



Facultad de Ciencias Económicas y de Administración
Universidad de la República

UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE CIENCIAS ECONÓMICAS Y ADMINISTRACIÓN

Tesis para optar al Título de Licenciado en Economía

LA EXTENSIÓN DEL TIEMPO DE CLASE EN EL AULA
¿Garantía para un mayor rendimiento académico?

INÉS MÉNDEZ 4.728.469-9

MATILDE PEÑAGARICANO 4.499.186-7

INÉS STURLA 4.727.309-0

Tutor: Graciela Sanroman

Montevideo, Uruguay

Setiembre 2011

PÁGINA DE APROBACIÓN

FACULTAD DE CIENCIAS ECONÓMICAS Y ADMINISTRACIÓN

El tribunal docente integrado por los abajo firmantes aprueba la Tesis de Investigación:

Título

La extensión del tiempo de clase en el aula. ¿Garantía para un mayor rendimiento académico?.....

Autores

Inés Méndez, Matilde Peñaricano e Inés Sturla.....

Tutor

Ec. Graciela Sanroman.....

Carrera

Lic. En Economía.....

Puntaje

.....

Tribunal

Profesor.....(nombre y firma).

Profesor.....(nombre y firma).

Profesor.....(nombre y firma).

AGRADECIMIENTOS

Queremos agradecer la ayuda brindada por Raúl Papa, Mariana Zerpa, Mari Sánchez (ANEP) y Adela Pereira (CES) por su colaboración con los datos; y un reconocimiento especial a Graciela Sanroman por todos los miércoles dedicados a nosotras gracias a los cuales esta investigación fue posible.

RESUMEN

La participación de Uruguay en las pruebas PISA 2009 dejó en evidencia problemas de deterioro del sistema educativo. Ante esto muchas opiniones surgieron sobre las causas y posibles soluciones. Una de las teorías que ha cobrado importancia en los últimos tiempos es el aumento de la cantidad de horas de clase como forma de mejorar la calidad educativa. Mientras que muchas investigaciones proveen evidencia convincente del efecto de varios insumos en la función de producción educativa, hay limitada evidencia del efecto del tiempo de clase. En esta investigación, siguiendo el estudio de Lavy (2010), estimamos el efecto del tiempo de clase sobre el rendimiento académico de los estudiantes de 15 años en matemática, ciencia y lectura, utilizando las pruebas PISA 2006 y 2009. La aproximación al tema se realiza mediante tres metodologías: en primer lugar se estima por MCO, luego se corrige el sesgo de selección que aparece debido que PISA no considera dentro de su universo a quienes desertaron del sistema educativo o aún se encuentran en primaria, y por último se realiza una estimación intra-individuo controlando de este modo por características inobservables de los alumnos y los centros educativos. Finalmente el estudio es enriquecido mediante un análisis de la descomposición de varianza.

Palabras claves: educación, PISA, capital humano, minutos de clase, inasistencias docente, estimador intra-individuo, descomposición de varianza.

TABLA DE CONTENIDO

PÁGINA DE APROBACIÓN.....	II
AGRADECIMIENTOS	III
RESUMEN.....	IV
I. INTRODUCCIÓN	1
II. ANTECEDENTES.....	5
ANTECEDENTES INTERNACIONALES.....	5
ANTECEDENTES NACIONALES.....	8
III. MARCO TEORICO	15
TEORÍA DEL CAPITAL HUMANO.....	15
EVOLUCIÓN DE LA SOCIOLOGÍA DE LA EDUCACIÓN	18
LAS ESCUELAS EFICACES	20
URUGUAY: EQUIDAD Y CALIDAD DE LA EDUCACIÓN BÁSICA.....	21
IV. DATOS.....	24
V. ESTRATEGIA EMPÍRICA.....	29
ESTIMACIÓN DEL EFECTO DEL TIEMPO DE CLASE	29
<i>Mínimos Cuadrados Ordinarios.....</i>	<i>30</i>
<i>Estimación con corrección de sesgo.....</i>	<i>30</i>
<i>Incorporación de variable Inasistencias</i>	<i>32</i>
<i>Estimador Intra-individuo</i>	<i>33</i>
ANÁLISIS DE DESCOMPOSICIÓN DE VARIANZA	35
VI. RESULTADOS.....	38
ESTIMACIÓN DEL EFECTO DEL TIEMPO DE CLASE	38
<i>Incorporación de variable Inasistencias</i>	<i>41</i>
ANÁLISIS DE LOS RESULTADOS DE DESCOMPOSICIÓN DE VARIANZA	42
VII. CONCLUSIONES.....	47
VIII. REFERENCIAS.....	49

IX. ANEXOS.....	53
ANEXO I - PRESENTACIÓN DE LAS VARIABLES Y ESTADÍSTICAS DESCRIPTIVAS	53
ANEXO II - CONSIDERACIONES METODOLÓGICAS	56
ANEXO III - RESULTADOS DE LA ESTIMACIÓN DEL EFECTO DEL TIEMPO DE CLASE .	66
ANEXO IV - RESULTADOS DE DESCOMPOSICIÓN DE VARIANZA	73
ANEXO V - ANÁLISIS DE LA INFLUENCIA DE FACTORES DEL ESTUDIANTE Y DEL CENTRO EDUCATIVO EN EL RENDIMIENTO EN LAS PRUEBAS PISA.....	75

I. INTRODUCCIÓN

Los resultados obtenidos en las pruebas PISA¹ 2009, así como otros indicadores que muestran estancamiento y en algunos casos deterioro de los resultados educativos, han generado preocupación en Uruguay y han puesto el tema de la educación en el centro del debate público.

En el año 2009 Uruguay participó por tercera vez en PISA (siendo las instancias previas en los años 2003 y 2006). Al igual que en los ciclos anteriores, obtuvimos una posición relativamente ventajosa en relación a los países latinoamericanos, pero distante de los promedios de puntaje que obtienen los estudiantes de los países de la OCDE. Más aún, al analizar los resultados a la interna del país, se observa una gran dispersión de los mismos según el entorno socio económico del centro educativo al que concurre el estudiante, lo que pone en evidencia una fuerte segmentación social del sistema educativo uruguayo (ANEP, 2009).

Lo anterior ha llevado a la discusión sobre las posibles causas del deterioro del sistema educativo en el país y sus potenciales soluciones. En relación a este tema, la mayoría de las investigaciones centran su análisis en la desigualdad social y sus consecuencias sobre el acceso a la educación y la obtención de

¹ Las pruebas PISA, dirigidas por la Organización para la Cooperación Económica y el Desarrollo (OCDE), son un conjunto de pruebas estandarizadas que evalúan cada tres años la comprensión lectora, matemática y científica, que alcanzan los alumnos de 15 años. PISA proporciona resultados para esas tres competencias básicas y analiza las prioridades y condicionantes de los sistemas educativos atendiendo a la relación entre los resultados y los contextos sociales, económicos y culturales, según las circunstancias individuales de los alumnos y conforme a la organización y funcionamiento de los centros. Más aún, proporciona un análisis de tendencias, por comparación con los datos recogidos en las evaluaciones de 2000, 2003, 2006 y 2009.

buenos resultados. Sin embargo, en los últimos años ha cobrado importancia el debate sobre el aumento de la cantidad de horas de clase como forma de mejorar la calidad educativa.

En este sentido, la información obtenida a partir de las pruebas PISA 2009 muestra la existencia de grandes diferencias entre países en cuanto al tiempo de instrucción que se dicta en los centros públicos. Mientras que en países como Francia, Italia, México, Corea del Sur y Austria, los alumnos de 15 años de edad reciben más de 1000 horas en el año de clases obligatorias, en Hungría, Suecia y China el promedio de horas que se dictan es sólo de 750 en el año (OCDE, 2010).

En Uruguay, por su parte, los jóvenes concurren al liceo público en promedio 824 hs en el año, siendo que los adolescentes de países de la OCDE se ven expuestos a 902 horas en el mismo período de tiempo. La diferencia se agrava si se tiene en cuenta que el promedio de inasistencias docentes en la educación secundaria pública uruguaya fue en el 2009 de 14.31 por ciento. Esto constituye una desventaja competitiva siendo que los desafíos del nuevo milenio demandan una sociedad cada vez más educada y preparada.

Lo anterior lleva a preguntarnos si estas diferencias pueden explicar las disparidades de resultados académicos encontrados entre países en las distintas materias dictadas, siendo la hipótesis orientadora de nuestro trabajo que efectivamente el tiempo de clase tiene un efecto positivo y significativo sobre los logros académicos de los estudiantes.

Consideramos que esta pregunta es de suma importancia principalmente porque permitiría identificar si las políticas educativas deberían estar enfocadas en aumentar la carga horaria de escuelas y liceos. En este sentido, en países como el nuestro, este tipo de políticas no debería ser de difícil aplicación dado el alto margen que existe para aumentar el tiempo de clase al año.

Mientras que muchas investigaciones proveen evidencia convincente del efecto de varios insumos en la función de producción educativa, hay limitada evidencia del efecto del tiempo de clase. El estudio más relevante en la materia es el realizado por Lavy (2010), el cual analiza la relación causal entre tiempo de clase y el rendimiento de los estudiantes de la OCDE e Israel. En esta investigación monográfica, se sigue como referencia el trabajo del autor y se aplica la metodología para el caso de Uruguay, a la vez que se enriquece el análisis mediante un estudio de descomposición de varianza.

En primer lugar se analiza la correlación simple entre las horas de clase y los resultados de las pruebas PISA 2006 y 2009 para estudiantes de 15 años, complementándose el estudio con la incorporación de datos sobre inasistencias docentes en liceos públicos para el año 2009 y observando particularmente los efectos de esta variable sobre los rendimientos educativos. Mediante este análisis no es posible inferir relación causal entre las variables de interés ya que existen posibles problemas de autoselección y endogeneidad.

Para solucionar los problemas anteriores procedemos a estudiar la variación intra-alumno del tiempo de clase. En este sentido, las pruebas PISA presentan una ventaja como fuente de información que es explotada para identificar el efecto del tiempo de clase sobre el rendimiento académico: incluyen resultados

de pruebas en múltiples materias para cada estudiante, y existe variación en el tiempo de clase entre estas asignaturas. Esto permite utilizar la estimación intra-alumno del efecto del tiempo de clase, controlando al mismo tiempo por características individuales invariantes en el tiempo que afectan de igual forma el rendimiento del estudiante en todas las materias (habilidades, contexto familiar, recursos del centro, entre otras).

Por último, con el objetivo de profundizar en el estudio de los datos, se realiza un análisis de descomposición de varianza; moviéndose el foco de estudio de este trabajo desde la búsqueda del efecto causal entre el tiempo curricular y el rendimiento de los alumnos, al análisis de cuánto de la variabilidad de los resultados puede explicarse por diferencias entre los alumnos.

El resto del trabajo se organiza del siguiente modo: en el segundo capítulo se realiza un resumen del trabajo de Lavy (2010), investigación internacional más relevante para nuestro estudio, a la vez que se sintetizan los principales trabajos nacionales que utilizan como base de datos a las pruebas PISA. En el tercer capítulo se estudia la importancia de la educación en la economía, y se realiza una sistematización bibliográfica de los distintos enfoques teóricos que han surgido respecto a las variables que influyen en la educación. En el cuarto capítulo se describen las bases de datos utilizadas y se realiza una primera aproximación a la información. En el quinto capítulo se presenta la estrategia empírica seguida. En el sexto capítulo se presentan los resultados alcanzados y por último en el séptimo se concluye.

II. ANTECEDENTES

En este capítulo se presentan los principales antecedentes del trabajo, dividiéndose el mismo en dos secciones. En la primera se realiza una revisión del trabajo de Lavy (2010), el cual se toma como referencia directa para la monografía. En la segunda sección, se sintetizan los principales trabajos nacionales que utilizan como base de datos a las pruebas PISA, buscando de este modo resumir los aportes que ya se hicieron en esa materia en el país.

Antecedentes Internacionales

En su trabajo de investigación, Lavy (2010) estudia la relación causal entre tiempo de clase y los logros académicos de los alumnos en ciertas materias.

Para esto utiliza dos bases de datos: por un lado las pruebas PISA 2006, las cuales se aplican a estudiantes de 15 años y en las que se miden sus habilidades en tres áreas de estudio (ciencias, matemática y lectura), así como se releva información sobre los estudiantes, sus familias y los centros educativos; y por el otro lado las pruebas GEMS (*Growth and Effectiveness Measures for Schools - Meizav* en Hebreo) tomadas en Israel a estudiantes de 5° y 8° grado en 2002-2005 en matemática, ciencia, inglés y hebreo.

Las fuentes de datos utilizadas presentan dos características fundamentales que le permiten estudiar el efecto del tiempo de clase en el rendimiento académico de los estudiantes. En primer lugar, las dos incluyen los resultados de los individuos en las pruebas realizadas sobre varias materias, existiendo diferencias relativamente grandes en el tiempo de clase entre esas disciplinas. De esta forma, es posible para el autor utilizar la variación del tiempo de clase

intra-alumno, controlando al mismo tiempo las características invariantes del estudiante que afectan de igual forma el rendimiento de éste en las materias (habilidad, contexto familiar, recursos de los centros, etc.). En segundo lugar, existe además considerable variación del tiempo de clase intra-alumno entre distintos países, lo que es utilizado por Lavy para estudiar si el efecto de la carga horaria difiere entre países desarrollados y en desarrollo.

En primer lugar, el autor realiza estimaciones del efecto de tiempo de clase en los países de la OCDE, utilizando en primera instancia el Método de Mínimos Cuadrados Ordinarios, y luego el estimador intra-individuo. Encuentra que el efecto del tiempo de clase es significativo, positivo, y de gran magnitud, y no dramáticamente sensible al conjunto de controles adicionales incluidos en la regresión.

En segundo lugar, presenta una serie de análisis de robustez y especificaciones alternativas que apoyan la interpretación causal encontrada. Para esto divide la muestra entre aquellos centros educativos que tienen programas especiales en ciertas disciplinas, los que separan las clases según las habilidades de los estudiantes, entre otros criterios. En particular, estudia la inclusión de centros privados como fuente potencial de sesgo, ya que los mismos tienen mayor tendencia a seleccionar a los alumnos en base a su historial académico. Para esto, realiza una estimación del efecto del tiempo de clase basado en una muestra que incluye sólo instituciones públicas. Los resultados obtenidos para esta estimación, así como para el resto de los análisis fueron similares a los originales, lo cual llevaría a pensar que no hay problemas de selección de estudiantes que estén influyendo los resultados.

En tercer lugar, para hacer un estudio más profundo de los efectos del tiempo de clase, explora los efectos heterogéneos de las horas de clase en distintos subgrupos de la muestra. Para esto reporta estimaciones separadas para niñas y niños, encontrando un impacto positivo del tiempo de clase en ambos géneros, pero con un efecto marginal mayor para las niñas. Además, fracciona la muestra según el nivel educativo de los padres, encontrando mayores resultados para los niños con padres de bajo nivel educativo. Finalmente, estudia los efectos en los inmigrantes, separando entre primera y segunda generación de inmigrantes. Los resultados son de gran magnitud para las primeras generaciones, pero aún más para los de segunda generación. Esto mismo lo realiza también para dos muestras de países: catorce países de Europa del Este y trece países en desarrollo (entre los que se encuentran Argentina, Brasil y Uruguay). Las estimaciones encontradas fueron en todos los casos mayores para los países de Europa del Este, a excepción del efecto del tiempo de clase en los estudiantes inmigrantes, el cual resultó superior para los países en desarrollo.

Por último, estudia ciertos factores estructurales del sistema educativo que pueden afectar la productividad del tiempo de clase, como son la responsabilidad del centro educativo y el grado de autonomía del mismo. Al incorporar estos controles encuentra que la productividad del tiempo de clase es mayor en las instituciones que implementan medidas de responsabilidad, y en las que tienen un mayor grado de independencia en cuanto a la asignación de sus recursos.

En cuanto al estudio de los datos de Israel, la estructura de GEMS le permite seguir a una muestra de alumnos desde 5° grado en el 2002 hasta 8° grado en el 2005. Para esto construye una base de datos longitudinal a nivel de estudiante para examinar cómo cambios en los resultados académicos de los estudiantes en las tres materias están asociados a cambios en la carga horaria de las mismas. Los resultados encontrados en Israel dotan de mayor credibilidad a los surgidos de PISA 2006, al obtener resultados similares del efecto del tiempo de clase en el rendimiento de los estudiantes.

