*, 65, 6, 348.

Ladd, H. (1996). "Holding Schools Accountable: Performance-Based Reform in Education". *Washington, DC: The Brookings Institution*.

Lamdin, D., (1996). "Evidence of student attendance as an independent variable in education production functions". *The Journal of educational research*, 89, 3, 155, *Heldref Publications, Bloomington, Ill.*

Lamdin, D., (1998). "Attendance revisited: A reply to Borland and Howsen". *The Journal of educational research*, 91, 4, 198, *Heldref Publications, Bloomington, Ill.*

McCarthy, K. (2002). "The effects of student activity participation, gender, ethnicity, and socio-economic level on high school student grade point averages and attendance". *United States Department of Education. Office of Educational Research and Improvement*, 62, 409-424.

Mineduc (1988). "Reglamento de Evaluación y Promoción del Alumnos de Educación General Básica, de Educación Media, de Educación de Adultos, y Proceso de Titulación en Educación Técnico-Profesional". *Reglamento Ministerio de Educación Pública*.

Mizala, A. y P. Romaguera (2000). "Determinación de Factores Explicativos de los Resultados Escolares en Educación Media en Chile". *Centro de Economía Aplicada, U. de Chile*.

Mizala A. & Romaguera P. (2001). "Regulación, Incentivos y Remuneraciones de los Profesores en Chile". *Documento de Trabajo N° 116, CEA, Universidad de Chile*.

Murray, S. (2002). "Raising school attendance". *Education Digest*, 67, 6, 54.

National Audit Office (2005). "Improving School Attendance". *Education Journal*.

Paredes, R., Lizama, O. (2006). "Restricciones, Gestión y Brecha Educativa en Escuelas Municipales". *Documento en proceso*.

- Paredes, R., Paredes, V. (2006). "Educational Performance and Management under a Rigid Labor Regime."
- Paredes, V. (2006). "Determinantes del Desempeño Educativo en Escuelas Municipales".
- Reid K. (2006). "An Evaluation of Inspection Reports on Primary School Attendance". *Educational Research*, 48, 3, 267.
- Roby, D. (2004). "Research on School Attendance and Student Achievement: A Study of Ohio Schools". *Educational Research Quarterly*, Vol. 28 (1), 3-14.
- Roderick M. (1997). "Habits Hard to Break: A New Look at Truancy in Chicago's Public High Schools". *Students Life in High School Projects*, <http://www.consortium-chicago.org>
- Romer, D. (1993). "Do Students Go to Class? Should They?" *The Journal of Economic Perspectives*, 7, 3, 167.
- Sapelli, C. (2002). "Introducción, La Economía de la Educación y el Sistema Educativo Chileno". *Cuadernos de Economía*, Año 39, N° 118, pp. 281-296.
- Sapelli, C., Vial, B. (2002), "The Performance of Private and Public Schools in the Chilean Voucher System". *Cuadernos de Economía*, Año 39, N° 118.
- Staiger D. & Stock J. (2007). "Instrumental Variables Regression with Weak Instruments". *Econometrica*, 65, 3, 557
- Steward, R. (2002). "School Attendance Revisited: A Study of Urban African-American Students' GPA and Coping Strategies". *Department of Counseling and Education Psychology, Michigan State University, Lansing, Michigan*. 1-20.
- Steenbergen & Bradford (2002). "Modelling Multilevel Data Structures". *American Journal of Political Science*, 41, 1, 218-237.
- Strickland V. (1998). "Attendance and Grade Point Average: A Study". *Libro*
- World Bank (1995). "Staff Appraisal Report: Honduras Basic Education Project". *World Bank, Washington, DC*.

ANEXOS

A. ANEXO A: CALENDARIO ESCOLAR

La Tabla A-1 muestra la cantidad de feriados de cada uno de los meses escolares hábiles del año 2005 (que caen entre el lunes y el viernes):

Tabla A-1: Feriados que cayeron en días escolares hábiles del 2005.

Mes	Número de feriados	Día feriado
Marzo	1	Viernes 25
Abril	0	
Mayo	1	Lunes 23
Junio	1	Lunes 27
Julio	0	
Agosto	1	Lunes 15
Septiembre	1	Lunes 19
Octubre	1	Lunes 10
Noviembre	1	Martes 1
Diciembre	1	Jueves 8
Total	8	

La Tabla A-2 muestra la cantidad de días trabajados, en teoría, por las escuelas. Este número varía ya que cada escuela puede definir su calendario escolar con relativa libertad.

Tabla A-2: Días laborables año 2005.

Mes	Número de días
Marzo	22
Abril	21
Mayo	21
Junio	21
Julio	11
Agosto	22
Septiembre	21
Octubre	20
Noviembre	21
Diciembre	11
Total	191

Nota: Se suponen dos semanas de vacaciones de invierno, que no hay vacaciones en septiembre y que las clases terminan las segunda semana de diciembre.

B. ANEXO B: PORCENTAJE DE INASISTENCIAS SEGÚN DÍA DE LA SEMANA

Para obtener los porcentajes de inasistencias según el día de la semana y el según el mes se consideraron sólo los alumnos que no se retiraron de la escuela durante el año 2005 y a los cuales se les tomó asistencia continua durante el año (al menos 17 días todos los meses).

