



EL COLEGIO DE MÉXICO

CENTRO DE ESTUDIOS ECONÓMICOS

MAESTRÍA EN ECONOMÍA

Trabajo de Investigación para obtener el grado de Maestro en Economía

Determinantes de la Diferencia en
Calificaciones entre Escuelas Públicas
y Privadas Usando la Prueba ENLACE

LUIS ALBERTO MARTÍNEZ CHIGO

PROMOCIÓN 2009-2011.

DR. RAYMUNDO MIGUEL CAMPOS VAZQUEZ

2012



Biblioteca Daniel Cosío Villegas
EL COLEGIO DE MÉXICO, A.C.

Determinantes de la Diferencia en Calificaciones entre Escuelas Privadas y Públicas Usando la Prueba Enlace

Resumen

El presente trabajo analiza la diferencia de calificaciones entre las escuelas privadas y públicas utilizando los resultados de la Evaluación Nacional del Logro Académico en Centros Escolares. La metodología desarrollada por DiNardo, Fortin, & Lemieux (1996) permite estimar diversos contrafactuales y estudiar distintos puntos de la distribución de las calificaciones como el percentil 10. En particular, se estima la distribución de calificaciones de estudiantes de escuelas públicas si tuvieran la distribución de características de estudiantes de escuelas privadas. Posteriormente se descompone la aportación de las características observables sobre grupos de variables que reflejan las características de los estudiantes, familiares, recursos escolares y factores institucionales. Los principales resultados sugieren que gran parte de la diferencia en calificaciones se explica vía diferencias en características observables (80 % - 95 %). Las contribuciones más importantes en estudiantes de bajo rendimiento son los factores escolares y características familiares mientras que en estudiantes de alto rendimiento los factores institucionales y características individuales adquieren relevancia. Esto significa que las aportaciones de los grupos antes mencionados varían conforme al rendimiento de los estudiantes.^a

^aLos errores son exclusiva responsabilidad del autor.

Índice general

1. Introducción	1
2. Revisión de la Literatura	5
3. Metodología.	10
3.1. Estimaciones de Densidad Kernel	13
3.2. Di Nardo, Fortin, Lemieux (1996)	13
3.3. Estrategia Empírica para DFL	16
3.4. Retornos	17
3.5. Una descomposición detallada	21
4. Datos y Estadísticas Descriptivas	24
4.1. Bases de Datos	24
4.2. Estadísticas Descriptivas	26
4.3. Diferencia en Características Observables.	38
4.4. Especificación Econométrica	39

5. Resultados	41
5.1. Di Nardo, Fortin & Lemieux	41
5.2. Descomposición Detallada	42
5.3. Descomposición Independiente	50
5.4. Diferencias en Retornos: McCrary <i>et al.</i> (2010)	53
5.5. El Impacto del Pasado	56
5.6. Implicaciones de Política Pública	58
6. Conclusiones	60
Appendices	69
A. Diversos	69

Capítulo 1

Introducción

La educación ha sido un tema central y muy estudiado en economía. Diferentes investigaciones han señalado la importancia de la educación sobre diversos aspectos individuales. Por ejemplo, [Card \(1993\)](#) encuentra una relación positiva entre educación y el ingreso, mientras que [Lleras-Muney \(2005\)](#) encuentra un efecto positivo de los años de educación sobre la salud de los individuos. Respecto a beneficios sociales, [Lochner & Moretti \(2004\)](#) afirman que existe una relación negativa entre educación y crimen, [De Gregorio & Lee \(2003\)](#) encuentran que mayor educación implica una mayor equidad en la distribución del ingreso. En términos de políticas públicas, la educación es parte esencial, sin duda alguna, en las agendas de todos los gobiernos.

En México, el gasto público en educación, de acuerdo con el Banco de México, representa el 5 % del Producto Interno Bruto (PIB) y 22 % del gasto público total. Esto posiciona a México en la cima de todos los países de la OECD respecto a porcentaje del gasto público¹ e inclusive encima del gasto promedio mundial como porcentaje del PIB.² El sistema educativo mexicano se conforma por Educación Básica, la cual se divi-

¹OECD (2010): *PISA 2009 Results: Executive Summary*

²The World Bank (2010): *Education Data*, <http://data.world.org/topic/education>

de en educación primaria (6 años) y secundaria (3 años), Educación Media Superior (3 años) y Educación Superior (2-5 años). Desde la década de 1990 el gobierno mexicano enfocó los esfuerzos a incrementar la matrícula escolar provocando un aumento en el número de estudiantes de 34.5 %.³ Sin embargo, a pesar de la considerable cantidad del gasto público destinado a educación y de ciertos objetivos alcanzados, en temas de calidad educativa aún falta mucho camino por recorrer. Prueba de ello, son los resultados que México ha obtenido en exámenes estandarizados como el Programa Internacional para la Evaluación de los Estudiantes (PISA). En PISA 2009, México se encuentra en el último lugar de todos los países de la OECD con una calificación promedio de 425 puntos. Esto significa que la brecha entre Corea, país ubicado en el primer lugar, y México es de 114 puntos, lo que equivale a más de dos años de educación.⁴