Tanto la evidencia de la muestra PISA 2006 de estudiantes de 15 años de 50 países, y de la muestra de niños de 10 y 13 años en Israel, demuestran de forma consistente que el tiempo de clase tiene un efecto positivo y significativo en los resultados de las pruebas.

Antecedentes Nacionales

Diversos estudios se han realizado en el Uruguay utilizando las pruebas PISA como base para el análisis. Nos interesa en esta sección presentar una síntesis de los más relevantes, haciendo énfasis en las metodologías utilizadas y los resultados encontrados por los mismos. Asimismo se presenta un trabajo de investigación realizado por Cerdan-Infantes y Vermeersch (2007) el cual resulta de interés debido a la pregunta que intenta contestar y la metodología utilizada.

Llambí y Perera (2008) estudian el problema en la identificación de los efectos de haber asistido durante la educación primaria a una Escuela pública de Tiempo Completo (ETC) sobre las competencias de los estudiantes a los 15 años de edad. La motivación del trabajo surge no sólo por la importancia de las

ETC dentro de la política educativa del país, sino por los resultados evidenciados en el informe “Uruguay en PISA 2006” (ANEP, 2007) que constatan que los estudiantes que concurrieron a ETC durante la mayor parte de primaria tuvieron resultados más bajos en las pruebas PISA, en el área de ciencias, que los niños que asistieron a escuelas comunes. El objetivo de los autores es valorar hasta qué punto lo anterior puede considerarse un efecto causal en un enfoque de Función de Producción Educativa (FPE).

Los autores encuentran dos limitaciones a la hora de trabajar con los datos PISA: en primer lugar la misma no observa a los estudiantes que abandonaron el sistema educativo a la edad de 15 años, lo cual conlleva un posible sesgo de selección en la estimación de la FPE; en segundo lugar sospechan posibles problemas de endogeneidad de algunas variables.

Para abordar el primer problema, utilizan el método de Heckman en dos etapas, mientras que para el tratamiento de la endogeneidad de la variable ETC realizan una estimación con Variables Instrumentales (VI) y un modelo de selección de Heckman (HECK).

De la estimación con corrección del sesgo de selección, obtienen resultados similares a los que surgen de la estimación por MCO, no afectándose mayormente la significación del efecto de las ETC. No obstante, las estimaciones que contemplan además la posible endogeneidad de la variable ETC, modifican los resultados constatándose un cambio de signo en el coeficiente respecto a las estimaciones anteriores que no abordan este problema.

Por lo tanto, los autores concluyen que de la información de PISA no se puede inferir un efecto causal negativo entre la asistencia a una ETC y las competencias evaluadas, como lo sugiere la correlación parcial entre ambas variables o la propia estimación por MCO de la Función de Producción Educativa.

En un segundo trabajo, Llambí, Perera y Messina (2009), realizan un diagnóstico de la evolución de la desigualdad en el acceso y en los logros educativos de la población uruguaya en los últimos veinte años, siendo el estudio acotado a los logros a nivel de la educación obligatoria (Primaria y Primer Ciclo de Educación Media). Para esto utilizan en una primera etapa microdatos de la Encuesta Continua de Hogares (ECH) del Instituto Nacional de Estadística entre los años 1991 y 2007, en la cual se relevan un conjunto de características socioeconómicas y demográficas de los hogares y sus miembros. Siguiendo el concepto de igualdad de oportunidades de Roemer (1998), los autores definen y estiman un índice de desigualdad de oportunidades, así como distintos indicadores de logros educativos de la población joven. Para esto, en primer lugar, definen “variables de circunstancias” y analizan el grado de condicionamiento que ejercen sobre los logros alcanzados por los jóvenes. En segundo lugar, realizan una nueva cuantificación de la desigualdad de oportunidades en educación incorporando al análisis un conjunto de factores escolares, a partir de los datos de las pruebas PISA 2006. Para esto se sigue la misma lógica que con la ECH, pero se procede a la estimación de una medida de desigualdad de oportunidades diferente, propuesta en Bourguignon, Ferreira y Menéndez (2003), la cual se

basa en la estimación de un modelo para la variable de resultado y la utilización de un indicador de desigualdad.

En cuanto a los resultados encontrados, se destaca el hecho de que el mayor nivel de desigualdad de oportunidades se constata en la tasa de deserción del Ciclo Básico, seguido por la tasa de completitud en tiempo del Ciclo Básico obligatorio. El análisis de la contribución de los distintos factores sobre el índice de desigualdad de oportunidades les permite advertir el elevado peso de la transmisión intergeneracional de las mismas. Por su parte, del análisis basado en los datos de PISA 2006 surge que el conjunto considerado de variables escolares explican entre el 38 y el 43 por ciento de la desigualdad observada en los resultados de las pruebas en las tres disciplinas evaluadas. En lo que refiere al efecto parcial de determinadas circunstancias se observa que entre un 17 y 22 por ciento de la desigualdad de oportunidades se explica por factores individuales y familiares, mientras que el contexto escolar, medido como el nivel socioeconómico medio de los alumnos del centro, resulta el factor de mayor importancia, explicando entre el 29 y el 33 por ciento de la desigualdad de las competencias.

Méndez y Zerpa (2009) en su trabajo de investigación monográfico, evalúan en qué medida los sistemas educativos de Chile y Uruguay contribuyen a la equidad. Para ello estiman una regresión de los resultados alcanzados por los estudiantes de 15 años en la prueba de PISA 2006, a partir de variables vinculadas al contexto familiar del estudiante, el nivel socioeconómico de sus compañeros, y características del centro educativo al que asiste. Más aún, estiman indicadores de la desigualdad de los resultados y en qué magnitud

incide cada tipo de factores en la misma. En cuanto a los resultados, encuentran que en ambos países el contexto socioeconómico del estudiante tiene un efecto significativo y el contexto socioeconómico de los compañeros tiene un efecto aún mayor. En lo que respecta a las variables escolares, las autoras concluyen que para los dos casos se obtendría una reducción de la desigualdad si estas variables se distribuyeran homogéneamente.

Por otro lado, el trabajo de investigación de Da Rocha, Martínez y Rímoli (2010) estudia los principales factores asociados a la enseñanza primaria pública que afectan el desarrollo y adquisición de las competencias de los estudiantes, utilizando como herramienta la estimación de una Función de Producción Educativa (FPE) y como base las pruebas PISA 2006. Las estimaciones paramétricas de la FPE las realizan por el método de Heckman en dos etapas de modo de corregir el posible sesgo de selección que presenta la base PISA. La conclusión más importante que surge del documento es que los resultados obtenidos en las pruebas PISA 2006 por los alumnos que estudiaron en escuela pública están explicados o determinados por la interacción de variables asociadas al individuo y su entorno familiar, a variables referidas al centro educativo de enseñanza media y a variables vinculadas a su período escolar primario.

Por último, la investigación de Cerdan-Infantes y Vermeersch (2007) estima el impacto del programa de Escuelas de Tiempo Completo (ETC) instalado en Uruguay desde principios de la década de los 90', sobre los logros académicos de estudiantes de sexto año de escuela. Para esto utilizan tres bases de datos: La Evaluación Nacional de Aprendizajes realizada a estudiantes de 6° año de

escuela para los años 1996, 1999 y 2002 (evalúan los conocimientos de los niños en matemática y lenguaje e incorpora preguntas para el estudiante, el jefe de familia, la maestra de clase y el director de la escuela.), el Censo Nacional de Escuelas, e información especial recabada por ANEP.

El objetivo de su análisis es identificar el efecto de los centros convertidos a ETC en los rendimientos de los estudiantes. En particular, como la exposición al programa para los niños que tomaron las pruebas varía entre 0 y 6 años (1996 y 2002), les interesa conocer los cambios en los resultados de las pruebas asociados a los años de exposición al programa.

La estimación del efecto anterior puede tener problemas de sesgo si no se consideran todas las características que se utilizan para incluir a las escuelas en el programa. Por lo general, las mismas son seleccionadas de acuerdo al contexto socioeconómico, pero pueden existir características inobservables del centro que estén incorporadas en el criterio de selección.

Para minimizar el problema de sesgo anterior, los autores utilizan dos metodologías: Efectos Fijos y *Propensity Score Matching* (PSM). Con la primera, logran controlar las características observables e inobservables, invariantes en el tiempo, de las escuelas. Incluyen además variables dummies para los años estudiados, de modo de controlar cambios inobservables comunes a todas las observaciones en un año particular, características de los estudiantes y de los maestros. La única fuente potencial de sesgo que se deriva de esta especificación proviene de características de las escuelas que varían en el tiempo.

Para solucionar los problemas del método anterior, incorporan la metodología de PSM. Construyen para esto grupos de control utilizando distintos métodos de emparejamiento. Al no existir una regla explícita para la aplicación del programa de ETC en las escuelas, los autores utilizan un modelo PSM para identificar aquellas escuelas que dadas sus características observables, tienen un similar rango de probabilidad de participar en el modelo de ETC. De este modo, estiman tanto para las escuelas de tiempo completo como para las que no lo son, la probabilidad de participar en el programa, y utilizan la probabilidad predicha para formar tres grupos de comparación alternativos.

De la comparación de estos grupos, encuentran que los estudiantes en escuelas desfavorecidas fueron mejorando sus resultados en las pruebas luego de cada año de exposición al programa. Por lo tanto, el programa de Escuelas de Tiempo Completo, parece haber afectado positivamente los logros académicos de los estudiantes, especialmente en las escuelas más desfavorecidas.

III. MARCO TEORICO

La pregunta que da origen a esta investigación refiere a los factores que tienen mayor influencia sobre el rendimiento académico de los estudiantes de 15 años en el Uruguay, y si los mismos corresponden a variables propias del adolescente y su contexto sociocultural, o al centro educativo al cual concurren.

Para responder esta pregunta buscamos en primera instancia estudiar la importancia de la educación en la economía, lo cual se realiza a través de una aproximación a la teoría del capital humano, y en segundo lugar requiere del estudio de cuáles son esos factores que influyen en el rendimiento académico de los estudiantes. Para esto último, se realiza una sistematización bibliográfica que busca cubrir los distintos enfoques teóricos que han surgido respecto a las variables que influyen en la educación.

Teoría del capital humano

Los primeros estudios empíricos sobre los determinantes del crecimiento económico encuentran que el factor capital y trabajo (considerado como horas trabajadas) no llegan a explicar el 20 por ciento del crecimiento económico. El trabajo de Solow (1957) desarrollado en los Estados Unidos (período 1909-1949), obtiene que sólo el 12,5 por ciento del crecimiento económico de ese período se puede atribuir a la acumulación de capital y trabajo, siendo el 87,5 por ciento restante debido al progreso técnico. Esto genera un gran interés por determinar qué está detrás del factor residual y qué factores determinan el progreso técnico.

Ante la incapacidad teórica por conocer el origen del crecimiento económico del modelo de crecimiento tradicional neoclásico, es que surge la teoría del capital humano. Ya en los trabajos de Adam Smith y John Stuart Mill se encuentran asociaciones entre educación definida en sentido amplio, y crecimiento económico. No obstante, el punto de partida formal se da en la década del sesenta con los estudios e investigaciones de Mincer (1958), Schultz (1961), Becker (1962) y Denison (1962) en sus teorías del capital humano y la función de producción en educación.

La teoría del capital humano, que surge con los aportes de estos autores, considera la educación como una inversión que realizan los individuos y que les permite aumentar su dotación de capital (humano). Esta inversión aumenta su productividad, y al desarrollarse dentro del marco neoclásico de mercados competitivos, también sus ingresos futuros. Se establece de este modo una relación causal entre educación, productividad e ingresos, de forma que un aumento de la educación producirá un mayor nivel de renta y un mayor crecimiento económico (Oroval y Escardíbul, 1998).

En 1961, Schultz (1961) propone calcular los rendimientos de la inversión en educación. Su trabajo para medir la contribución de la educación al crecimiento sugiere que una proporción sustancial de la tasa de crecimiento del producto en los Estados Unidos se debe a la inversión en educación.

Denison (1974) estudia el crecimiento económico de los Estados Unidos entre los años 1929 y 1969 e intenta aislar la importancia de la educación y el progreso técnico en el crecimiento de la renta nacional. Su objetivo era calcular qué parte del crecimiento del producto era producido por el incremento del

capital y el trabajo, y qué parte quedaba como residuo no explicado de dicho crecimiento. Siguiendo la teoría del capital humano, este autor da a la educación un papel de bien de inversión y calcula la contribución de la misma al crecimiento económico suponiendo que los mayores ingresos de los trabajadores más educados son una medida de su mayor productividad debido a su mayor educación. De este modo encuentra que efectivamente los niveles salariales más altos son consecuencia de la mayor educación de los trabajadores (Oroval y Escardíbul, 1998).

Por su parte Becker (1964) analiza los retornos privados y sociales que genera la educación, y argumenta que los retornos sociales excederán los retornos privados por los efectos positivos de la educación sobre aspectos como el crimen, la salud, la mejora en la participación ciudadana, y en el crecimiento y productividad de la economía (Hanushek y Woessmann, 2007).

Por último y como crítica a la teoría del capital humano, cabe mencionar que en esta, la existencia de una correlación estadísticamente significativa y positiva entre la cantidad de años de escolaridad de una persona y los ingresos que obtiene, está establecida como una regularidad empírica. Esto lleva implícita la idea que sólo la educación formal influye en la formación de capital humano. Sin embargo, existen factores fuera de la educación formal que afectan directamente el desarrollo del conocimiento. Más aún, la calidad de la educación, que depende de la eficiencia del sistema educativo, de la calidad de los docentes, de la infraestructura y de otros factores, es sustantivamente más importante para el crecimiento económico que la cantidad de educación (Hanushek y Woessmann, 2007).

Evolución de la sociología de la educación

Hasta 1960 las teorías que dominaban la sociología de la educación eran las denominadas funcionalistas, las cuales tenían su base filosófica en el Positivismo. Emile Durkheim, creador de la teoría, sostenía que la función colectiva de la educación era adaptar al niño al medio social y convertirlo en un individuo útil dentro de la sociedad, a la vez que resaltaba la contribución del sistema educativo al mantenimiento del orden social. Según esta teoría, la función de la educación era preparar a las personas para el medio particular al que están destinados, teniendo a la escuela y a la familia como instituciones esenciales para ejercer la función de la reproducción social, justa y legítima.

En dicha década, se empieza a prestar más atención a las desigualdades de oportunidades educativas. En estos años el Departamento de Educación de los Estados Unidos encomienda a James Coleman y otros profesionales la realización de un estudio sobre la igualdad en la educación en el país (conocido más tarde como Reporte Coleman). Se suponía que el mismo documentaría lo que la mayoría de la gente asumía como cierto, que los niños de menor nivel socioeconómico tenían bajos rendimientos educativos debido a la falta de recursos en las escuelas. Sin embargo, el reporte descubrió que las diferencias de recursos entre las escuelas no eran tan grandes y que el impacto de estos factores en el rendimiento académico de los niños era modesto comparado con el impacto del contexto familiar de los mismos (Gamoran y Long, 2006).

La conclusión más controversial encontrada por Coleman (1966) fue que, una vez controlado el contexto familiar del niño, los recursos de las escuelas tenían

pocos efectos en los rendimientos educativos. Esto derribó la concepción de las escuelas como instrumentos igualadores y constituyó un disparador para la investigación sobre desigualdades educativas.

Al Reporte Coleman le siguieron una gran cantidad de estudios que buscaban reexaminar sus hallazgos. Entre estos se encuentran Averch et al. (1974), quienes no encontraron resultados consistentes referentes a los recursos educativos. La mayoría de estos estudios encontraron que el efecto del contexto familiar era mucho mayor al efecto de la escuela. Para los 90' se habían realizado cientos de estudios utilizando la función de producción educativa.

Una de las limitaciones detectadas por el Reporte Coleman consistía en que en los estudios para Estados Unidos sólo un 20 por ciento de la variación en los resultados académicos de los estudiantes se encontraba entre escuelas, mientras que el 80 por ciento restante ocurría dentro de las mismas. Esto podía implicar que el efecto de las escuelas fuera pequeño y difícil de detectar. En los países en desarrollo por su parte, la variación del rendimiento entre escuelas es mayor, por lo que los efectos de la escuela podían ser más importantes y fáciles de identificar (Casassus et al., 2001).