En la Tabla B-1 se indica la cantidad de veces que se tomó la asistencia cada uno de los días de la semana, los porcentajes de inasistencia según el día y si son significativamente distintos entre ellos.

Tabla B-1: Porcentajes de inasistencias según el día de la semana.

Día de la semana	Número de días que se tomó la asistencia	Porcentajes ^a	Test F diferencia con lunes ^b	Test F diferencia con viernes ^c
Lunes	554.466	0,066***		
Martes	641.894	0,059***	208,83***	307,24***
Miércoles	673.995	0,06***	164,56***	252,56***
Jueves	665.730	0,057***	378,17***	516,37***
Viernes	625.830	0,067***		
Total	3.161.915			

Notas: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$

a. Los porcentajes de obtienen de regresionar la variable inasistencia con una serie de variables dummies que indican el día de la semana, sin incluir la constante.

b. Se testea si el porcentaje de inasistencia del lunes es significativamente mayor al de los otros días de la semana.

c. Se testea si el porcentaje de inasistencia del viernes es significativamente mayor al de los otros días de la semana.

C. ANEXO C: MODELO DE MEDIA INCONDICIONAL (EFECTOS ALEATORIOS)

El modelo incondicional, a nivel de estudiante, queda especificado por:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ijk} \quad (\text{C-1})$$

Donde se sume que $\varepsilon_{ijk} \sim N(0, \sigma^2)$, para $i=1, \dots, n_j$ alumnos en la escuela $j=1, \dots, 287$, donde σ^2 representa la varianza a nivel de estudiante. El nivel 2 (escuela) se caracteriza como:

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \quad (\text{C-2})$$

Donde se asume que $\mu_{0j} \sim N(0, \tau_{00})$, donde τ_{00} representa la varianza a nivel de escuela.

Los estimadores de la varianza obtenidos con máxima verosimilitud del modelo incondicional con intercepto y del modelo que además incluye pendiente aleatoria (de la asistencia) de dos niveles se reportan en la Tabla C-1.

Tabla C-1: Resultados de modelo de media incondicional

	Modelo 1	Modelo 2
Efectos Fijos		
Promedio de media de escuelas	235,973*** (1,256)	242,851*** (1,328)
Asistencia		-0,804*** (,060)
Efectos Aleatorios		
Varianza entre escuelas (intercepto)	361,945 (37,065)	349,675 (41,574)
Varianza entre escuelas (Asistencia)		0,092 (0,075)
Cov (intercepto, Asistencia)		-0,228 (1,452)
Varianza intra-alumnos	2.483,866 (31,481)	2.439,079 (31,168)
Otros		
Número de alumnos	12.725	12.725
Número de escuelas	285	285
Deviance	136.112	135.908
AIC	136.118	135.920
BIC	136.141	135.965
Pseudo-R ² _g (R ² entre escuelas)		0,034
Método – Grados de Libertad	ML – 3	ML – 6
Dif. Deviance		204***

Nota: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$

Lo coeficientes de la varianza no incluyen el cero en el intervalo de confianza al 95%. Debido a que esta prueba no es confiable se ocupa la diferencia en los Deviance para ver si los nuevos componentes ajustan mejor el modelo, dado que en todas las estimaciones se ocupó máxima verosimilitud.

Se puede ver que el modelo con pendiente aleatoria se ajusta mejor que el modelo sólo con intercepto aleatorio, ya que la diferencia entre el Deviance de ambos modelos, que sigue una distribución Chi-cuadrado con grados de libertad según la cantidad extra de parámetros en el nuevo modelo (3 en este caso) es de 204,24, que es mayor que 16,27 (chi-cuadrado con tres grados de libertad). En la investigación sólo se incluirá el intercepto aleatorio, ya que al incluir más variable explicativas esta diferencia en el ajuste (de la Deviance) se hace no significativo ($p > 0,05$), y debido a la complicación del modelo, no se asegura convergencia.

El promedio del puntaje de matemáticas entre las escuelas es 236 (intercepto). El componente asociado al intercepto aleatorio de la escuela (varianza) es 362, estimador que es sustancialmente mayor que su error estándar, por lo que se puede apreciar que hay una variación significativa entre escuelas. Lo mismo se aprecia en la Figura C-1, donde los coeficientes calculados para las distintas escuelas varían enormemente.