Además, las diferencias en la calidad de la educación se presentan a nivel nacional. Las escuelas en México se diferencian en términos de quien proveen sus recursos: el sector privado o el sector público. Para 2009, sólo 9.3 % de los jóvenes en educación básica se encuentran en escuelas pertenecientes al sector privado y tal porcentaje se incrementa a 18.5 % para el nivel medio superior.⁵ En México, la educación privada es asociada con una educación de mayor calidad. Por ejemplo, en PISA 2009 las escuelas privadas obtuvieron 10 % más calificación que las escuelas públicas en todas las áreas evaluadas, lo cual equivale a un año escolar de diferencia. [Guichard \(2005\)](#) usando los resultados de PISA 2004 señala diferencias similares y sugiere que pueden reflejar el estado socioeconómico de los estudiantes.

Las grandes diferencias tanto internacionales como nacionales han hecho de la calidad educativa un tema relevante en México. Muestra de ello es la creación de la Evaluación Nacional del Logro Académico en Centro Escolares (ENLACE). ENLACE es un exa-

³SEP (2010): *Principales Cifras Ciclo Escolar 2008-2009*

⁴OECD (2010): *PISA 2009 Results: Executive Summary*, pág. 6

⁵SEP (2010): *Principales Cifras Ciclo Escolar 2008-2009*

men estandarizado aplicado de forma centralizada desde 2006 para Educación Básica y 2008 para Educación Media Superior. Durante su aplicación se incluyen cuestionarios de contexto a estudiantes, padres de familia, directores y docentes. Estos cuestionarios se aplican aleatoriamente a nivel escuela a uno de cada cuatro estudiantes y su objetivo es recabar información sobre características individuales (sexo y promedio del nivel educativo anterior), características del hogar (número de cuartos y libros en el hogar) y recursos escolares (número de laboratorios y estado de las aulas).

El objetivo principal del trabajo es investigar los factores causantes de la diferencia en calificaciones entre las escuelas privadas y escuelas públicas en Educación Media Superior usando los resultados de la prueba ENLACE y la información de los cuestionarios de contexto para 2008. Respecto a la metodología, seguimos la propuesta de [DiNardo, Fortin, & Lemieux \(1996\)](#) para analizar las distribuciones de las calificaciones de estudiantes en escuelas de Educación Media Superior. La metodología nos permite estimar contrafactuales bajo ciertos supuestos. En particular, es de interés el contrafactual “¿cómo se distribuirían las calificaciones de estudiantes en escuelas públicas si tuvieran las características de los estudiantes en escuelas privadas?”.

Los resultados indican que cerca de un 90 % de la diferencia en promedio entre estudiantes de escuelas privadas y públicas se explica por la diferencia en características individuales, familiares, escolares e institucionales. Además, en estudiantes de bajo desempeño los factores escolares tienen un efecto mayor al señalado por la literatura. En cambio, en estudiantes de alto rendimiento los recursos escolares pierden importancia y son los factores asociados a aspectos sociales de la escuela los que influyen de forma positiva e importante. Para estudiantes en la media, los resultados se encuentran en línea con los obtenidos por la literatura previa.

A pesar de que las características individuales decrecen en importancia conforme avanzamos en la distribución, estas explican buena parte de la diferencia en calificación.

nes (17 %-44 %). En una segunda etapa, clasificamos las características del estudiante en dos categorías: características pasadas y características presentes, cuya finalidad es revisar la importancia de inversiones educativas en una etapa temprana de los estudiantes (Heckman & Masterov, 2007). Los resultados señalan que las características presentes no contribuyen a la diferencia en calificaciones entre estudiantes de escuelas privadas y públicas. Además, utilizamos la metodología desarrollada por Leibbrandt, Levinsohn, & McCrary (2010) para corroborar que la diferencia en calificaciones no se debe a distintas funciones de producción educativa. Las estimaciones indican que la diferencia en características observables es la explicación más plausible.

En términos de políticas públicas, los resultados anteriores sugieren una revisión de los programas encaminados a mejorar el desempeño de los estudiantes en dos puntos. Primero, aquellos programas cuyo objetivo es mejorar los factores escolares deben enfocarse en aquellas escuelas con una mayor proporción de estudiantes de bajo rendimiento y a su vez acompañarse de programas que mejoren el ambiente institucional de la escuela. Segundo, en escuelas con estudiantes de alto rendimiento deben crearse programas que fomenten la integración de los estudiantes y mejoren los aspectos institucionales, como asociaciones estudiantiles o mejorar la relación con los profesores. En pocas palabras, el presente trabajo permitirá el diseño de programas más focalizados.