Varios estudios internacionales demostraron lo anterior, sugiriendo que los resultados encontrados por Coleman sobre los bajos efectos de los recursos educativos una vez controlado el contexto familiar del niño sólo se aplicaba a países que habían pasado un cierto nivel de recursos escolares básicos.

Por otro lado, en la década del setenta surgen las teorías de la reproducción, las cuales entendían que la educación era un medio a través del cual se reproducían las relaciones sociales desiguales. Bourdieu y Passeron (1971) presentan una teoría del sistema de enseñanza como violencia simbólica. Según estos autores, el individuo hereda un capital cultural de acuerdo a la clase social a la que pertenece su familia. La escuela se encarga de legitimar y reproducir el capital cultural de las clases dominantes, dejando en desventaja a aquellos estudiantes que provienen de otras clases sociales y tienen un menor contacto con ese capital cultural que es legitimado (Giroux, 1985).

Las escuelas eficaces

Desde la aparición del Reporte Coleman y por casi una década, una gran cantidad de investigaciones habían fallado en el intento de encontrar “efectos propios de la escuela” sobre el aprendizaje. Una vez controlado el contexto familiar del niño, las escuelas parecían hacer poca diferencia en los aprendizajes.

Sin embargo, las teorías tampoco podían explicar cómo algunas escuelas lograban que sus alumnos tuvieran altos niveles académicos a pesar de que proviniesen de medios sociales pobres y con bajos estímulos culturales. La investigación sobre los determinantes sociológicos del aprendizaje desató un cambio profundo en el pensamiento científico cuando a fines de los años setenta fueron publicados los primeros artículos fundacionales sobre “las escuelas eficaces” (Fernández Aguerre, 2004).

Con los trabajos de Rutter et al. (1979), Edmonds (1979), Brookover (1979) y el Nuevo reporte Coleman (1982), aparece un nuevo enfoque dentro de la

sociología de la educación, el cual se centra en el “efecto de la escuela” (Fernández Aguerre, 1999).

Una primera versión de este enfoque surge de los estudios de Edmonds (1979) y Rutter et al. (1979). La eficacia, según estos autores, se encontraba en la capacidad de las escuelas para lograr buenos resultados en las áreas básicas del conocimiento a pesar del contexto social de sus estudiantes. Estas investigaciones originaron una teoría de la organización social de la escuela, la cual identifica una cierta configuración organizacional que caracteriza a las “escuelas eficaces” en aspectos relacionados a la toma de decisiones, la estabilidad, el clima organizacional y el currículum (Méndez y Zerpa, 2009).

A partir del segundo reporte Coleman (1982) surge una segunda versión de “la escuela eficaz”. El mismo encontró que las escuelas privadas, y específicamente las católicas, presentaban mejores resultados que las públicas, una vez controlado el nivel sociocultural de las familias. Con esto, se empezó a estudiar los resultados poniendo énfasis en el ámbito institucional. Según este enfoque, es el mecanismo de competencia el que lleva a las escuelas privadas a mejorar la calidad de la educación que brindan, logrando así una mayor eficacia.

Uruguay: equidad y calidad de la educación básica

El caso específico de Uruguay puede ser estudiado a partir del marco conceptual desarrollado por Vegas y Petrow (2007). Este analiza el aprendizaje como el resultado de interacciones entre los niños, el centro educativo y las instituciones que determinan la política educativa.

Según los autores, los alumnos llegan al establecimiento escolar con una serie de activos y conductas que influyen en su aprendizaje y que son adquiridos de sus familias y del contexto sociocultural en el que se desarrollan. Por otro lado, las escuelas están dotadas de ciertos activos y conductas que afectan lo que le entregan al estudiante y que son influidos por los docentes, las condiciones específicas del establecimiento y la gestión e influencia de los equipos administrativos. Finalmente, las instituciones que fijan la política educativa y administran los establecimientos pueden afectar los aprendizajes.

Para facilitar el análisis, Vegas y Petrow (2007) descomponen el marco conceptual entre aspectos relacionados a los estudiantes, los establecimientos y las instituciones:

Por el lado de los estudiantes, los activos y conductas de los alumnos están influenciados por factores específicos a su realidad (habilidades innatas, desarrollo cognitivo temprano, estimulación, etc.), factores específicos del hogar (recursos, educación y empleo de los padres, estatus socioeconómico) y factores comunitarios (tales como valores en torno a los niños y la educación e infraestructura de apoyo al aprendizaje).

Por el lado de los establecimientos educativos, la calidad de la infraestructura y los recursos, el esfuerzo y conocimiento de los profesores, la capacidad de gestión y el liderazgo de los directores, así como la manera de financiar los centros educativos pueden influir en el rendimiento de los estudiantes.

Por el lado de las instituciones, entre los factores que pueden afectar el aprendizaje están el nivel de administración, la capacidad de gestión, la

cantidad y calidad de la participación de los padres y las comunidades en los establecimientos, el grado de provisión privada, y la capacidad de que disponen las familias para seleccionar las escuelas a las que asisten sus hijos, tanto en el sector público como en el privado.

Según estos autores, la calidad de los procesos de aprendizaje responde a interacciones entre los estudiantes, el establecimiento y las instituciones del sistema educativo. Esto facilita el análisis de políticas vigentes y las posibles reformas que podrían tener efecto en la equidad y la calidad de la educación en el contexto de Uruguay y de otros países.

IV.DATOS

Para la realización de este trabajo fueron utilizadas las bases PISA correspondientes a los años 2006 y 2009, así como información del porcentaje de clases no dictadas en la enseñanza pública para el último año, aportados por el Consejo de Educación Secundaria.

El Programa Internacional para la Evaluación de Estudiantes (PISA por sus siglas en inglés) surge en el marco de la Organización para la Cooperación y Desarrollo Económicos (OCDE) y la UNESCO, para la elaboración de indicadores educativos y en respuesta a las demandas de los países miembros para disponer de datos fiables sobre los conocimientos y las aptitudes de sus estudiantes y el rendimiento de sus sistemas educativos.

Su principal objetivo es describir en qué grado los estudiantes de 15 años de edad, que están concluyendo la etapa de educación obligatoria, están preparados para afrontar los desafíos de la sociedad del conocimiento. Para esto, se busca observar las habilidades de los estudiantes a la hora de analizar, razonar y aplicar conocimientos en la resolución de problemas en materias claves (OCDE, 2009).

PISA mide el rendimiento de los alumnos en tres disciplinas: lectura, matemática y ciencia, a la vez que indaga respecto a sus motivaciones para aprender, la concepción que éstos tienen sobre las áreas evaluadas, sus estrategias de aprendizaje, su familia y las características del centro educativo al que concurre.

Se trata de un estudio comparativo internacional con ciclos de tres años, estando cada uno de estos enfocado especialmente en una de las áreas de estudio. La primera edición realizada en el 2000 tuvo su foco en lectura, mientras que la de 2003 y 2006 en matemática y ciencia respectivamente. En el 2009 el foco pasó a ser nuevamente lectura.

Todos los países miembros de la OCDE participaron de las pruebas, así como también ciertos países asociados. En la primera oportunidad se implementó en 43 países, 41 en el segundo ciclo (2003), 57 en el tercero (2006) y 67 en el cuarto (2009).

Los estudiantes que participan de la evaluación conforman una muestra representativa y son sorteados a partir de una muestra aleatoria de centros públicos y privados de todo el país. Se utiliza para su selección una técnica de muestreo estratificado en dos etapas. La primera resulta en una muestra aleatoria de centros educativos con estudiantes de 15 años, la cual debe ser de un mínimo de 150 centros por país. En la segunda etapa se genera una muestra de 35 estudiantes de 15 años, de cada uno de los centros seleccionados previamente, teniendo cada alumno la misma probabilidad de selección.

Las pruebas incluyen tanto preguntas de múltiple opción, como preguntas que requieren la elaboración de respuestas. Cada materia es evaluada a través de problemas con distintos niveles de dificultad, de modo de representar un indicador completo de las habilidades de los estudiantes.

Usando la teoría de respuesta al ítem, se asignan los resultados en cada materia, con una escala cuya media a nivel de países de la OCDE es de 500 puntos y su desviación estándar de 100 puntos (OCDE, 2009).

PISA trata además el rendimiento de los estudiantes en las pruebas como una variable aleatoria inobservable, con una distribución condicionada a su rendimiento en la prueba estandarizada, así como a la información sobre su entorno (OCDE, 2006). Como consecuencia, PISA no provee un único estimador del rendimiento, sino que para cada materia se obtienen cinco resultados de la distribución condicional de la variable llamados valores plausibles, que luego son incorporados en la estimación de estadísticos (ver anexo II.1).

En materia de resultados, Uruguay se ha mantenido relativamente estable en los tres ciclos que ha participado (ver cuadro A.I.1 con resultados promedio para el año 2006 y 2009). A pesar de que se ubica entre los países de mejor nivel de América Latina, sigue rezagado respecto a los promedios de puntaje obtenidos por los estudiantes de la OCDE (ANEP, 2009).

A sí mismo, existen asimetrías entre los resultados obtenidos por estudiantes uruguayos cuando se analiza la muestra según características de los individuos y de los centros educativos. Se desprende del cuadro A.I.1 para el año 2006, que las mujeres obtuvieron mejores resultados que los varones únicamente en lectura. Sin embargo, en el año 2009 las mujeres lograron además superar a los varones en ciencias. En cuanto al lugar de residencia, los estudiantes que habitan en pequeñas localidades obtuvieron en promedio peores resultados que los que habitan en grandes ciudades. En relación a los antecedentes

académicos, los resultados promedio en las tres disciplinas de los individuos que repitieron primaria fueron más bajos que los que no lo hicieron. Por otro lado, en cuanto al contexto familiar, los estudiantes cuyas madres no culminaron la enseñanza primaria presentaron un peor rendimiento en las tres asignaturas.

En relación al tipo de institución al cual concurren los individuos seleccionados, en el año 2006 el 78.77 por ciento asistía a centros de enseñanza media públicos, mientras que en el 2009 el porcentaje aumentó a 82.52 por ciento. Del cuadro A.I.2 se desprende que los resultados promedio de los estudiantes del sector público fueron inferiores a los obtenidos por aquellos que asistieron a centros privados.

La segunda base de datos utilizada corresponde a información de las inasistencias docentes, la cual surge de los datos relevados por el Consejo de Educación Secundaria (CES) durante el año 2009 y refleja el porcentaje de la cantidad de horas no dictadas en cada centro educativo público, sobre las que se debieron dictar. Estos datos, que nos fueron cedidos de manera exclusiva por el CES, si bien representan un aporte novedoso, tienen a su vez ciertas limitaciones. En primer lugar, los datos no discriminan las inasistencias por materia sino que refleja el porcentaje de horas no dictadas en el centro educativo en su conjunto, restringiendo su utilización. En segundo lugar, la información corresponde únicamente a inasistencias docente en los centros educativos públicos, lo cual nos obligó a centrar el estudio a este tipo de instituciones a la hora de incorporar esta variable al análisis. Por último, los datos de varios centros públicos fueron omitidos en la base por ser

considerados por el CES como atípicos, lo que puede estar distorsionando los resultados obtenidos a partir de la utilización de esta información.

Centrándonos en las variables de interés de este estudio, en el cuadro A.I.3 se presenta la distribución del tiempo curricular de clase a la semana en nuestro país, para cada una de las tres materias que evalúa la prueba. Como se puede observar, en el año 2006 matemática era la asignatura con mayor carga horaria, siendo sustituida en 2009 por ciencias. En promedio, los niños declararon en 2006 tener un 20 por ciento menos de clase de esta asignatura, situación que se revierte en la evaluación siguiente cuando pasan a tener un 72 por ciento más de carga horaria destinada a la misma. En cuanto a los desvíos, lectura es en ambos años, la que presenta la menor dispersión.

Por último, el cuadro A.I.4 nos aporta información sobre las inasistencias docentes a centros públicos en el año 2009. En promedio, los docentes no dictan un 14.31 por ciento de las clases que deberían dictar en un año académico, siendo el máximo en la muestra de 37.4 por ciento. En el cuadro A.I.5 además, se describe cómo es el comportamiento de esta variable en relación al tamaño de la locación, pudiéndose concluir que las pequeñas localidades presentan no sólo un menor promedio de clases no dictadas, sino además una menor dispersión de las mismas.

V. ESTRATEGIA EMPÍRICA

Estimación del Efecto del Tiempo de Clase

El objetivo central del presente trabajo es estudiar el efecto del tiempo de clase sobre los resultados de las pruebas PISA 2006 y 2009, buscando de este modo evaluar si más carga horaria en la educación media se traduce en mayores logros académicos de los estudiantes.

Siguiendo a Lavy (2010) partimos del modelo:

$$Y_{ijk} = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + (\varepsilon_j + \eta_k) + u_{ijk} \quad (1)$$

Donde Y_{ijk} es el resultado en la prueba del estudiante i en el centro j en la materia k , H_{kj} es el tiempo de clase de la materia k en el centro j , X_{ij} es un vector de características del estudiante i en el centro j , S_j es un vector de características del centro j , ε_j y η_k representan características inobservables de la escuela y la materia respectivamente, u_{ijk} es el término de error restante y μ_i es el efecto fijo del estudiante que captura el contexto familiar, habilidades, motivación, y otras características no cognitivas del niño.

El parámetro clave a estimar en nuestro modelo es γ asociado a la variable H_{kj} (a la cual llamamos *promediomin*). Esta representa el tiempo de clase regular a la semana (expresado en minutos) que tiene cada estudiante en cada una de las materias de las pruebas. La misma se construyó a partir de respuestas de los alumnos a preguntas del cuestionario PISA (referentes al tiempo de clase en cada disciplina), y se calculó como el promedio de minutos

de clase por materia, por centro, que tiene cada niño, buscando de este modo neutralizar los valores atípicos producto de respuestas erróneas de los individuos.

Mínimos Cuadrados Ordinarios

En primer lugar se estimó la regresión antes planteada mediante el método de Mínimos Cuadrados Ordinarios. El mismo permite obtener estimadores de los coeficientes asociados a las variables e inferir sobre sus verdaderos valores, pero para esto deben cumplirse ciertos supuestos, entre los que se encuentra la correcta especificación del modelo.

En relación a esto, la base de información utilizada presenta una limitación, y es que sólo dispone de datos relevados para una muestra de los estudiantes de 15 años que asisten actualmente al sistema educativo formal post-primaria. Este problema de inobservabilidad de los estudiantes desertores del sistema educativo en la muestra de datos, refiere a una situación de truncatura incidental y potencialmente introduce un sesgo de selección si se pretende estimar los parámetros poblacionales a partir del modelo presentado anteriormente (Llambí, Perera y Messina, 2009).

Estimación con corrección de sesgo

En segundo lugar, corregimos el sesgo de selección aplicando el método en dos etapas de Heckman (1976) explicado en el anexo II.3. En este caso, la ecuación de regresión válida para las observaciones muestrales (las cuales corresponden a los individuos que asisten al sistema educativo post-primaria) se transforma en:

$$Y_{ijk} = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + \rho \hat{\lambda}_i + w_{ijk} \quad (2)$$

Donde $w_{ijk} = v_{ijk} + \rho(\lambda_i - \hat{\lambda}_i)$

y $\lambda_i = \lambda(Z_i, \omega) = \frac{\phi(Z_i, \omega)}{\Phi(Z_i, \omega)}$ es el Inverso del Ratio de Mills (siendo Z un vector de

variables que determinan la asistencia)

Mediante el método de Heckman, llegamos a la obtención de estimadores consistentes y asintóticamente normales.

En lo que refiere a nuestra investigación, se decidió calcular el Inverso del Ratio de Mills (tanto para el 2006 como para el 2009) utilizando los resultados obtenidos por Méndez y Zerpa (2009).

Para la estimación del $\lambda(Z_i, \hat{\omega})$, las autoras debieron identificar en primera instancia la asistencia (o inasistencia) al sistema educativo formal post-primaria (y las variables que la determinan), información que obtuvieron de la Encuesta Nacional de Hogares Ampliada (ENHA) 2006. Para estimar la ecuación (2) es necesario replicar el valor de $\lambda(Z_i, \hat{\omega})$ para los adolescentes presentes en la muestra de PISA, por lo que debieron tomar de la ENHA variables Z que se encontraran también en los cuestionarios PISA. Es decir, debido a que es necesario trabajar con dos muestras diferentes (las de la encuesta de hogares y las de PISA), para incorporar dentro de la ecuación (2) el sesgo producido por la probabilidad de estar o no en la muestra, de manera de obtener estimadores insesgados de los parámetros de la ecuación original, es necesario contar con las mismas variables Z en ambas encuestas (Méndez y Zerpa, 2009).