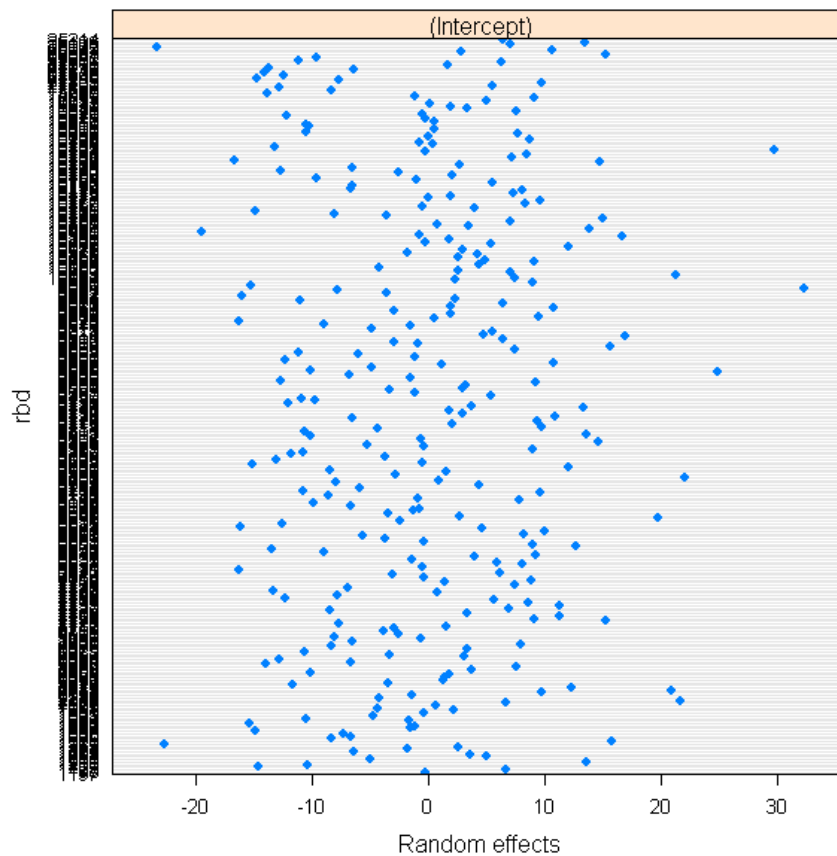


Figura C-1: Variación del intercepto para las distintas escuelas.

En este tipo de modelos, escuelas con un intercepto más alto se estima tendrán alumnos con un mayor logro académico que aquellas con un intercepto bajo, debido a características propias de la escuela.

Por último, el coeficiente de correlación intra-clase es de 0,127, lo que estaría indicando que un 12,7% de la variabilidad de los datos estaría a nivel de escuela, por lo que un análisis mediante MCO llevaría a resultados equivocados.

D. ANEXO D: EVALUACIÓN DE LOS SUPUESTOS DEL MODELO DE EFECTOS ALEATORIOS

Antes de hacer inferencias sobre el modelo de efectos mixtos ajustado se debe chequear si los supuestos de distribución de los residuos son válidos para los datos. Se examinarán dos supuestos básicos de los modelos multinivel:

1. Supuesto 1: Los errores inter-grupo son independientes e idénticamente distribuidos de forma normal, con media cero y varianza σ^2 , y son independientes de los efectos aleatorios.
2. Supuesto 2: Los efectos aleatorios están distribuidos de forma normal, con media cero y matriz de covarianza ψ (no dependiente del grupo) y son independientes para grupos distintos (Pinheiro y Bates, 2000).

D.1 Inspección del error inter-grupo (within)

En la Figura D-1 se muestran los residuos por grupo para 75 escuelas seleccionados al azar. Se puede apreciar que en general los errores están centrados en cero (i.e., $E[\varepsilon]=0$); que hay una varianza constante entre los distintos escuelas (i.e. $\text{Var}(\varepsilon_{ij})= \sigma^2$), aunque se observan outliers para algunas escuelas; y que son independientes de los niveles de grupo.