Para la estimación con los datos PISA 2006, utilizamos los coeficientes $\hat{\omega}$ estimados por las autoras, y a partir de los mismos obtuvimos nuestro propio $\hat{\lambda}_i$. Así, ya con la estimación del Inverso del Ratio de Mills, procedimos a estimar los parámetros de la ecuación (2) obteniendo estimadores consistentes y asintóticamente normales. Esto fue posible para el modelo del año 2006, ya que las autoras basan su estudio para ese año.

El desafío mayor se nos presentó por tanto, a la hora de usar la base PISA 2009. Una primera opción, fue recurrir a la Encuesta Nacional de Hogares Ampliada 2009 y estimar nuevamente el Inverso del Ratio de Mills. Esto no fue posible debido al cambio en las preguntas que existe de una encuesta a la otra, no pudiendo así reproducir un $\hat{\lambda}_i$ que fuera comparable entre 2006 y 2009. Una segunda opción, la que finalmente se llevó adelante, fue asumir que los coeficientes $\hat{\omega}$ permanecieron constantes en dicho período. Si bien el hecho de que un niño asista o no a la educación formal se ve influido por factores macroeconómicos y los mismos han cambiado en los últimos años, estos cambios son captados por el vector de variables Z , avalando la utilización nuevamente de los coeficientes obtenidos por Méndez y Zerpa (2009). De esta manera, el problema del sesgo en las bases PISA tanto 2006 como 2009 se resuelve de la misma forma.

Incorporación de variable Inasistencias

Como tercer paso, se profundizó en el estudio del efecto del tiempo de clase sobre los logros académicos, mediante la introducción de la variable *inasistencias* a la regresión original. Tal como fue explicado, refleja el porcentaje de la cantidad de horas no dictadas en cada centro educativo, sobre

las que se debieron dictar. La base presenta información del año 2009 únicamente para los liceos públicos, por lo que debimos restringir el estudio a este tipo de centros educativos.

Para esto, tomamos las pruebas PISA 2009 quedándonos únicamente con la información de los estudiantes que concurren a liceos públicos, e incorporamos los datos de inasistencias a la base.

La aproximación a este modelo fue realizada de la misma forma que se desarrolló anteriormente, comenzando por una estimación por MCO y luego corrigiendo el sesgo de selección mediante el método de Heckman.

A partir de la introducción de esta nueva variable explicativa (*inasistencias*), intentamos identificar no sólo el efecto de los minutos de clase que debieron ser dictados (*promediomin*) sobre el rendimiento de los estudiantes que asisten a centros públicos, sino también el efecto de las horas no dictadas.

Estimador Intra-individuo

Como señala Lavy (2010), el problema de los métodos de estimación antes planteados es que no permiten inferir causalidad debido a posibles problemas de selección y endogeneidad. Los efectos de factores inobservables y correlacionados confunden generalmente el efecto del tiempo de clase en el resultado académico. Situaciones de este tipo pueden surgir si la selección de estudiantes en los centros educativos está influida por los recursos de los mismos, o si existe correlación entre el tiempo de clase y otras características de los centros que pueden afectar los logros académicos.

Para sobreponerse a estos dos problemas, el autor utilizar la variación intra-alumno del tiempo de clase en las materias. De este modo, examina si las diferencias de los resultados de un alumno entre materias están asociadas sistemáticamente a diferencias en el tiempo de clase. La idea central es que las características de los individuos y el ambiente del centro educativo son las mismas para cada estudiante en las tres materias, mientras que la carga horaria varía entre las asignaturas.

Siguiendo a Lavy, presentamos como cuarto paso, estimadores intra-alumno del efecto del tiempo de clase en los resultados individuales de las pruebas (ver anexo II.4), partiendo del siguiente modelo:

$$Y_{ijk} = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + (\varepsilon_j + \eta_k) + \rho \lambda(Z_i \hat{\omega}) + u_{ijk} \quad (3)$$

Para estimar el efecto causal del tiempo de clase de forma consistente a la presencia de heterogeneidad individual invariante en el tiempo, una alternativa es restar a la ecuación original las medias de las variables que presentan variación intra-individuo, quedando la regresión a estimar como:

$$\tilde{Y}_{ijk} = \gamma \tilde{H}_{kj} + (\eta_k - \bar{\eta}) + \tilde{u}_{ijk} \quad (4)$$

Donde $\tilde{Y}_{ijk} = Y_{ijk} - \bar{Y}_{ij}$, $\tilde{H}_{kj} = H_{kj} - \bar{H}_j$ y $\tilde{u}_{ijk} = u_{ijk} - \bar{u}_{ij}$

Mediante el uso de la variación intra-alumno no sólo se logra controlar los efectos fijos de los estudiantes, sino también los del centro educativo y el término asociado al sesgo de selección (inversa del ratio de Mills). Por lo que la utilización de este método de estimación permite controlar varias fuentes

generadoras de sesgo relacionadas con las características inobservables de los centros, los estudiantes o sus interacciones.

Al igual que Lavy debemos hacer tres observaciones importantes referentes a la estrategia implementada. En primer lugar, el método de identificación asume que el efecto del tiempo de clase es el mismo para las tres asignaturas, por lo que γ no puede variar entre materias. Siguiendo al autor, se estudió cuán restrictivo es el supuesto comparando las estimaciones obtenidas al tomar subgrupos de las materias. Las conclusiones obtenidas se detallan en el capítulo de resultados. En segundo lugar, los logros académicos en cierta materia están influidos también por el tiempo que se le dedica a otras asignaturas (por ejemplo, el tiempo de clase de Inglés puede influir en los resultados de las pruebas en lectura). Por último, puede ocurrir que los alumnos que tienen mayores destrezas en matemática elijan centros educativos con mayor carga horaria de esa materia, sesgando así los resultados. Lavy encuentra que esto es poco relevante en la base PISA, ya que los centros educativos a la edad de 15 años no tienen por lo general programas diferenciados.

Análisis de Descomposición de varianza

Como segundo objetivo del trabajo, y de modo de enriquecer el análisis de los datos, decidimos realizar un análisis de descomposición de varianza.

Una de las ventajas de los modelos de heterogeneidad inobservable con componentes de error, es que nos permiten ir más allá de solucionar los problemas planteados por la regresión MCO, ofreciendo la posibilidad de

estimar qué proporción de la varianza total de la variable dependiente es atribuible a cada uno de los niveles implicados en el análisis.

Así, el foco de estudio de este trabajo se mueve desde la búsqueda del efecto causal entre el tiempo curricular y el rendimiento de los alumnos, al análisis de cuánto de la variabilidad de los resultados puede explicarse por diferencias entre los estudiantes.

Tal como se explica en Arellano (2003) y es aplicado en Sanroman (2010) es posible partir de una especificación sencilla del modelo:

$$y_{ij} = \mu + \eta_i + v_{ij} \quad (5)$$

Donde i representa los individuos, j el tiempo, μ el intercepto, η las características inobservables de los alumnos y v_{ij} un componente que varía aleatoriamente entre individuos y materias.

Bajo los supuestos:

supuesto n°1: $\eta_i \sim \text{iid}(0, \sigma_\eta^2)$

supuesto n°2: $v_{ij} \sim \text{iid}(0, \sigma^2)$

supuesto n°3: η_i y v_{ij} son independientes uno de otro.

Es posible expresar nuestro parámetro de interés como:

$$\rho = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma^2} \quad (6)$$

En lo que respecta a esta investigación puntual, podemos afirmar que ρ representará el porcentaje de la variabilidad total que se explica por las diferencias entre los individuos que se mantienen constantes a lo largo de las tres materias. Además, el análisis de descomposición de varianza se llevó a cabo en tres pasos, estudiándose el efecto sobre la varianza total de cada etapa como la diferencia entre los valores obtenidos para cada $\hat{\rho}$ respectivamente. Así mismo, el estudio se realizó para varios países, en los años 2006 y 2009.

El análisis comienza con la estimación de una ecuación sólo con la constante (c), a la que se le agregó en primer lugar un vector de variables del centro (x_{ij}), en segundo lugar un vector de variables del niño (z_{ij}) alcanzándose así la tercera ecuación de estudio; con v_{ij} , w_{ij} y u_{ij} los términos de error.

Primera ecuación: $y_{ij} = c + \eta_i + v_{ij}$

Segunda ecuación: $y_{ij} = c + \beta x_{ij} + \eta_i + w_{ij}$

Tercera ecuación: $y_{ij} = c + \beta x_{ij} + \delta z_{ij} + \eta_i + u_{ij}$

VI.RESULTADOS

Estimación del Efecto del Tiempo de Clase

A partir de la aplicación de los distintos métodos explicados en el capítulo anterior, procedimos a estimar el efecto del promedio de minutos de clase semanales por materia, sobre el rendimiento de los estudiantes en las pruebas PISA. En el cuadro 6.1 se presentan los resultados obtenidos en la estimación del coeficiente asociado a la variable *promediotmin* para las distintas especificaciones del modelo.

Cuadro 6.1: Resultados de la estimación del efecto del tiempo en clase sobre el rendimiento de los estudiantes en las pruebas PISA 2006 y 2009.

Año	Matemáticas		Ciencia		Lectura		Intra-individuo
	MCO (I)	HECK (II)	MCO (I)	HECK (II)	MCO (I)	HECK (II)	
2006	0.25 (0.08)	0.25 (0.08)	0.13 (0.05)	0.13 (0.05)	0.14 (0.11)	0.12 (0.11)	0.01 (0.04)
2009	-0.05 (0.11)	-0.02 (0.11)	0.11 (0.03)	0.10 (0.03)	0.02 (0.07)	0.03 (0.07)	-0.01 (0.02)

Las cifras entre paréntesis corresponden al desvío estándar.

Los cuadros con los resultados completos pueden encontrarse en el anexo III.

- (I) Representa la estimación por MCO del modelo sin corrección del sesgo de selección. Se incluyen además en esta regresión como variables explicativas: *mujer*, *iestatusesco*, *tamañooc*, *privado*, *tamañoocentro*, *escasezdoc*, *propdoctitulados*, *propdocuniv*, *grado*, *dummies por tipo de centro primaria y ciclo básico*. En 2009 no se incorporan estas dos últimas variables.
- (II) Representa la estimación del modelo por el método bietápico de Heckman. Se incluyen como explicativas las mismas variables detalladas en el punto anterior y la estimación del Inverso del ratio de Mills.

Tal como fue mencionado, la estimación por Mínimos Cuadrados Ordinarios de nuestro modelo de partida presenta un sesgo debido a la naturaleza no aleatoria de la muestra PISA (Llambí y Perera, 2008). Por este motivo, en lo que sigue, los resultados presentados son los obtenidos en la estimación del modelo que incorpora el sesgo de selección a través del método en dos etapas de Heckman (modelo II del cuadro 6.1).

Para el año 2006, los coeficientes estimados para las tres materias resultaron positivos pero no siempre estadísticamente significativos. Tanto para matemáticas como para ciencias, *promediomin* resultó significativa a un nivel de confianza del 95 por ciento. Por lo tanto, en línea con los resultados obtenidos por Lavy (2010) en la estimación MCO de su modelo, es posible afirmar que el aumento en una unidad del promedio de minutos en estas asignaturas, dado todo lo demás constante, genera una mejora de 0.25 puntos y 0.13 puntos en los resultados de las pruebas respectivamente. Sin embargo, *promediomin* no es significativa para explicar el resultado en la prueba de lectura. Es decir, no hay evidencia suficiente para afirmar que el efecto de esta variable en el resultado de la prueba es positivo.

Por otro lado, para el año 2009 *promediomin* es significativa únicamente para explicar los resultados obtenidos en ciencias. Dado todo lo demás constante, el aumento en una unidad del promedio de minutos a la semana de esta asignatura genera una mejora en 0.10 puntos en el resultado de esta prueba. Respecto al resto de las materias, nuevamente no hay evidencia suficiente para concluir sobre el efecto de la variable de estudio en el resultado de las pruebas.

La última columna del cuadro 6.1 presenta el estimador intra-individuo del efecto en estudio para el año 2006 y 2009. Este método, que se introduce como solución al sesgo generado en la estimación por Mínimos Cuadrados Ordinarios, se justifica en nuestro modelo por la presencia de factores inobservables que tienen influencia sobre el rendimiento y están relacionados

tanto a características del estudiante, como del centro educativo y de las materias.

Para la aplicación del método, siguiendo a Lavy (2010), debimos suponer que el efecto del tiempo de clase es el mismo para las tres materias. Es decir, asumimos que la productividad de las clases no varía por asignatura (β es el mismo para las tres disciplinas). Además, estudiamos cuan restrictivo es el supuesto, comparando las estimaciones obtenidas a partir de tres submuestras que incluyeron combinaciones de sólo dos de las asignaturas. En los cuadros A.III.7 y A.III.8 se presentan estos resultados para los años 2006 y 2009 respectivamente.

En ambos años las estimaciones para matemática y ciencias fueron superiores que las encontradas para matemática y lectura, que a su vez fueron mayores a las de ciencia y lectura. Esto podría indicar que la productividad media del tiempo de clase es mayor en matemática que en ciencia y lectura. No obstante no podemos extraer conclusiones contundentes, ya que las mismas no resultaron significativas a un nivel del 5 por ciento. De todos modos, el promedio de las tres estimaciones obtenidas para las tres submuestras es muy parecido al coeficiente estimado en la muestra original, lo que permitiría validar el uso del supuesto antes mencionado.

En cuanto a los resultados obtenidos para la variable de interés, *promediomin*, se observa en el cuadro 6.1 que la misma no es significativa para explicar el rendimiento de los alumnos en las pruebas, en ninguno de los años de estudio. A pesar de que se obtuvieron estimadores insesgados y consistentes, no encontramos evidencia suficiente para concluir sobre la existencia de una

relación causal entre el tiempo en clase y los logros académicos. Por lo tanto, no es posible descartar que las diferencias intra-individuo en la cantidad de horas a la semana por materia tienen efecto alguno sobre los resultados obtenidos en las pruebas.

Incorporación de variable Inasistencias

En el cuadro 6.2 se presentan los resultados obtenidos para la estimación del modelo por MCO y del modelo bietápico de Heckman, luego de agregar como variable explicativa a las inasistencias docentes en los centros educativos públicos (*inasistencias*). Al no poseer el dato diferenciado por materia, no fue posible incluirla en la estimación intra-individuo, incorporándola únicamente como un regresor en el modelo estimado por MCO y en el modelo en dos etapas de Heckman.

Cuadro 6.2: Resultados de la estimación del efecto del tiempo en clase y las inasistencias docentes sobre el rendimiento de los estudiantes en las pruebas PISA 2009

Año	Matemáticas		Ciencia		Lectura	
	MCO (I)	HECK (II)	MCO (I)	HECK (II)	MCO (I)	HECK (II)
<i>promediomin</i>	0.14 (0.16)	0.14 (0.13)	0.04 (0.04)	0.04 (0.04)	-0.11 (0.09)	-0.11 (0.09)
<i>inasistencias</i>	-0.19 (0.51)	-0.25 (0.15)	-0.49 (0.29)	-0.57 (0.29)	0.12 (0.41)	0.13 (0.38)

Las cifras entre paréntesis corresponden al desvío estándar.

Los cuadros con los resultados completos pueden encontrarse en el anexo III.

(I) Representa la estimación por MCO del modelo sin corrección del sesgo de selección. Se incluyen además en esta regresión como variables explicativas: *mujer*, *iestatuseco*, *tamañoloc*, *privado*, *tamañocentro*, *escasezdoc*, *propdoctitulados*, *propdocuniv*, *grado*.

(II) Representa la estimación del modelo mediante el método bietápico de Heckman. Se incluyen como explicativas las las mismas variables detalladas en el punto anterior y la estimación del Inverso del ratio de Mills.

Se observa que nuestra variable de interés, *inasistencias*, resulta significativa para explicar los resultados de las pruebas de ciencias y matemática, a un nivel

de significación del 5 y 10 por ciento respectivamente. En particular, se infiere un efecto negativo sobre las competencias antes mencionadas, que oscila entre 0,57 para ciencias y 0,25 para matemática.

Es decir, el porcentaje de clases perdidas por ausencias docentes en el año afecta negativamente al rendimiento de los estudiantes de centros públicos en las pruebas de ciencias y matemática.

Por su parte, *promediomin* resulta no significativa en este modelo para explicar el rendimiento de los estudiantes en las tres competencias evaluadas.

Análisis de los resultados de descomposición de varianza

En esta sección se interpretan los resultados obtenidos del análisis empírico de la descomposición de la varianza. Considerando que “teóricamente, los aprendizajes no solamente se ven influidos por las características individuales y sociofamiliares de los alumnos, sino también por el tipo de recursos culturales, económicos y lingüísticos que son activados en las interacciones que se verifican en el proceso de enseñanza y aprendizaje, los grupos de pares y otros eventos cotidianos de la vida escolar” (Férrandez y Blanco, 2004); se cree que los mismos son un aporte interesante a esta investigación.

Basándonos en el planteo anterior, y apoyándonos en la estratificación del marco conceptual que realizan Vegas y Petrow (2007) (desarrollado en el capítulo de marco teórico) se tomó la decisión de aproximarnos al análisis de la varianza a través de factores asociados a características del niño y de su centro educativo. En lo que respecta al nivel de las instituciones, no fue posible incluirlo debido a limitaciones en nuestra base de datos. El sistema educativo

uruguayo a nivel público, se caracteriza por la falta de autonomía de sus instituciones a la hora de tomar decisiones; y por programas, políticas y presupuestos que son aplicados de forma general en todo el sistema. Por lo tanto, se consideró que la inclusión de variables institucionales no aportaría al análisis uruguayo, y distorsionaría su comparación con otros países (que sí muestran mayor autonomía).

Se busca estudiar de qué forma influyen los dos grupos de factores sobre nuestro parámetro de interés $\hat{\rho}$ (porcentaje de variabilidad total que se explica por las diferencias entre los individuos que se mantienen constantes a lo largo de las tres materias). Para esto, se decidió hacer una aproximación a los datos por dos vías. Por un lado, se comparan los resultados de $\hat{\rho}$ obtenidos por diez países con diferentes desempeños en los rankings PISA (Finlandia, Nueva Zelanda, Estados Unidos, España, Italia, Chile, Uruguay, México, Argentina y Colombia), dándole especial importancia a aquellos que son de la región y/o comparten “características educativas” con nuestro país. Por otro lado, se estudian dichos resultados para los años 2006 y 2009, últimas dos pruebas PISA disponibles.

Cuadro 6.3: Resultados de la descomposición de varianza.

Países	2006					2009				
	Ranking	Puntaje	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$	Ranking	Puntaje	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$
Finlandia	1	566	0.79	0.79	0.76	3	536	0.82	0.82	0.80
NZ	7	530	0.85	0.84	0.80	7	521	0.88	0.87	0.83
EEUU	29	489	0.87	0.87	0.83	17	500	0.89	0.87	0.85
España	31	488	0.78	-	-	33	481	0.82	0.81	0.78
Italia	36	475	0.75	0.70	0.68	29	486	0.84	0.81	0.78
Chile	40	438	0.73	0.66	0.62	44	449	0.81	0.78	0.76
Uruguay	43	428	0.66	0.61	0.48	47	426	0.80	0.75	0.71
México	49	410	0.75	0.68	0.66	48	425	0.83	0.79	0.76
Argentina	51	391	0.71	0.61	0.57	58	398	0.85	0.82	0.78
Colombia	53	388	0.59	0.53	0.49	51	413	0.77	0.74	0.71

Los valores estimados de RHO1 corresponden a un modelo sólo con constante, los de RHO2 a un modelo con constante y variables del centro, y por último los de RHO3 al modelo completo: constante más variables del centro y del niño.

Una primera conclusión, que se hace evidente al observar el cuadro 6.2, es que existe una fuerte relación entre el puesto ocupado en el ranking PISA y el valor de $\hat{\rho}$ obtenido: a mayores puestos en el ranking, más alto el valor de $\hat{\rho}$. De esto, es posible afirmar que aquellos niños que pertenecen a países cuyo puntaje en las pruebas supera la media de la OCDE (500 puntos) tienden a desempeñarse bien en todas las materias cuando lo hacen en una de ellas; obteniendo así rendimientos poco dispersos entre pruebas. Por el contrario, los alumnos de países con menores valores de $\hat{\rho}$ (con puntajes significativamente peores), presentan resultados más dispersos entre pruebas, no siendo ya tan evidente que el buen rendimiento en una materia asegure un rendimiento similar en las restantes.

Lo dicho anteriormente se reafirma en el cuadro A.IV.1, donde mediante sencillas regresiones MCO se vincula la posición en el ranking con los diferentes valores de $\hat{\rho}$. Tanto para el año 2006 como para el año 2009, en la

mayoría de los casos el efecto es significativo y ampliamente negativo (en línea con lo planteado).

En segundo lugar, tal como se explicó en el capítulo de estrategia empírica, los valores de $\hat{\rho}$ obtenidos responden a tres regresiones distintas que se diferencian por el nivel de información que incorporan (desde la constante sola, hasta la constante, la información del centro y la del niño). Como se puede observar en el cuadro 6.2, para ambos años y para todos los países se cumple la relación: $\hat{\rho}_1 > \hat{\rho}_2 > \hat{\rho}_3$. En todos los casos estudiados, al incluirse más información a la regresión, el porcentaje de variabilidad explicada por diferencias entre niños (constantes en las tres materias) cae. Además, es destacable también, y se cumple en la gran mayoría de los países, que la información relacionada a características de los niños contribuye a “controlar” en mayor medida la variabilidad de los datos que la información relacionada a características del centro.

Más aún, para los países que se encuentran en los primeros lugares del ranking, si bien el valor de $\hat{\rho}_2$ es menor del de $\hat{\rho}_1$, es muy cercano a este último. En estos casos, el centro cumple un rol menos importante, teniendo menor influencia sobre la variabilidad de los datos que en los países por debajo de la media OCDE (donde la diferencia entre $\hat{\rho}_2$ y $\hat{\rho}_1$ llega a ser hasta ocho veces mayor).

En esta línea, si bien es posible identificar efectos de los centros muy diferentes según el desempeño del país en las pruebas, no es posible afirmar lo mismo en relación a las características de los niños. La introducción de dicha

información a la regresión, aunque sí influye y ayuda a controlar la variabilidad tal como se dijo anteriormente, no produce efectos diferenciados entre los países aquí analizados.

En último lugar, en lo que respecta al estudio comparativo entre años es posible señalar que para el 2009, aunque todas las conclusiones anteriores se cumplen, las diferencias se atenúan tanto entre los países estudiados como entre las tres regresiones analizadas (ver gráficos anexo IV). Si bien los valores de $\hat{\rho}$ aumentan en todos los casos en comparación con 2006, el efecto del centro no es tan fuerte como en dicho año, aproximándose al efecto del niño sobre la variabilidad de los datos.

VII. CONCLUSIONES

En la presente investigación se estudió el efecto del tiempo de clase sobre el rendimiento académico de los estudiantes, siendo nuestra hipótesis orientadora que dicho efecto existe y es significativo.

En primer lugar, una vez resuelto el problema del sesgo de selección, la evidencia empírica obtenida a partir de las bases PISA 2006 y 2009 señala que el efecto del promedio de los minutos de clase no es igualmente significativo en todas las materias. Mientras que en el año 2006, la influencia de la variable *promediomin* se muestra tanto en matemática como en ciencias, pero no en lectura; para el año 2009 dicha significación se reduce sólo a ciencias.

En líneas generales, un minuto más de clase aumenta el rendimiento de los estudiantes entre 0.10 y 0.25, lo que transformado a tiempo hora hace más visible aún la importancia de esta variable. Si bien el aumento de la cantidad de horas de clase como forma de mejorar la calidad educativa ha tomado relevancia en el debate público, consideramos que es crucial al mismo tiempo asegurarse que se dicte la cantidad de horas de clases que se supone deben ser dictadas.

Es respecto a esto último, que este trabajo se propone además colaborar con el debate mediante la introducción de una variable que considera el porcentaje de clases que no fueron dictadas en los centros públicos durante el año 2009. Los resultados obtenidos para la misma son igual de significativos pero aún más contundentes que para el promedio de minutos.

En segundo lugar, además de resolver la limitación que presenta la base PISA, se avanzó en el control de las características inobservables de los alumnos y las materias, mediante estimadores intra-individuo. Al estudiar la variación intra-alumno del tiempo de clase, se buscó demostrar la existencia de un efecto causal entre este y los logros académicos de los individuos.

La no significación de la variable *promediomin* en dicha regresión no hace posible la afirmación de que existe el efecto causal, no pudiéndose probar nuestra hipótesis orientadora más allá de lo que permiten los resultados obtenidos del modelo con corrección de sesgo de selección.

En lo que respecta al desempeño de Uruguay en las pruebas, en comparación con el de otros países y sus posiciones en el ranking OCDE, la herramienta de descomposición de varianza nos aportó resultados interesantes.

Por un lado, se obtuvo que los alumnos que pertenecen a países cuyo puntaje en las pruebas supera la media de la OCDE, tienden a desempeñarse bien en todas las materias cuando lo hacen en una de ellas; obteniendo así rendimientos menos dispersos entre pruebas que aquellos estudiantes cuyos países están por debajo de dicha media.

Por otro lado, se evidencia un efecto mayor de los centros educativos sobre la varianza total de los datos en el caso de los países con peor desempeño (no pudiéndose afirmar lo mismo en el caso de las variables asociadas a los alumnos). Es así que podemos recalcar aún más la importancia del rol de los centros educativos en nuestro país.

VIII. REFERENCIAS

ANEP (2007): "Informe Nacional PISA 2006 Uruguay", ANEP, Montevideo.

ANEP (2009): "Uruguay en PISA 2009. Primeros resultados en Ciencias, Matemática y Lectura del Programa Internacional de Evaluación de Estudiantes. Informe Ejecutivo", ANEP, Montevideo.

Arellano, M. (2003): "Panel Data Econometrics", Oxford University Press

Averch, H.A., Carroll, S.J., Donaldson, T.S., Kiesling, H.J. y J. Pincus (1974): "How effective is schooling? A critical review of research", Englewood Cliffs, N.J.: Educational Technology Publications.

Becker, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis", Journal of Political Economy, 70(5), 9-49.

Becker, G.S. (1964): "Human Capital", Columbia University Press for NBER, New York.

Bourdieu, P., J.C. Passeron (1971): "La reproducción. Elementos para una teoría del sistema de enseñanza", Barcelona: Ed. Laia.

Bourguignon, F., Ferreira F.H.G. y M. Menéndez (2003): "Inequality of Outcomes and Inequality of Opportunities in Brazil", Policy research Working Paper 3174, The World bank, Washington D.C.

Brookover, W., (1979): "School Social Systems and Student Achievement: Schools Can Make a Difference", Praeger, New York.

Casassus, J., Cusato, S., Froemel, J. E., y J. C. Palafox (2001): "Primer Estudio Internacional Comparativo Sobre Lenguaje, Matemática y Factores Asociados, Para Alumnos Del Tercer y Cuarto Grado de la Educación Básica. Informe Técnico", Santiago, Chile: UNESCO.

Cerdan-Infantes, P. y C. Vermeersch (2007): "More Time Is Better: An Evaluation of the Full-Time School Program in Uruguay", Policy Research Working Paper 4167, World Bank.

Coleman, J.S., (1966): "Equality of Educational Opportunity", Arno Press.

Coleman, J.S., Hoffer, T. y S. Kilgore (1982): "High School Achievement: Public, Catholic and Private Schools Compared", Basic Books.

Da Rocha, P., Martínez, J.P. y P. Rímoli (2010): "La escuela importa: Incidencia de los factores asociados a la Escuela Pública sobre las competencias de los estudiantes de PISA 2006", Trabajo de Investigación monográfica, Facultad de Ciencias Económicas, Universidad de la República.

Denison, E.F. (1962): "The Sources of Economic Growth in the United States and the Alternatives Before Us", Committee on Economic Development, New York.

Denison, E.F. (1974): "Accounting for United States economic growth, 1929-1969", Washington D. C.: The Brookings Institution.

Edmonds, R. (1979): "Effective schools for the urban poor", *Educational Leadership*, 37(1), 15-24.

Fernández Aguerre, T. (1999): "Análisis Organizacional en Educación", Documento de Trabajo 46, Departamento de Sociología, Facultad de Ciencias Sociales, Montevideo.

Fernández Aguerre, T. (2004): "De las 'Escuelas Eficaces' a las Reformas Educativas", en *Estudios Sociológicos de El Colegio de México*, Vol 22 (65).

Fernández Aguerre, T. y E. Blanco (2004): "¿Cuánto importa la escuela? El caso de México en el contexto de América Latina", en *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 2 (1), 1-27.

Gamoran, A. y D. A. Long (2006): "Equality of Educational Opportunity: A 40-year retrospective", WCER Working Paper 2006-9, Wisconsin Center for Education Research.

Giroux (1985): "Teorías de la reproducción y la resistencia en la nueva sociología de la educación: un análisis crítico", en *Cuadernos Políticos*, México, Era, 36-65.

Hanushek, E. A. y L. Woessmann (2007): "The role of education quality for economic growth," Policy Research Working Paper 4122, The World Bank.

Lavy, V. (2010): "Do Differences in school's instruction time explain international achievement gaps in math, science and reading? Evidence from developed and developing countries", NBER Working Papers 16227, National Bureau of Economic Research.

Llambí, C., M. Perera (2008): "La Función de Producción Educativa: el posible sesgo en la estimación de efectos "institucionales" con los datos PISA. El caso de las escuelas de Tiempo Completo", Trabajo realizado para una investigación

financiada por el Fondo Concursable “Carlos Filgueira” del Programa Infancia, Adolescencia y Familia del Ministerio de desarrollo Social, edición 2008, CINVE.

Llambí, C., Perera, M. y P. Messina (2009): “Desigualdad de oportunidades y el rol del sistema educativo en los logros de los jóvenes uruguayos”, Investigación financiada por el Fondo “Carlos Filgueira”, del Programa Infancia, Adolescencia y Familia – Ministerio de Desarrollo Social.

Martínez Arias, R. (2006): “La metodología de los estudios PISA”, en Revista de Educación, 111-129, Universidad Complutense.

Méndez, N. y M. Zerpa (2009): “Desigualdad en las capacidades educativas en Uruguay y Chile”, Trabajo de Investigación monográfica, Facultad de Ciencias Económicas, Universidad de la República.

Mincer, J. (1958): “Investment in Human Capital and Personal Income Distribution”, *Journal of Political Economy*, 66, 281-302.

Nunnally, J. y I. Bernstein (1995): “Teoría psicométrica”, México: McGraw-Hill.

OCDE (2006): “PISA 2003 Manual de análisis de datos. Usuarios de SPSS”, OCDE e INECSE, Madrid.

OCDE (2009): “PISA 2006 Technical report”, PISA/OECD.

OCDE (2010): OECD Education at Glance, 2010:

<http://dx.doi.org/10.1787/888932310472>

Oroval E., y J. Escardibul Ferra (1998): “Economía De La Educación”, Madrid, Encuentro Ediciones.

Roemer, J.E. (1998): “Equality of Opportunity”, Cambridge, Harvard University Press.

Ruiz de Miguel, C. y M. Castro Morera (2006): “Un Estudio Multinivel Basado en PISA 2003: Factores de Eficacia Escolar en el área de Matemáticas”, Archivos Analíticos de Políticas Educativas, 14 (29), Universidad Complutense.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., y A. Smith (1979): “Fifteen thousand hours-secondary schools and their effects on children”, Cambridge, MA: Harvard University Press.

Sanroman, G. (2010): "Intergenerational Educational Mobility: evidence from three approaches for Brazil, Chile, Uruguay and the USA (1995-2006)", Working Paper 0110, Departamento de Economía, Facultad de Ciencias Sociales, Montevideo.

Schultz, T.W. (1961): "Education and Economic Growth" en HENRY, N. (ed.) *Social Forces Influencing American Education*, Chicago: National Society for the Study of Education, 46-88.

Solow, R. (1957): "Technical Change and the Aggregate Production Function", *The Review of Economics and Statistics*, 39 (3), 312-320, Cambridge, Massachussets Institute of Technology, Press.

Vegas, E. y J. Petrow (2007): "Raising student achievement in Latin America: the challenge for the 21st century", Latin American Development Forum, Banco Mundial, Washington, D.C.

Wooldridge, J.M. (2002): "Econometric Analysis of Cross Section and Panel Data", Cambridge, Massachussets Institute of Technology, Press.