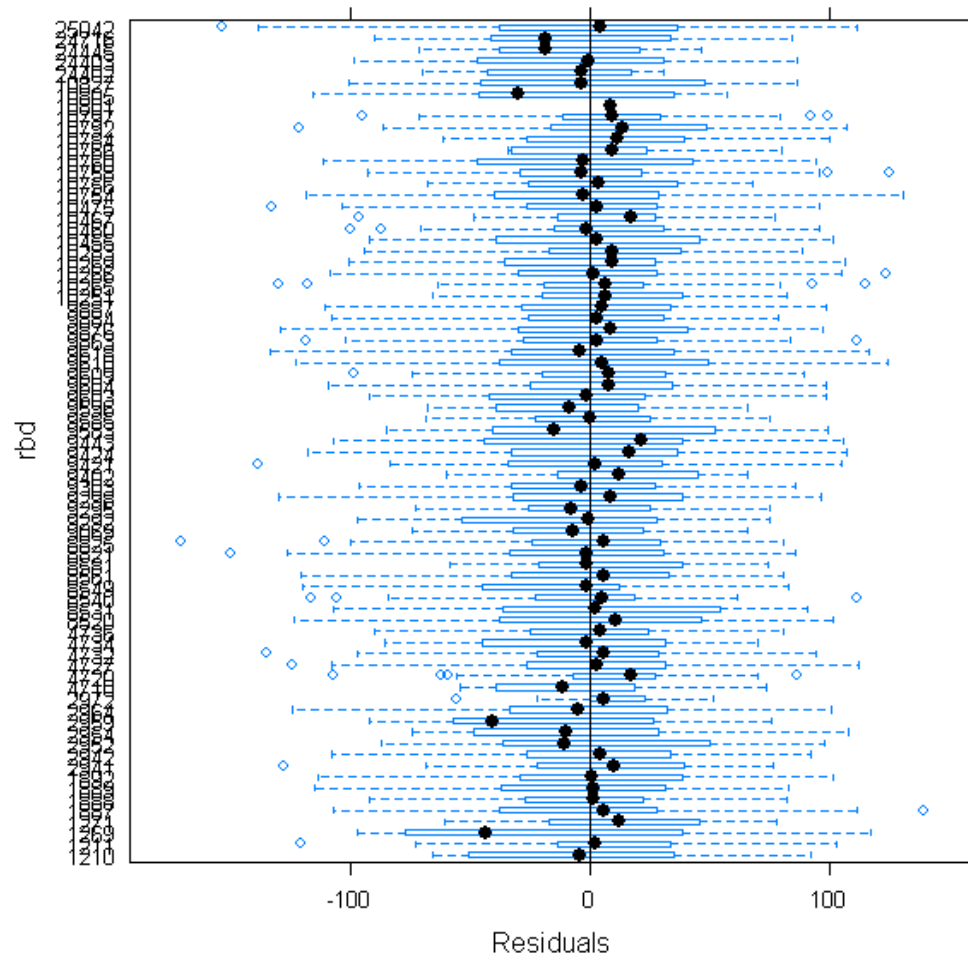


Figura D-1: Boxplot de los residuos por escuela (efectos aleatorios).

La Figura D-2 muestra los residuos normalizados versus los valores ajustados, que es muy útil al momento de investigar la heterocedasticidad y desviaciones de la normalidad. Se puede apreciar que los residuos distribuyen simétricamente alrededor de cero, con excepciones en los extremos, y varianza aproximadamente constante, aunque con la misma excepción anterior.

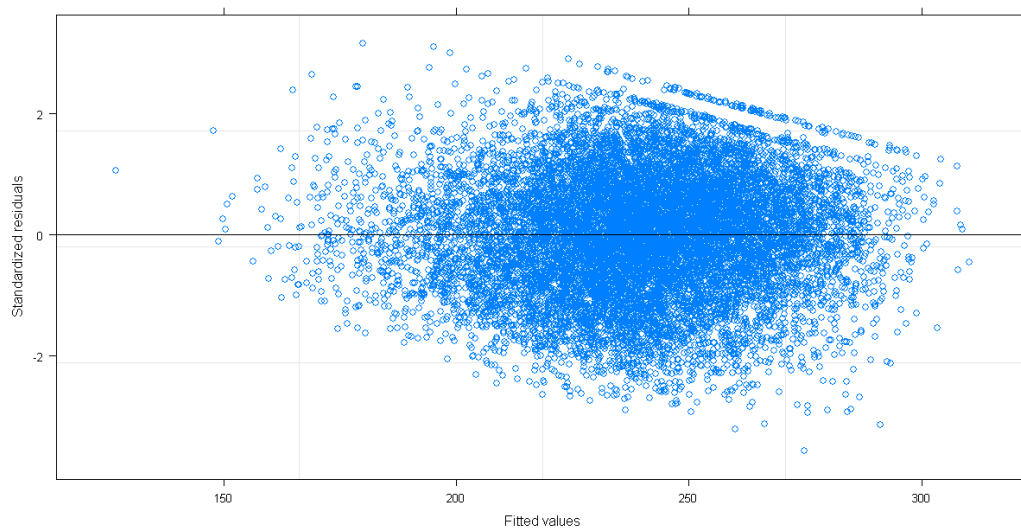


Figura D-2: Residuos estandarizados versus valores ajustados (efectos aleatorios).

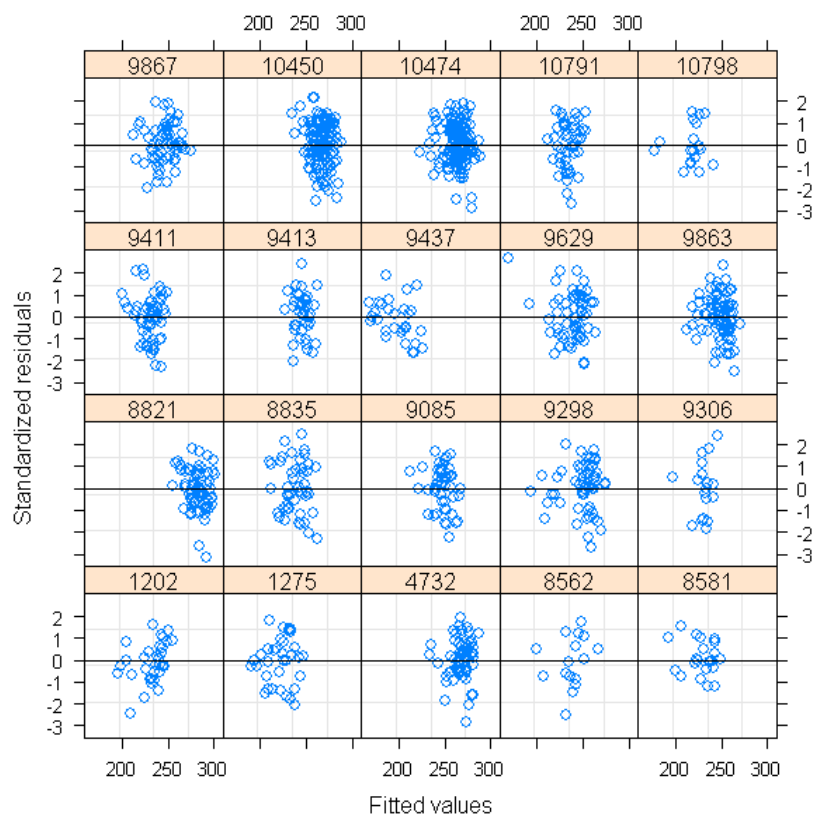


Figura D-3: Residuos estandarizados versus valores ajustados para 20 escuelas elegidas al azar (efectos aleatorios).