IX. ANEXOS

Anexo I - Presentación de las variables y estadísticas descriptivas

VARIABLE	INDICADOR
VARIABLES DEL ALUMNO Y DEL CONTEXTO SOCIOCULTURAL	
sexo	Mujer (0= Varón; 1=Mujer)
iestatuseco	Índice de estatus socio-económico y cultural construido por PISA
escolaridadm	Madre no pasó primaria (0= Pasó primaria; 1= No pasó primaria)
areares	Reside en el interior (0=Montevideo; 1=Interior)
tamañoloc	Tamaño de la localidad (0= Ciudades grandes o medianas; 1= Pequeñas Localidades)
grado	Año que cursa al momento de tomar la prueba
institucionp	Tipo de institución en el que cursó primaria (dummy por tipo de centro): ETC pública Escuela pública ETC privada Escuela privada Escuela Rural (dummy omitida)
institucioncb	Tipo de institución en el que cursó ciclo básico (dummy por tipo de centro): P86 público P96 público Ciclo básico privado Ciclo básico rural Ciclo básico técnico No cursó ciclo básico (dummy omitida)
repeticionp	Repitió primaria al menos una vez (0 = No repitió; 1= Repitió al menos una vez)
FACTORES ESCOLARES	
promediomin	Promedio de minutos de clase por materia por escuela
publicoprivado	Tipo de centro (0= público; 1= privado)
tamañocentro	Cantidad estudiantes matriculados en el centro educativo
escasezdoc	Índice de escasez de personal docente
propdoctitulados	Proporción de docentes titulados
propdocuniv	Proporción de docentes con nivel ISCED 5A (título universitario)
inasistencias	Porcentaje de cantidad de horas no dictadas sobre las que se debieron dictar en centros educativos públicos.

Cuadro A.I.1 – Media y desvío estándar de los resultados obtenidos según características relevantes de los individuos bajo estudio.

	PISA 2006			PISA 2009		
	Mat.	Cien.	Lect.	Mat.	Cien.	Lect.
Resultados	435.62 (99.12)	437.54 (94.46)	423.73 (119.80)	423.81 (92.87)	424.63 (96.70)	423.13 (100.57)
Mujeres	428.36 (96.49)	436.43 (90.70)	441.56 (114.51)	418.22 (89.80)	425.69 (92.88)	442.43 (93.84)
Varones	443.81 (101.42)	438.79 (98.52)	403.57 (122.45)	430.06 (95.82)	423.44 (100.81)	401.52 (103.42)
Madre no terminó primaria	394.65 (92.71)	397.93 (85.13)	379.06 (113.80)	361.26 (80.32)	364.38 (84.88)	354.93 (87.47)
Madre terminó primaria	457.99 (95.43)	459.43 (92.10)	448.43 (115.94)	429.47 (91.52)	430.30 (95.36)	429.96 (98.73)
Reside en grandes localidades	453.39 (121.68)	450.55 (97.67)	441.00 (121.68)	446.81 (102.47)	447.06 (100.72)	448.25 (102.47)
Reside en pequeñas localidades	422.32 (116.85)	427.62 (90.79)	410.45 (116.85)	409.71 (96.23)	410.92 (91.50)	407.79 (96.22)
Repitió primaria	340.17 (84.12)	349.05 (77.00)	320.49 (105.36)	339.88 (68.77)	343.74 (77.30)	332.67 (78.44)
No repitió primaria	462.52 (88.38)	463.54 (85.15)	454.16 (108.52)	451.45 (83.20)	451.83 (87.09)	452.57 (89.41)

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Los cálculos fueron elaborados utilizando el valor plausible 3

Fuente: Elaboración propia en base a la información de PISA 2006 y 2009

Cuadro A.I.2 – Media y desvío estándar de los resultados obtenidos en las pruebas por tipo de centro.

	PISA 2006			PISA 2009		
	Mat.	Cien.	Lect.	Mat.	Cien.	Lect.
Público	418.76 (96.35)	421.31 (90.97)	404.01 (117.16)	407.98 (86.93)	408.42 (91.07)	405.69 (94.4)
Privado	497.15 (84.38)	496.26 (83.52)	494.37 (102.58)	498.56 (83.10)	501.16 (85.15)	505.52 (87.02)

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Los cálculos fueron elaborados utilizando el valor plausible 3

Fuente: Elaboración propia en base a la información de PISA 2006 y 2009

Cuadro A.I.3 - Media y desvío estándar de la cantidad de minutos de clase a la semana por materia.

Años	Matemáticas	Ciencia	Lectura
2006	206.02 (104.06)	160.36 (116.16)	171.66 (93.59)
2009	191.83 (55.33)	330.23 (149.61)	184.21 (54.88)

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Fuente: Elaboración propia en base a la información de PISA 2006 y 2009

Cuadro A.I.4 - Estadísticas descriptivas para el año 2009 de la variable inasistencias docentes.

Media	Desvío	Max.	Min.
14.31	5.64	1.28	37.40

Percentiles

10%	25%	50%	75%	90%
6.3	10.7	14.31	15.4	19

Nota: La variable inasistencias fue construida a partir de los datos brindados por el Consejo de Educación Secundaria y refleja el porcentaje de la cantidad de horas no dictadas en cada centro educativo público, sobre las que se debieron dictar en el año 2009.

Cuadro A.I.5 - Inasistencias docentes promedio según el tamaño de la localidad en la que reside el estudiante.

Tamaño de la localidad	Inasistencias
Grandes o medianas ciudades	15.28 (5.73)
Pequeñas localidades	13.91 (5.64)

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Fuente: Elaboración propia en base a PISA 2009 y a la información brindada por el Consejo de Educación Secundaria.

Anexo II - Consideraciones metodológicas

II.1 - Valores Plausibles: metodología para su utilización

Al estudiar el rendimiento de un estudiante, PISA pretende inferir la capacidad subyacente inobservable (competencias cognitivas) a partir de una serie de ítems que se encuentran en el cuestionario. Para realizarlo, se enfrenta al problema de que cada sujeto responde solamente a un número limitado de ítems del test y es necesario estimar como sería el comportamiento en el total de los ítems que se utilizan para la evaluación (Martínez Arias, 2006). Para solucionar este problema, PISA realiza el cálculo de los resultados de cada prueba a través de valores plausibles tomando en consideración la teoría de respuesta al ítem (TRI) basada en el modelo de Rasch. A diferencia de la teoría clásica que estima el nivel de atributo como la sumatoria de respuestas de ítems individuales, la TRI utiliza un patrón de respuesta y logra obtener un conjunto de valores plausibles que determinarán dicha capacidad (Nunnally y Bernstein, 1995).

El modelo de Rasch consiste en crear una distribución a posteriori a partir de los puntajes individuales de cada prueba para cada sujeto con sus probabilidades asociadas. A través de una función de probabilidad, se logra crear una distribución continua donde quedan alocadas las dificultades de cada ítem y las habilidades de los estudiantes. La dificultad relativa de cada ítem resulta de la comparación entre ítems y es independiente de las habilidades del estudiante. Por lo tanto, se crea una escala donde se ubican las dificultades relativas para que posteriormente pueda computarse la puntuación de los estudiantes (OCDE, 2009).

A partir de la definición anterior se deriva que la metodología para el trabajo con valores plausibles consiste en dos pasos. En primer lugar, debe obtenerse matemáticamente las distribuciones posteriores alrededor de los valores que se obtienen como resultado de cada prueba. En segundo lugar, se le debe asignar a cada individuo un conjunto de valores aleatorios que se extraen a partir de éstas distribuciones. Una estimación individual de las habilidades de un

estudiante puede derivarse de distribuciones posteriores, ésta estimación derivada para cada individuo es llamada “*expected a posteriori estimator*” (EAP). Se considera a ésta última como la media de un conjunto infinito de valores plausibles para un estudiante en particular (OCDE, 2009). En lugar de estimar directamente las capacidades de un alumno se estima una distribución de probabilidad, y se seleccionan cinco valores aleatorios de dicha función, con una probabilidad asociada a cada valor (Ruiz de Miguel y Castro Morera, 2006).

Para poder trabajar con valores plausibles, los modelos de estudio deben considerar cada valor plausible por separado y luego agregar los resultados para determinar: un estimador final, determinado como un promedio de las estimaciones realizadas con los cinco valores plausibles, y una varianza total de la estimación que combina la varianza de imputación, medida del margen de error que se introduce por la estimación de habilidades a través de pruebas con un número limitado de preguntas, y la varianza muestral.

Un factor adicional a tomar en cuenta para el cálculo de estadísticos y de intervalos de confianza es que, en lugar de realizar un muestreo aleatorio simple, PISA realiza el muestreo en dos etapas: en primer lugar se seleccionan los centros a partir de un procedimiento sistemático y en segundo lugar se seleccionan los estudiantes dentro de cada centro de forma aleatoria (Méndez y Zerpa, 2009). Los centros con mayor tamaño tienen mayor probabilidad de ser seleccionados en la muestra pero los estudiantes que asisten a estos centros tienen menor probabilidad dentro del centro de ser seleccionados que los que asisten a centros de menor tamaño (OCDE, 2009). Debido a que existe el reconocimiento de que algunas unidades son más importantes que otras, se procede a introducir ponderaciones asociadas a cada centro y cada estudiante en la muestra. Los pesos que se asignan son inversamente proporcionales a la probabilidad de selección (OCDE, 2009).

En el caso de aplicación de un muestreo en dos etapas, para evitar la subestimación de la varianza muestral en las estimaciones de las varianzas se debe proceder a introducir métodos de replicación. Estos últimos consisten en estimar la varianza del parámetro de interés mediante el uso de un gran

número de submuestras replicadas a partir de la muestra completa. La base de datos de PISA contiene 80 pesos replicados para cada estudiante generados a partir de la versión desarrollada por Fay del método BRR (*Balanced Repeated Replication*). Fay sugiere multiplicar los pesos por un deflactor k que se encuentre entre 0 y 1, con el segundo factor definido como 2 menos k . PISA establece el deflactor k en 0.5 y determina el estadístico de interés a partir de toda la muestra y de cada replica (OCDE, 2009). La varianza muestral se determina entonces por la comparación entre lo estimado para el total de la muestra con lo estimado para cada replica, es decir:

$$\sigma_{(\hat{\theta})}^2 = \frac{1}{G(1-k)^2} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\theta})^2 = \frac{1}{80(1-0.5)^2} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\theta})^2 = \frac{1}{20} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\theta})^2$$

Siendo:

G Cantidad de replicas que decide generar PISA

k Deflactor

$\hat{\theta}_{(i)}$ Coeficiente estimado del parámetro de interés para la réplica i

$\hat{\theta}$ Coeficiente estimado del parámetro de interés

La metodología para los valores plausibles combinada con el método aplicado para las replicas, requiere que el parámetro a estimar sea computado 80 veces para cada uno de los cinco valores plausibles, es decir 405 en total, para obtener el estimador final y su error estándar. Tal como se explica en la segunda edición del Manual para análisis de datos PISA con SPSS (OCDE, 2009), el cálculo de un estadístico, así como el cálculo de los coeficientes de una regresión, puede resumirse en seis pasos:

- 1) El estadístico requerido y su correspondiente error estándar debe ser calculado para cada valor plausible como fue explicado. Cualquier análisis que requiera el cálculo con los cinco valores plausibles requerirá 405 estimaciones.
- 2) La media se calcula como el promedio de la media calculada para los cinco valores plausibles:

$$\hat{\mu} = \frac{1}{5}(\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4 + \hat{\mu}_5)$$

En el caso de la estimación de los coeficientes de las variables explicativas de una regresión, el coeficiente de la regresión final se determina como el promedio de los coeficientes calculados para cada uno de los cinco valores plausibles.

$$\hat{\beta} = \frac{\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \hat{\beta}_4 + \hat{\beta}_5}{5}$$

- 3) La varianza muestral se calcula como el promedio de las varianzas calculadas utilizando los 80 pesos replicados para los cinco valores plausibles:

$$\sigma_{(\mu)}^2 = \frac{1}{5}(\sigma_{(\mu_1)}^2 + \sigma_{(\mu_2)}^2 + \sigma_{(\mu_3)}^2 + \sigma_{(\mu_4)}^2 + \sigma_{(\mu_5)}^2)$$

Para el caso de estimación de los coeficientes de una regresión:

$$\sigma_{(\hat{\beta})}^2 = \frac{1}{5}(\sigma_{(\hat{\beta}_1)}^2 + \sigma_{(\hat{\beta}_2)}^2 + \sigma_{(\hat{\beta}_3)}^2 + \sigma_{(\hat{\beta}_4)}^2 + \sigma_{(\hat{\beta}_5)}^2)$$

- 4) La varianza de imputación se calcula como:

$$\sigma_{(test)}^2 = \frac{1}{M-1} \sum_{i=1}^M (\mu_i - \mu)^2 = \frac{1}{4} \sum_{i=1}^5 (\mu_i - \mu)^2$$

Siendo:

M Número de valores plausibles

Para el caso de estimación de los coeficientes de una regresión:

$$\sigma_{(test)}^2 = \frac{1}{4} \sum_{i=1}^5 (\hat{\beta}_i - \hat{\beta})^2$$

- 5) La varianza de imputación y la varianza muestral se combinan para obtener la varianza final de la siguiente forma:

$$\sigma_{(error)}^2 = \sigma_{(\mu)}^2 + (1.2 \sigma_{(test)}^2)$$

$$\text{con } 1.2 = \left(1 + \frac{1}{M}\right)$$

Para el caso de estimación de los coeficientes de una regresión:

$$\sigma_{(error)}^2 = \sigma_{(\beta)}^2 + (1.2\sigma_{(test)}^2)$$

- 6) El error estándar es igual a la raíz cuadrada de la varianza final calculada en el punto anterior:

$$SE = \sqrt{\sigma_{(error)}^2}$$

II.2 - Mínimos Cuadrados Ordinarios y Sesgo de Selección

El modelo del que partimos puede ser expresado como sigue:

$$Y_{ijk} = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + (\varepsilon_j + \eta_k) + u_{ijk} \quad (1)$$

Donde Y_{ijk} es el resultado en la prueba del estudiante i en la escuela j en la materia k , H_{kj} es el tiempo de clase de la materia k en el centro j , X_{ij} es un vector de características del estudiante i en el centro j , S_j es un vector de características del centro j , ε_j y η_k representan características inobservables del centro y la materia respectivamente, u_{ijk} es el término de error restante y μ_i es el efecto fijo del estudiante que captura el contexto familiar, habilidades, motivación, y otras características no cognitivas del niño.

Al quedar las características inobservables del centro y de las materias agregadas en el término de error, podemos reescribir el modelo como:

$$Y_{ijk} = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + v_{ijk} \quad (2)$$

Siendo $v_{ijk} = (\varepsilon_j + \eta_k) + u_{ijk}$

Sesgo de Selección

La regresión planteada para la población objetivo está dada por:

$$Y_{ijk}^* = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + v_{ijk} \quad (3)$$

Donde $E[v_{ijk} | H_{kj}, X_{ij}, S_j] = 0$

El problema radica en que la muestra no es representativa de la población objetivo, ya que existe una regla de selección que determina la observabilidad de Y_{ijk}^* , que se puede escribir como:

$$T_i = 1[Z_i \omega + h_i > 0] \quad (4)$$

Donde T_i es una variable que vale 1 cuando el individuo asiste al sistema educativo post-primaria y 0 en otro caso, el vector Z contiene variables que determinan la asistencia, y h_i es un término de error.

La probabilidad del individuo i de estar en el sistema educativo post-primaria es:

$$P(T_i = 1) = \Phi Z_i \omega \quad (5)$$

Esto implica admitir ciertos supuestos (Méndez y Zerpa, 2009):

1) T_i y Z_i son observados siempre, mientras que Y_{ijk}^* es visto únicamente para aquellos individuos en los que se cumple $T_i = 1$ (no son desertores).

2) el vector (v, h) es independientes de Z_i y tiene media igual a 0.

3) $h_i \sim \text{Normal}(0,1)$

4) $v_i = \rho h_i + \varphi_i$, con $\varphi_i \sim \text{Normal}(0,1)$ independiente de h_i

Al estimarse los parámetros de la ecuación original a partir de los datos muestrales, existirá un sesgo que se puede expresar como:

$$\begin{aligned}
E[v_i | T_i = 1, Z_i] &= E[v_i | \langle h_i > -Z_i\omega \rangle, Z_i] = \\
&= \underbrace{\rho E[h_i | \langle h_i > -Z_i\omega \rangle, Z_i]}_{\rho\lambda(Z_i\omega)} + \underbrace{E[\varphi | \langle h_i > -Z_i\omega \rangle, Z_i]}_0 \quad (6)
\end{aligned}$$

Resumiéndose el sesgo de selección como:

$$E[v_i | T_i = 1, Z_i] = \rho\lambda(Z_i\omega) \quad (7)$$

Donde $\lambda(Z_i\omega) = \frac{\phi(Z_i\omega)}{\Phi(Z_i\omega)}$ es el inverso del Ratio de Mills

En síntesis, existe sesgo de selección distinto de 0 en la estimación de los parámetros de la ecuación original cuando $\rho \neq 0$, lo cual sucede cuando v_i y h_i están correlacionados.

II.3 - Estimación con corrección del sesgo de selección

La estimación de la ecuación original mediante Mínimos Cuadrados Ordinarios trae aparejado la aparición de un sesgo y conduce a estimadores inconsistentes si se omite el regresor $\lambda(Z_i\omega)$.

Para corregir esto es que se aplica el método en dos etapas de Heckman (1976).

El modelo es planteado a partir de las ecuaciones:

$$Y_{ijk}^* = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + v_{ijk} \quad (3)$$

(Ecuación Estructural)

$$T_i = 1[Z_i\omega + h_i > 0] \quad (4)$$

(Ecuación de selección)

Siendo la ecuación de regresión válida para las observaciones muestrales (para las cuales $T_i = 1$):

$$Y_{ijk} = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + \rho \lambda(Z_i \hat{\omega}) + w_{ijk} \quad (8)$$

Donde $w_{ijk} = v_{ijk} + \rho(\lambda_i - \hat{\lambda}_i)$

El método bietapico de Heckman consiste en (Wooldridge, 2002):

- 1- Estimar un modelo Probit de T_i sobre Z_i y obtener los estimadores ω consistentes para una muestra representativa de toda la población. Luego debe estimarse el Inverso del Ratio de Mills, como $\lambda(Z_i \hat{\omega}) = \frac{\phi(Z_i \hat{\omega})}{\Phi(Z_i \hat{\omega})}$ para cada individuo presente en la muestra en la cual se observa la variable Y_{ijk} .
- 2- Se estiman los parámetros de la ecuación (8) a través de MCO.

Como resultado, los estimadores obtenidos son consistentes y asintóticamente normales.

II.4 - Estimador Intra-Individuo

Utilizando estimadores intra-individuo, es posible controlar por todas las características de los individuos (incluso aquellas omitidas y las que no se pueden medir) siempre que las mismas no cambien en el tiempo, no incurriendo en el sesgo antes mencionado. Esto se debe a que el método se centra en la variación intra-individuo de las variables, ignorando la variación entre individuos.

La implementación de este método requiere de dos condiciones:

- 1- Cada individuo en la muestra debe tener dos o más mediciones en la misma variable dependiente.

2- Para al menos algunos de los individuos, los valores de las variables independientes de interés deben presentar valores distintos en al menos dos de las ocasiones de medición.

Partimos de la siguiente especificación:

$$Y_{ijk} = \mu_i + \gamma H_{kj} + \beta X_{ij} + \delta S_j + (\varepsilon_j + \eta_k) + \rho \lambda(Z_i \hat{\omega}) + u_{ijk}$$

Donde el subíndice i refiere a diferentes individuos y k a diferentes mediciones dentro de un mismo individuo, y nuevamente Y_{ijk} es el resultado en la prueba del estudiante i en el centro j en la materia k, H_{kj} es el tiempo de clase de la materia k en el centro j, X_{ij} es un vector de características del estudiante i en el centro j, S_j es un vector de características del centro j, ε_j y η_k representan características inobservables del centro y la materia respectivamente, $\rho \lambda(Z_i \hat{\omega})$ representa el sesgo de selección, u_{ijk} es el término de error restante y μ_i es el efecto fijo del estudiante que captura el contexto familiar, habilidades, motivación, y otras características no cognitivas del niño.

Para deshacerse de los efectos fijos individuales, se resta a la ecuación original las medias de las variables que presentan variación intra-individuo:

$$Y_{ijk} - \bar{Y}_{ij} = \gamma(H_{kj} - \bar{H}_j) + (\eta_k - \bar{\eta}) + (u_{ijk} - \bar{u}_{ij})$$

$$\text{Donde } \bar{Y}_{ij} = \frac{1}{K} \sum_{k=1}^K Y_{ijk} \quad \bar{H}_j = \frac{1}{K} \sum_{k=1}^K H_{kj} \quad \bar{\eta} = \frac{1}{K} \sum_{k=1}^K \eta_k \quad \text{y} \quad \bar{u}_{ij} = \frac{1}{K} \sum_{k=1}^K u_{ijk}$$

Lo anterior puede expresarse como:

$$\tilde{Y}_{ijk} = \gamma \tilde{H}_{kj} + (\eta_k - \bar{\eta}) + \tilde{u}_{ijk}$$

$$\text{Siendo } \tilde{Y}_{ijk} = Y_{ijk} - \bar{Y}_{ij}, \quad \tilde{H}_{kj} = H_{kj} - \bar{H}_j \quad \text{y} \quad \tilde{u}_{ijk} = u_{ijk} - \bar{u}_{ij}$$

Para controlar las características inobservables η_k se introducen K-1 variables dummy, quedando el modelo expresado como sigue:

$$\tilde{Y}_{ijk} = \gamma \tilde{\Pi}_{kj} + \delta_1 d1_k + \delta_2 d2_k + \dots + \delta_{K-1} d(K-1)_k + \tilde{u}_{ijk}$$

Una vez eliminados los efectos fijos del individuo y controladas las características inobservables η_k , el coeficiente de interés γ es estimado mediante MCO.

Para que $\hat{\gamma}_{\text{intraindividuo}}$ tenga características deseables deben realizarse los siguientes supuestos (Wooldridge, 2002):

1- Exogeneidad estricta: $E(\tilde{u}_{ijk} | \tilde{\Pi}_{kj}, \delta_1 d1_k, \delta_2 d_k, \dots, \delta_{K-1} d(K-1)_k) = 0$

2- Homoscedasticidad:

$$Var(\tilde{u}_{ijk} | \tilde{\Pi}_{kj}, \delta_1 d1_k, \delta_2 d_k, \dots, \delta_{K-1} d(K-1)_k) = Var(\tilde{u}_{ijk}) = \sigma_u^2 \quad \forall k = 1, \dots, K$$

3- $Cov(\tilde{u}_{ijk}, \tilde{u}_{ijs} | \tilde{\Pi}_{kj}, \delta_1 d1_k, \delta_2 d_k, \dots, \delta_{K-1} d(K-1)_k) = 0$

4- Las variables explicativas deben variar por k y no debe existir relación lineal perfecta entre dichas variables.

Bajo los supuestos enumerados puede afirmarse que el estimador $\hat{\gamma}_{\text{intraindividuo}}$ obtenido es insesgado y consistente.

Anexo III - Resultados de la estimación del efecto del tiempo de clase

Cuadro A.III.1 - Resultados para el año 2006 de la estimación del modelo por MCO

	Matemáticas	Ciencia	Lectura
Promediomin	0.25 (0.08)	0.13 (0.05)	0.14 (0.11)
Mujer	-25.49 (3.33)	-11.94 (3.55)	29.55 (4.75)
Iestatusco	16.36 (1.94)	13.60 (1.44)	13.87 (2.04)
Tamañoloc	-8.98 (5.00)	-1.95 (4.63)	-8.84 (6.98)
Privado	-2.35 (10.39)	4.75 (10.66)	36.16 (13.90)
Tamañocentro	0.01 (0.01)	0.02 (0.01)	0.03 (0.01)
Escasezdoc	2.05 (2.22)	1.65 (2.04)	1.41 (2.59)
propdoctitulados	12.66 (13.06)	20.63 (10.51)	19.16 (17.19)
Propdocuniv	42.34 (33.96)	25.46 (37.44)	-40.55 (50.75)
Grado	37.88 (2.42)	38.40 (2.44)	46.02 (3.28)
ETCpública	-22.52 (12.36)	-21.01 (6.77)	-16.39 (11.51)
Escuelapublica	3.79 (10.93)	3.55 (5.97)	5.13 (8.61)
ETCprivada	16.45 (12.52)	17.00 (9.53)	17.80 (11.20)
Escuelaprivada	6.91 (12.13)	0.58 (7.07)	6.38 (9.34)
P86publico	-51.57 (29.98)	-13.52 (22.69)	-71.48 (33.55)
P96publico	-55.46 (30.86)	-17.64 (22.75)	-72.66 (33.57)
ciclobasicotecnico	-70.36 (29.37)	-43.39 (21.93)	-86.07 (33.09)
ciclobasicoprivado	-45.61 (31.68)	-4.20 (24.73)	-68.49 (36.96)
ciclobasicorural	-75.49 (36.50)	-14.01 (25.02)	-53.79 (41.95)
Cons	81.40 (37.61)	52.85 (32.68)	-6.99 (47.81)

Número de Observaciones: 4326

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Cuadro A.III.2 - Resultados para el año 2009 de la estimación del modelo por MCO

	Matemáticas	Ciencia	Lectura
Promediomin	-0.05 (0.11)	0.11 (0.03)	0.02 (0.07)
Mujer	-24.95 (2.93)	-12.48 (2.98)	30.49 (3.17)
lestatusedo	12.77 (1.47)	13.63 (1.39)	14.21 (1.40)
Tamañooc	-11.49 (5.57)	-7.12 (4.67)	-8.18 (4.57)
Privado	40.82 (11.08)	36.81 (8.27)	49.60 (7.44)
Tamañoocentro	0.001 (0.003)	0.01 (0.002)	0.01 (0.003)
Escasezdoc	0.02 (2.59)	3.35 (2.37)	-0.01 (2.36)
propdoctitulados	15.00 (13.07)	24.01 (12.26)	8.58 (11.74)
Propdocuniv	25.58 (39.94)	28.71 (29.80)	36.69 (34.46)
Grado	43.90 (2.44)	44.84 (2.51)	46.42 (2.25)
Cons	33.26 (34.70)	-43.46 (22.63)	-42.29 (32.35)

Número de Observaciones: 4179

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Cuadro A.III.3 - Resultados para el año 2006 de la estimación del modelo bietapico de Heckman

	Matemáticas	Ciencia	Lectura
Promediomin	0.25 (0.08)	0.13 (0.05)	0.12 (0.11)
Mujer	-44.53 (4.83)	-30.08 (4.98)	6.38 (6.20)
Iestatusesco	8.07 (2.31)	6.48 (2.06)	3.97 (3.21)
Tamañoloc	-6.05 (5.09)	-0.52 (4.72)	-5.40 (7.07)
Privado	-1.40 (11.40)	9.44 (11.69)	40.43 (15.37)
Tamañocentro	0.01 (0.01)	0.02 (0.01)	0.03 (0.01)
Escasezdoc	1.68 (2.26)	1.87 (2.03)	1.81 (2.71)
propdoctitulados	16.05 (12.66)	21.94 (10.54)	19.86 (16.57)
Propdocuniv	38.48 (37.25)	19.65 (41.29)	-60.76 (53.06)
Grado	31.07 (3.04)	31.41 (2.93)	37.02 (3.83)
ETCpública	-23.12 (13.23)	-20.74 (6.99)	-11.93 (11.75)
Escuelapublica	0.98 (11.98)	1.52 (6.42)	4.03 (9.03)
ETCprivada	13.39 (13.66)	13.93 (10.24)	15.20 (12.10)
Escuelaprivada	1.89 (13.05)	-4.03 (7.78)	1.87 (10.48)
P86publico	-61.64 (29.77)	-15.73 (22.76)	-71.57 (34.98)
P96publico	-63.31 (30.67)	-17.18 (22.90)	-69.69 (35.28)
ciclobasicotecnico	-78.70 (28.86)	-47.28 (22.00)	-86.13 (34.64)
ciclobasicoprivado	-57.27 (31.21)	-10.94 (25.39)	-70.45 (38.13)
ciclobasicorural	-89.12 (36.41)	-22.41 (26.36)	-58.80 (44.56)
lambda	-110.39 (16.39)	-100.08 (17.61)	-130.44 (22.11)
Cons	248.79 (50.25)	208.70 (47.31)	192.13 (62.51)

Número de Estudiantes: 4326

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Cuadro A.III.4 - Resultados para el año 2009 de la estimación del modelo bietapico de Heckman.

	Matemáticas	Ciencia	Lectura
Promediomin	-0.02 (0.11)	0.10 (0.03)	0.03 (0.07)
Mujer	-42.52 (4.26)	-22.41 (4.81)	14.58 (4.49)
lestatuseco	4.93 (1.82)	9.98 (1.83)	7.59 (1.73)
Tamañoloc	-11.47 (5.74)	-7.26 (5.06)	-8.40 (4.36)
Privado	39.46 (10.69)	34.97 (8.46)	47.21 (7.13)
Tamañocentro	0.001 (0.003)	0.01 (0.002)	0.01 (0.003)
Escasezdoc	-0.83 (2.58)	2.78 (2.47)	-1.34 (2.30)
propdoctitulados	9.24 (12.27)	20.65 (12.55)	2.16 (11.14)
Propdocuniv	18.85 (38.50)	27.76 (30.89)	35.44 (32.17)
Grado	36.58 (2.88)	40.14 (3.14)	39.88 (2.54)
Lambda	-98.91 (15.34)	-56.43 (16.08)	-88.31 (15.10)
Cons	188.81 (46.07)	60.15 (37.87)	101.08 (39.49)

Número de Estudiantes: 4179

Nota: Las cifras entre paréntesis corresponden al desvío estándar.

Cuadro A.III.5 - Resultados para el año 2009 de la estimación por MCO del modelo después de la incorporación de la variable Inasistencias.

	Matemáticas	Ciencia	Lectura
Promediomin	0.14 (0.16)	0.04 (0.04)	-0.11 (0.09)
Inasistencias	-0.19 (0.51)	-0.49 (0.29)	0.12 (0.41)
Mujer	-29.51 (3.48)	-16.67 (3.74)	25.58 (3.45)
Iestatusesco	11.30 (1.68)	12.96 (1.71)	13.19 (1.58)
Tamañoloc	-9.74 (5.17)	-5.28 (4.70)	-6.78 (4.43)
Tamañocentro	0.001 (0.003)	0.01 (0.003)	0.01 (0.004)
Escasezdoc	-0.48 (2.65)	2.58 (2.45)	-0.24 (2.32)
Propdoctitulados	8.41 (14.70)	0.93 (15.92)	-13.13 (15.01)
Propdocuniv	20.88 (43.71)	3.79 (33.12)	31.26 (45.53)
Grado	49.38 (2.46)	45.67 (2.37)	46.61 (2.38)
Cons	-46.19 (43.09)	-1.08 (26.48)	-0.57 (41.04)

Número de Estudiantes: 2775

Notas: Las cifras entre paréntesis corresponden al desvío estándar.

El modelo fue estimado únicamente para los centros educativos públicos en el año 2009. La variable inasistencias fue construida a partir de los datos brindados por el Consejo de Educación Secundaria y refleja el porcentaje de la cantidad de horas no dictadas en cada centro educativo público, sobre las que se debieron dictar.

Cuadro A.III.6 - Resultados para el año 2009 de la estimación del modelo bietapico de Heckman después de la incorporación de la variable Inasistencias.

	Matemáticas	Ciencia	Lectura
Promediamin	0.14 (0.13)	0.04 (0.04)	-0.11 (0.09)
Inasistencias	-0.25 (0.15)	-0.57 (0.29)	0.13 (0.38)
Mujer	-40.32 (5.19)	-21.87 (5.76)	17.71 (5.19)
Iestatusco	6.07 (2.22)	11.13 (2.16)	9.80 (2.06)
Tamañoloc	-9.02 (5.15)	-4.79 (4.74)	-6.74 (4.48)
Tamañocentro	0.001 (0.003)	0.01 (0.003)	0.01 (0.004)
Escasezdoc	-1.15 (2.61)	1.97 (2.53)	-1.34 (2.36)
Propdoctitulados	3.44 (14.48)	-3.95 (15.93)	-18.35 (15.07)
Propdocuniv	13.85 (39.98)	0.87 (33.16)	31.52 (41.51)
Grado	44.19 (2.59)	43.19 (2.55)	42.66 (2.57)
Lambda	-68.51 (17.27)	-33.39 (18.96)	-48.36 (17.73)
Cons	66.80 (46.32)	59.61 (39.93)	84.16 (48.02)

Número de Estudiantes: 2775

Notas: Las cifras entre paréntesis corresponden al desvío estándar.