La Figura D-3 muestra los residuos estandarizados versus los valores ajustados para 20 escuelas elegidas al azar, donde se puede apreciar que los residuos están centrados en cero y parecieran tener una variabilidad similar entre escuelas.

D.2 Supuesto: Normalidad

A pesar que se pueden llevar a cabo varios tests formales para normalidad, se procederá a hacer una inspección visual de las distribuciones de los residuos. Para cada uno de los residuos, tanto a nivel 1 y 2, se examina si los gráficos de probabilidad normal son rectas diagonales, donde desvíos de la linealidad indicarían desvíos en la normalidad. Se puede observar en la Figura D-4 que el residuo a nivel 1 tiene una violación del supuesto de normalidad para la cola superior e inferior. Sin embargo, como las colas parecieran distribuir simétricamente, los estimadores de los efectos fijos no deberían variar sustantivamente con respecto a los obtenidos en una distribución más adecuada de los errores, más aún, este efecto estaría llevando a errores estándar más conservadores para los efectos fijos (Pinheiro y Bates p.180, 2000). El nivel 2 también muestra violaciones al supuesto de normalidad.

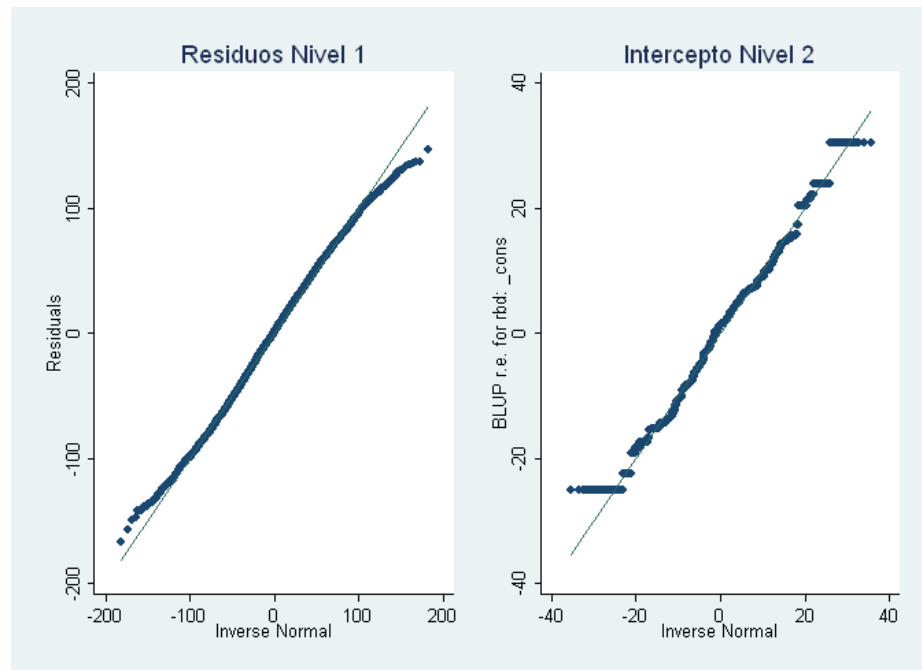


Figura D-4: Examen del supuesto de normalidad para los dos niveles (efectos aleatorios).

D.3 Supuesto: Homocedasticidad

Para evaluar este supuesto se grafican los residuos versus los valores predichos por el modelo, para los dos niveles considerados. Se busca que la variabilidad en los residuos sea aproximadamente similar para todos los valores de cada uno de los predictores.

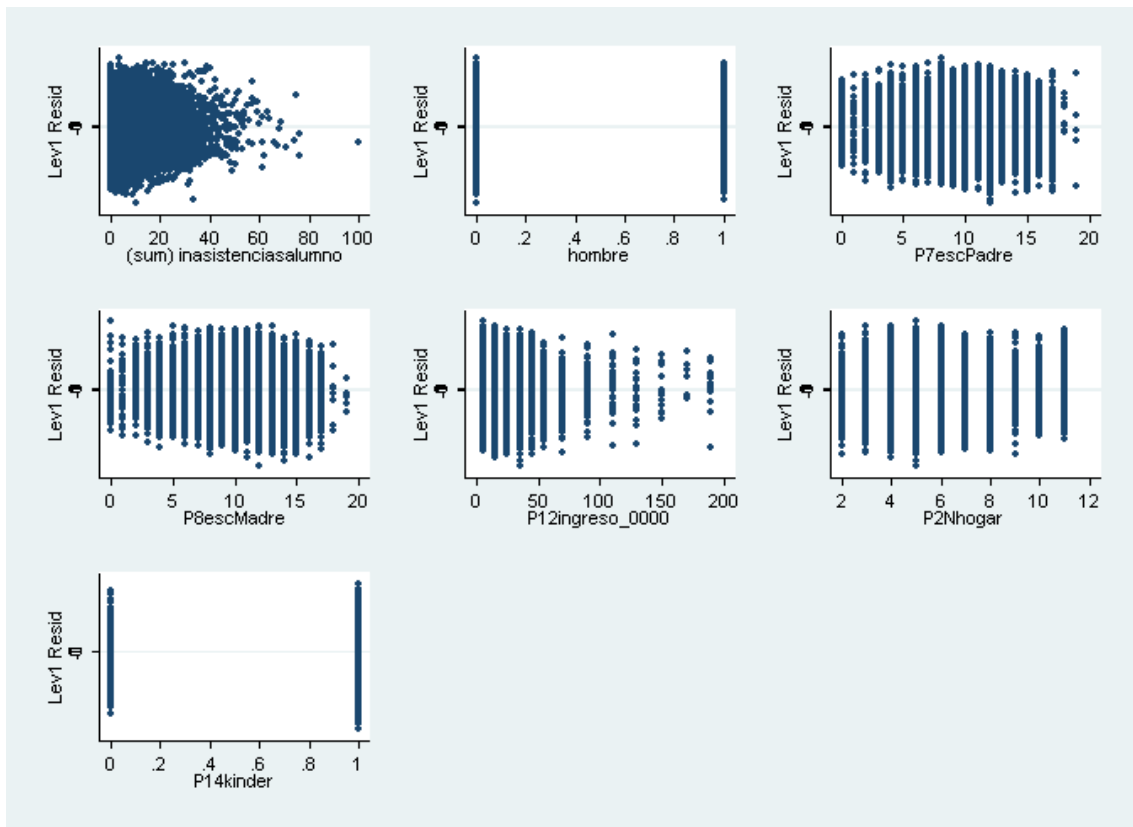


Figura D-5: Examen de heterocedasticidad para variables del nivel alumno (efectos aleatorios).

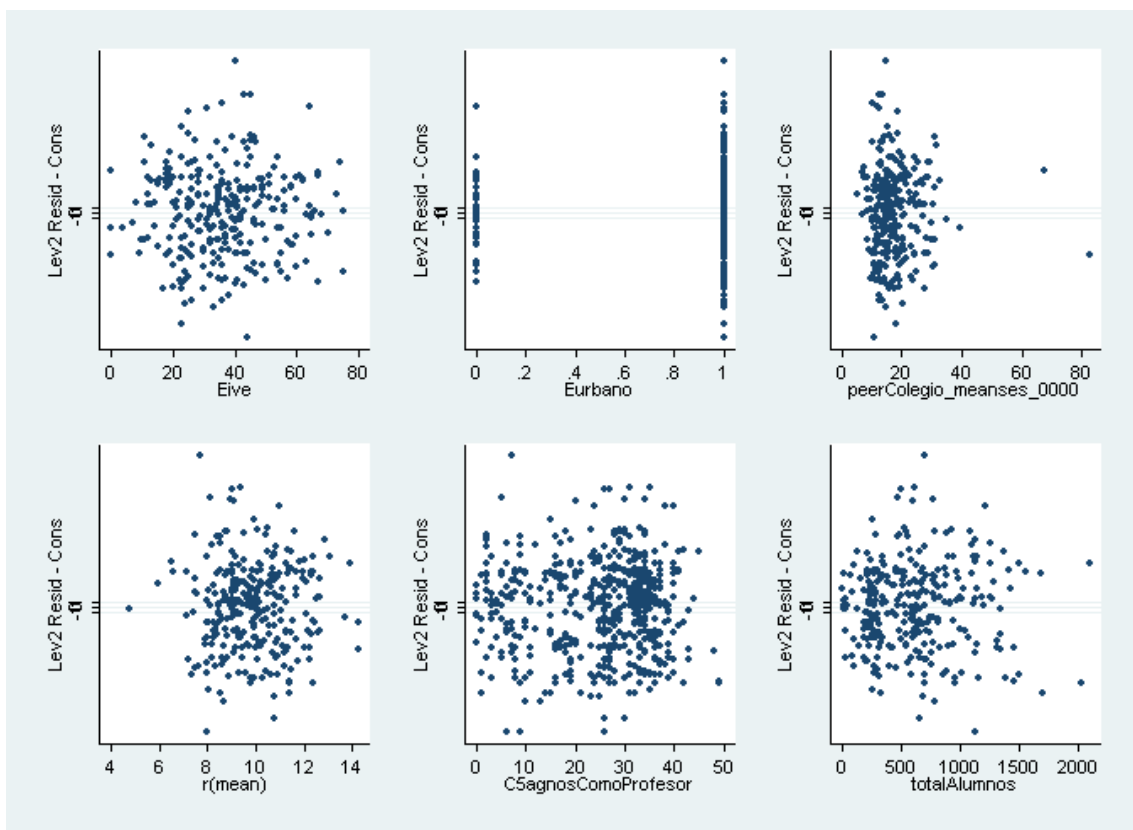


Figura D-6: Examen de heterocedasticidad para variables del nivel escuela (efectos aleatorios).

Se puede ver, de la Figura D-5 y Figura D-6, que tanto la variable que indica la asistencia a clases del alumno como la que indica el ingreso familiar del alumno son la que sugieren la presencia de heterocedasticidad en el nivel 1. Para el nivel de escuela, la media de los ingresos de los establecimientos y la zona (rural o urbana) también sugieren la presencia de heterocedasticidad.

E. ANEXO E: EVALUACIÓN DE LOS SUPUESTOS DEL MODELO DE EFECTOS FIJOS

Antes de hacer inferencias sobre el modelo de efectos fijos ajustado se debe chequear si los supuestos de distribución de los residuos son válidos para los datos.