El modelo fue estimado únicamente para los centros educativos públicos en el año 2009. La variable inasistencias fue construida a partir de los datos brindados por el Consejo de Educación Secundaria y refleja el porcentaje de la cantidad de horas no dictadas en cada centro educativo público, sobre las que se debieron dictar.

Cuadro A.III.7 - Resultados para el año 2006 de la estimación intra-alumno del efecto del tiempo en clase sobre el resultado de las pruebas PISA.

	Mat.+Cien.+Lect.	Mat.+Cien.	Mat+Lect.	Cien.+Lect.
Promediomin	0.01 (0.04)	0.05 (0.05)	0.01 (0.07)	-0.02 (0.06)
Dmat	-1.88 (2.36)	-3.64 (2.38)	13.92 (3.17)	-
Dcien	-15.80 (2.47)	-	-	-15.32 (2.61)
Cons	426.36 (7.31)	420.65 (8.27)	410.53 (13.12)	430.90 (9.65)

Número de Estudiantes: 4326

Notas: Las cifras entre paréntesis corresponden al desvío estándar.

Cuadro A.III.8 - Resultados para el año 2009 de la estimación intra-alumno del efecto del tiempo en clase sobre el resultado de las pruebas PISA.

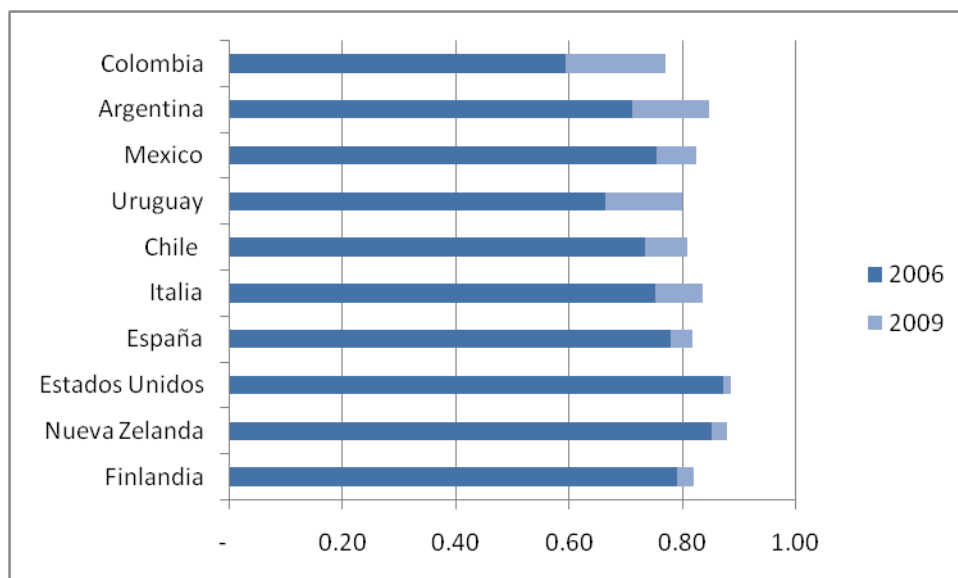
	Mat.+Cien.+Lect.	Mat.+Cien.	Mat+Lect.	Cien.+Lect.
Promediomin	-0.01 (0.02)	0.02 (0.02)	0.00 (0.06)	-0.03 (0.02)
Dmat	-1.52 (3.09)	1.61 (3.65)	0.86 (1.62)	-
Dcien	-2.46 (3.27)	-	-	-6.35 (3.80)
Cons	429.77 (7.26)	422.32 (8.73)	425.46 (11.1)	438.53 (8.15)

Número de Estudiantes: 4179

Notas: Las cifras entre paréntesis corresponden al desvío estándar.

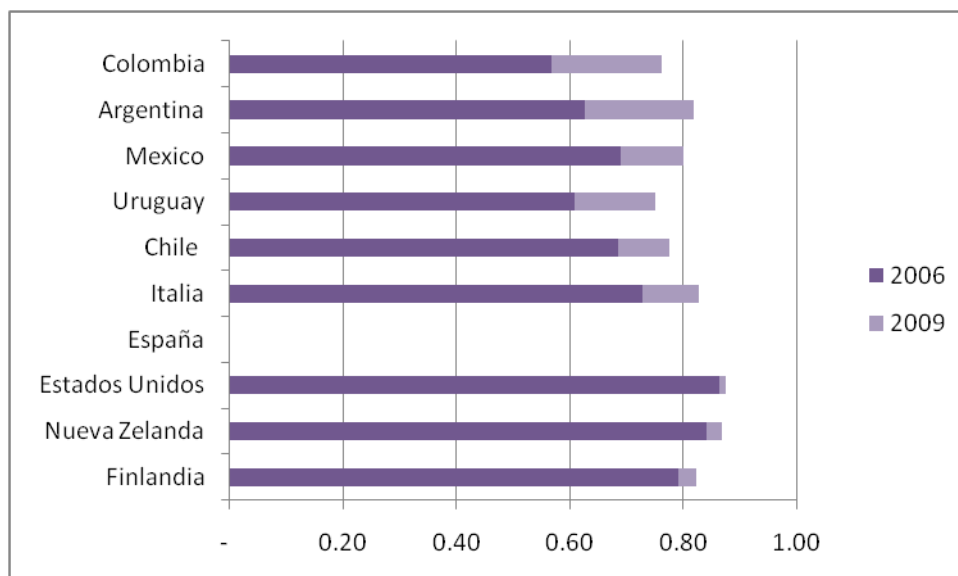
Anexo IV - Resultados de descomposición de varianza

Grafica A.IV.1 - Comparación del $\hat{\rho}_1$ obtenido por país entre años de estudio.



Notas: Los valores estimados de $\hat{\rho}_1$ corresponden al modelo únicamente con constante.

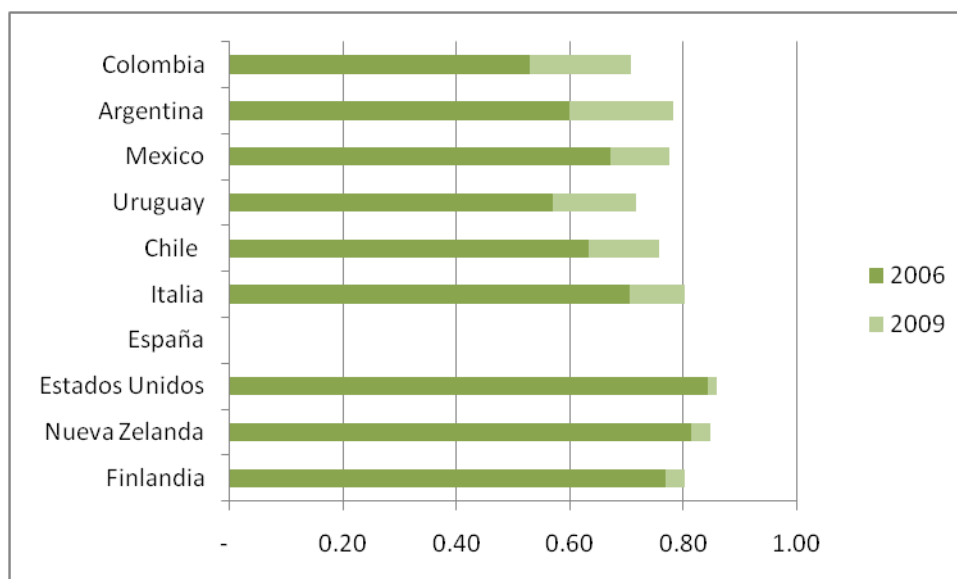
Grafica A.IV.2 - Comparación del $\hat{\rho}_2$ obtenido por país entre años de estudio.



Notas: Los valores estimados de $\hat{\rho}_2$ corresponden a un modelo con constante y variables de centro.

España no presenta resultados en este caso debido a la falta de datos correspondientes a las variables de centro.

Grafica A.IV.3 - Comparación del $\hat{\rho}_3$ obtenido por país entre años de estudio.



Notas: Los valores estimados de $\hat{\rho}_3$ corresponden al modelo completo: constante más variables de centro y del estudiante. España no presenta resultados en este caso debido a la falta de datos correspondientes a las variables de centro.

Cuadro A.IV.1 - Resultados de la estimación por MCO de regresiones que vinculan el puesto en el ranking con los $\hat{\rho}$ presentados.

Años	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$
2006	-149.62 (53.78)	-132.19 (37.60)	-111.81 (36.63)
2009	-276.05 (161.63)	-309.39 (111.77)	-315.35 (102.07)

Número de países: 10

Notas: Las cifras entre paréntesis corresponden al desvío estándar.

Anexo V - Análisis de la influencia de factores del estudiante y del centro educativo en el rendimiento en las pruebas PISA.

Para realizar la estimación del efecto del tiempo en clase sobre los resultados de los estudiantes en las pruebas PISA, se incorporaron al modelo variables que tienen relación con características del estudiante y con características del centro educativo al que asiste. En esta sección, se analiza la influencia de dichos factores en los resultados académicos obtenidos por los estudiantes en los años 2006 y 2009 a través de la estimación del modelo bietápico de Heckman.

En los cuadros A.III.3 y A.III.4 se detallan las variables que fueron incorporadas en la estimación de los modelos para los años 2006 y 2009 respectivamente y los resultados de los mismos. La diferencia en la cantidad de variables consideradas en cada año se debe a que, al momento de presentación de este trabajo, no se encuentra publicada la base nacional para el año 2009.

Factores del Estudiante

En primer lugar analizamos los coeficientes asociados a las variables de los estudiantes, buscando reflejar de qué forma estas características de los individuos que son independientes del centro educativo al que asisten al momento de realizar la prueba, impactan sobre su rendimiento académico.

Para los dos años bajo estudio, la variable sexo tuvo efectos significativos sobre el rendimiento de los estudiantes en matemáticas y en ciencias. Se observa que el hecho de ser mujer, tiene un efecto negativo sobre los resultados de éstas dos disciplinas. En cambio para lectura, el efecto de ser mujer es significativo y positivo únicamente para explicar el resultado en el año 2009.

Para estudiar el efecto del contexto sociocultural del alumno se incorporó la variable Índice de estatus económico, social y cultural del hogar (iestatuseco). El coeficiente asociado a este índice resultó positivo y significativo en matemáticas y ciencias para los dos años. Sin embargo para lectura, el

coeficiente, a pesar de que se mantiene positivo, es significativo únicamente para el 2009. Estos resultados sugieren que cuanto mejor es el contexto familiar en el que se encuentre un estudiante, mejores son los resultados que obtiene en las tres asignaturas.

La variable Tamaño de la localidad (tamañoloc) fue incluida para identificar en qué medida el lugar de residencia del estudiante puede afectar sus resultados académicos. Mientras que en el año 2006 esta variable no resultó significativa en ninguna de las tres disciplinas, para el 2009 tuvo efectos significativos para explicar el rendimiento en matemática (a un nivel de confianza del 95 por ciento) y el rendimiento en lectura (a un nivel de confianza del 90 por ciento).

Por la forma en que fue construida esta variable, el signo negativo de los coeficientes implica que, en promedio, los estudiantes de pequeñas o medianas localidades obtienen peores resultados en sus pruebas de matemática y lectura que los habitantes de grandes ciudades. Sin embargo, no hay evidencia suficiente para afirmar que lo mismo ocurre en relación a los resultados en ciencias.

Por su parte, las variables relacionadas al tipo de centro en el que los estudiantes realizaron sus estudios previos, fueron observadas únicamente en el año 2006, incluyéndose por este motivo exclusivamente en el modelo correspondiente a ese año.

En cuanto al centro educativo en el cual el niño cursó la enseñanza primaria, únicamente la variable Escuela Tiempo Completo Pública (ETCpublica) resultó significativa para las asignaturas matemática (a un nivel de confianza del 90 por ciento) y ciencias (a un nivel de confianza del 95 por ciento), siendo el efecto de haber asistido a estas escuelas, negativo sobre el rendimiento de los estudiantes en las competencias evaluadas. No obstante, posibles problemas de endogeneidad de esta variable no nos permiten concluir sobre su efecto en las competencias de los estudiantes a los 15 años. Como fue desarrollado en el capítulo de antecedentes, Llambí y Perera (2008) tratan la endogeneidad de la variable ETC mediante el uso de Variables Instrumentales y un modelo de selección de Heckman (HECK), llegando a la conclusión de que a partir de la

base PISA no puede inferirse un efecto causal negativo entre la asistencia a una ETC y las competencias evaluadas.

El efecto del centro en el que cursó o está cursando ciclo básico resultó ser dispar para las tres asignaturas y no siempre significativo. En el caso de matemática, los resultados sugieren que el tipo de plan en el que cursó o está cursando ciclo básico tiene efectos significativos en el resultado de la prueba. En promedio los estudiantes que asistieron o asisten a Ciclo Básico Privado obtienen mejores resultados en esta prueba que el resto. Por el contrario, en promedio quienes asistieron o asisten a Ciclo Básico Rural obtienen peores resultados. En cuanto a la competencia lectura, el tipo de plan es significativo para explicar los resultados en esta prueba en todos los casos a excepción de la variable Ciclo Básico Rural. De los resultados obtenidos para esta asignatura se puede deducir que, la probabilidad de obtener mejores resultados es mayor si los estudiantes asistieron o asisten a Ciclo Básico Público y menor si asistieron o asisten a Ciclo Básico Técnico. No hay evidencia suficiente para concluir sobre los resultados obtenidos en la prueba ciencias.

Por último, como era de esperarse, se observa que la variable grado tiene un efecto significativo y positivo en la explicación de los resultados de las tres materias en ambos años. Es decir, cuanto más alto sea el año de escolarización alcanzado por el estudiante al momento de tomar la prueba, mejor será su rendimiento.

Factores del Centro Educativo

Luego de haber presentado las características asociadas a los estudiantes, procedemos a analizar los resultados obtenidos para los insumos educativos.

En cuanto al Tipo de Centro, que indica la asistencia del estudiante a un centro educativo privado, para el año 2006 dicha variable resultó significativa únicamente para explicar los resultados obtenidos en la prueba lectura. A sí mismo, el coeficiente positivo asociado a la variable indica que los individuos que asisten a centros privados obtienen en promedio 40.43 puntos más en la prueba lectura dado todo lo demás constante. Para el año 2009, el efecto del

tipo de centro es significativo y positivo para explicar el rendimiento en las tres asignaturas. Los estudiantes que asisten a centros privados obtienen en promedio, dado todo lo demás constante, 39.46 puntos más en matemática, 34.97 en ciencias y 47.21 en lectura.

Por su parte, el efecto del tamaño del centro educativo es positivo pero bastante pequeño, y no siempre resulta estadísticamente significativo. En los dos años bajo análisis, el tamaño del centro es significativo únicamente para explicar el resultado en las competencias ciencia (a un nivel de confianza de 95 por ciento para el 2009) y lectura (a un nivel de confianza del 95 por ciento para los dos años).

Por otro lado, respecto al resto de las variables incluidas en el modelo, se encuentra que ni el ratio de Escasez Docente (*escasezdoc*) ni la Proporción de Docentes con Título Universitario (*propdocuniv*) resultan significativos en ninguno de los dos años. No obstante, la Proporción de Docentes Titulados (*propdoctitulados*) para el año 2006 y 2009 resulta significativa para explicar los resultados en ciencias a niveles de confianza del 95 por ciento y 90 por ciento respectivamente. Es decir, a mayor cantidad de docentes titulados en el centro educativo, mejores serán los resultados obtenidos por los estudiantes en ciencias.

Finalmente, debemos presentar los resultados de la estimación del parámetro ρ asociado a la variable Inverso del Ratio de Mills (*lambda*). Este coeficiente cuantifica en qué medida los resultados educativos no son independientes de la probabilidad de asistir a la enseñanza media (Méndez y Zerpa, 2009).

En los dos años bajo estudio, se obtienen para las tres disciplinas, coeficientes negativos y significativos a un nivel de confianza del 95 por ciento. Tal como señalan Méndez y Zerpa (2009), al tomar en cuenta a los jóvenes desertores en el modelo, se reduce el sesgo de selección y los resultados se ajustan a la baja.