E.1 Inspección del error inter-grupo (within) y normalidad

La Figura E-1 muestra los residuos normalizados versus los valores ajustados, que es muy útil al momento de investigar la heterocedasticidad y desviaciones de la normalidad. Se puede apreciar que los residuos distribuyen simétricamente alrededor de cero, con excepciones en los extremos, y varianza aproximadamente constante, aunque con la misma excepción anterior.

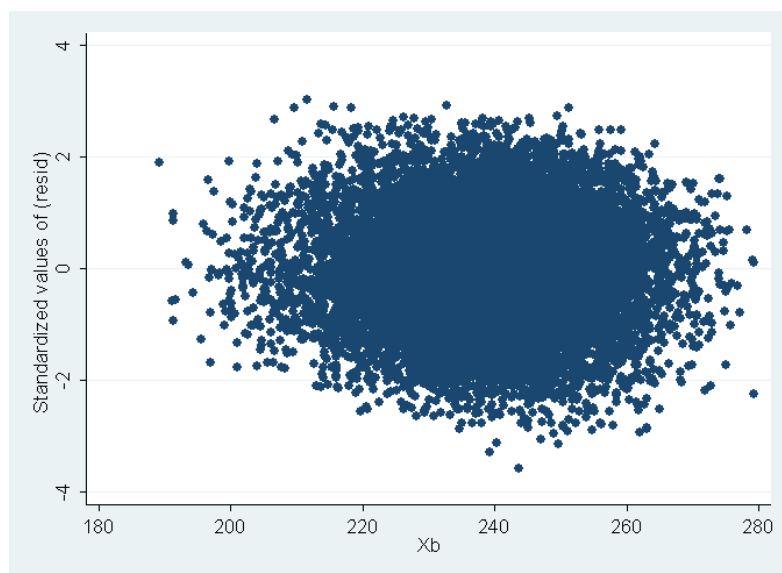


Figura E-1: Residuos estandarizados versus valores ajustados (efectos efectos fijos).

En la Figura E-2 se puede apreciar que el residuo a nivel de alumno tiene una violación del supuesto de normalidad para la cola superior e inferior, que de igual forma, aunque más pronunciada, se observa en el nivel de escuela. Sin embargo, se aprecia también que las colas parecieran distribuir simétricamente.

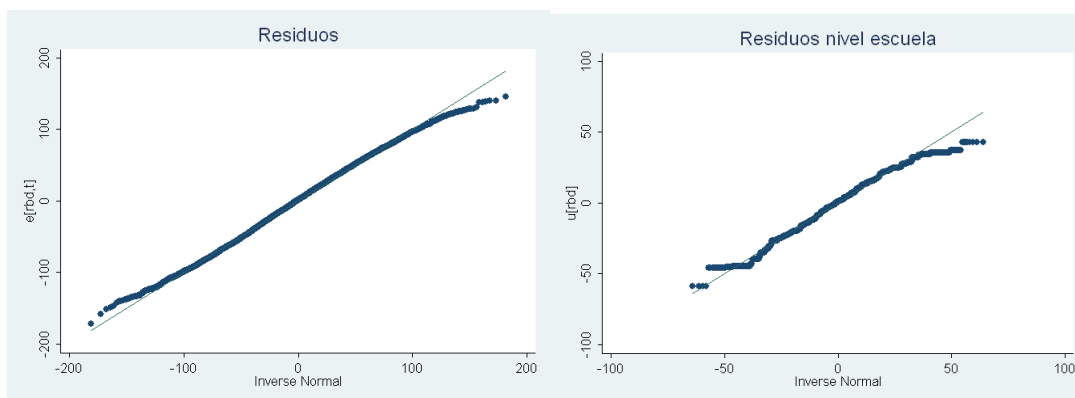


Figura E-2: Examen del supuesto de normalidad para los dos niveles (efectos fijos).

E.2 Homocedasticidad

Para evaluar este supuesto se grafican los residuos versus los valores predichos por el modelo. Se busca que la variabilidad en los residuos sea aproximadamente similar para todos los valores de cada uno de los predictores.

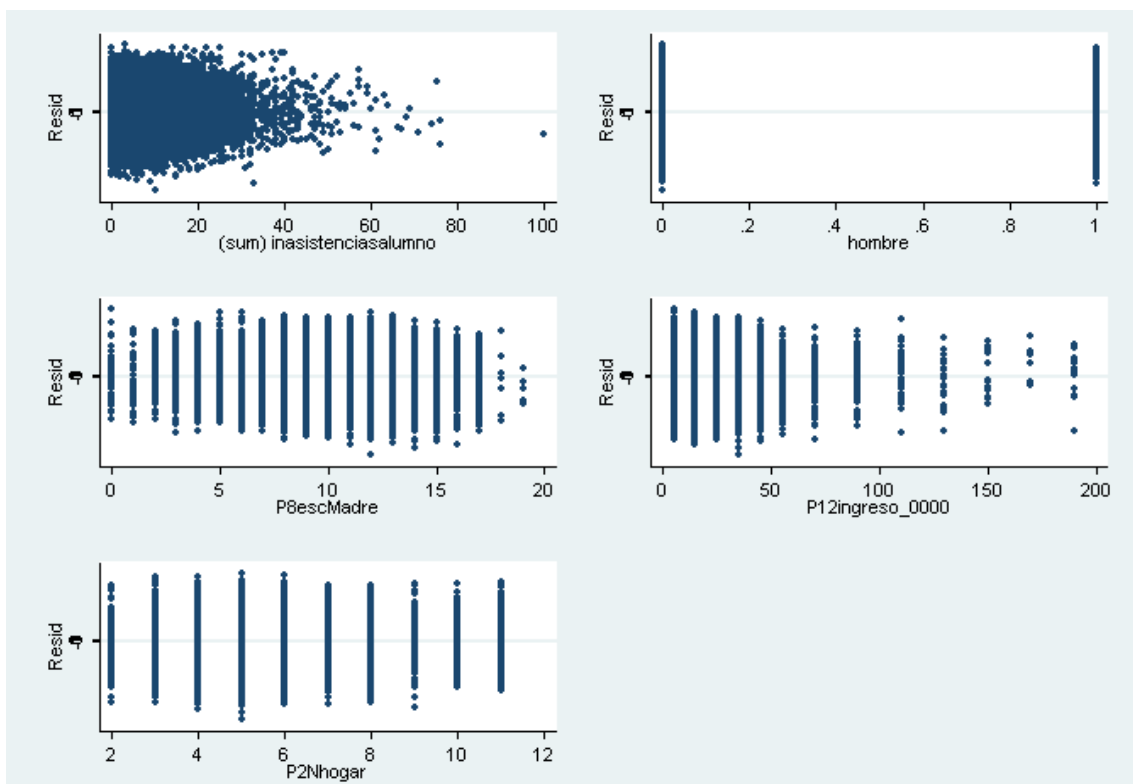


Figura E-3: Examen de heterocedasticidad (efectos fijos).

Se puede ver de la Figura E-3, que tanto la variable que indica la asistencia a clases del alumno como la que indica el ingreso familiar del alumno son la que sugieren la presencia de heterocedasticidad.

F. ANEXO F: TEST DE HANSEN PARA DETECCIÓN DE UMBRALES

El valor del estadístico LM es 23,6 con valor-p (bootstrap) de 0,048, por lo que se rechaza la hipótesis nula de la no existencia de un umbral. La Figura F-1 muestra la secuencia del estadístico LR normalizado como una función de la variable asistencia (*demeaned variable*). El estimador de mínimos cuadrados de γ es el valor que minimiza la función $LR_n^*(\gamma)$, que se observa en $\hat{\gamma} = 13$. El valor crítico asintótico al 95% está representado por la línea punteada. Su intersección con $LR_n^*(\gamma)$ muestra el intervalo de confianza $[-5,7 ; 31,9]$.

En el caso que no se incluya el término cuadrático de la asistencia, el estadístico LM es 28 con valor-p (bootstrap) de 0,002, para el mismo umbral detectado anteriormente. En este caso el intervalo de confianza está dado por $[-0,4 ; 15,9]$, tal como se muestra en la Figura F-2.

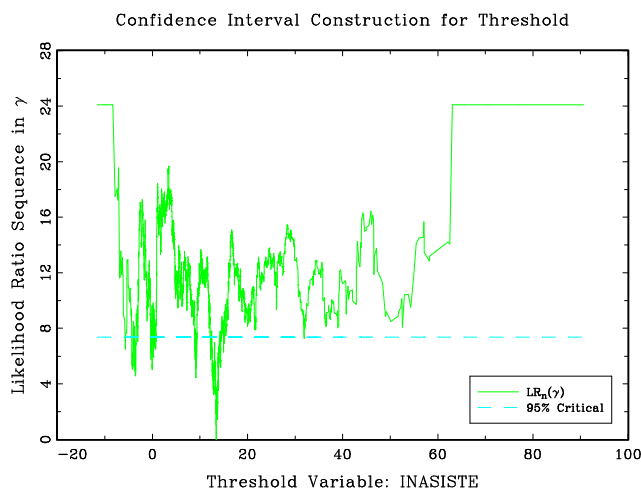


Figura F-1: Intervalo de confianza para el umbral con término cuadrático de la asistencia.

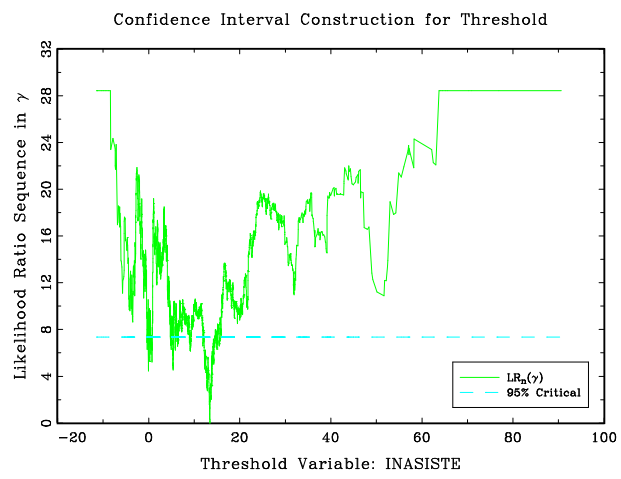


Figura F-2: Intervalo de confianza para el umbral sin término cuadrático de la asistencia.