El resto del trabajo está organizado como sigue. El desarrollo teórico y la estrategia empírica se incluyen en la sección 2. En la sección 3 se resume la evidencia empírica relacionada con el efecto de diversos factores tanto individuales y escolares sobre el desempeño escolar, y las investigaciones comparables. En la sección 4 se describe la base de datos y se muestran algunas estadísticas descriptivas de interés. La sección 5 resume las estimaciones y resultados encontrados. Las conclusiones e implicaciones de los resultados sobre la política pública se presentan en la sección 6.

Capítulo 2

Revisión de la Literatura

En economía de la educación la mayor parte de los estudios se han centrado en determinar qué factores son esenciales para el aprendizaje escolar. El primer problema con el que se topa la investigación es definir, de forma precisa, “¿Qué es el aprendizaje escolar?”. En este caso, no se pretende abordar esta discusión y se seguirán las líneas que ha marcado la literatura, la cual asocia el aprendizaje escolar con el desempeño de los estudiantes en exámenes estandarizados. El objetivo principal de los trabajos relacionados con el aprendizaje escolar es estimar una función de producción educativa de la siguiente forma:

$$P_i = m(Z_i, \xi_i) \quad (2.1)$$

donde P_i es la calificación obtenida en exámenes estandarizados, Z_i es un conjunto de características observadas y ξ_i un término de error. En la literatura la medida más usada como P_i son los resultados del Programa Internacional para la Evaluación de los Estudiantes mejor conocido como PISA. Para Z_i se puede enunciar a tres grupos de variables que la literatura considera relevante para explicar las calificaciones obtenidas por los estudiantes: características individuales y familiares, recursos e infraestructura

escolar y efectos institucionales.

En el grupo de características individuales y familiares se pueden encontrar variables como edad, sexo, grado de estudio de los padres, número de computadoras en el hogar, cantidad de libros en la casa, grado de marginación de la localidad y tutorías privadas. Por ejemplo, [Kang \(2007\)](#) usando el Panel de Educación y Empleo Coreano (KEEP) encuentra que incrementar un 10 % los gastos en tutorías privadas implica un aumento de 1.5 puntos porcentuales sobre el puntaje medio. [Woessmann \(2000\)](#) con los resultados del Tercer Estudio Internacional en Matemáticas y Ciencias (TIMSS) muestra que la cantidad de libros en el hogar es un buen proxy para el pasado familiar y puede resultar en una diferencia hasta de un 8 % más calificación sobre la media. [Fuchs & Woessmann \(2004\)](#) encuentran resultados similares a [Woessmann \(2000\)](#) usando datos del Programa Internacional para la Evaluación de Estudiantes a nivel internacional y para el año 2000.

A nivel de recursos e infraestructura escolar han sido incluidas variables como el número de computadoras en la escuela, experiencia de los profesores, número de profesores por alumno y grado de inasistencia del docente ([Angrist & Lavy, 1999](#); [Dearden, Ferri, & Meghir, 2002](#); [Backhoff, Bouzas, Contreras, Hernández, & García, 2007](#)). [Graddy & Stevens \(2005\)](#) usando datos para el Reino Unido encuentran que una disminución del 10 % sobre la razón estudiantes-profesores refleja un incremento del 4 % sobre la calificación media para el examen de ingreso a la universidad. [De Hoyos, Espino, & García \(2010\)](#) usando los resultados de ENLACE 2008 para México, muestra que los recursos escolares pueden explicar hasta un 18 % de la variación de las calificaciones. En general no existe un consenso sobre el efecto de las variables escolares sobre el desempeño de los estudiantes. [Hanushek \(2003\)](#) realiza un meta-análisis donde llega a la conclusión de que no existe evidencia contundente de que los recursos escolares impacten al desempeño de los estudiantes. Sin embargo, Hanushek hace énfasis en que lo anterior no implica que los recursos escolares nunca importen y que esta “falta de relación”

puede ser consecuencia de ineficiencias durante la aplicación de los programas.

Un enfoque más reciente se ha preocupado por incluir variables que reflejen el entorno social e institucional donde se encuentra inmersa la institución. Este enfoque introduce a la función de producción variables para determinar el impacto de los sindicatos, así como de la independencia de las escuelas en toma de decisiones, sobre la calificación de los estudiantes. [Alvaréz, García-Moreno, & Patrinos \(2007\)](#) con datos de PISA 2003 para México, encuentra que los estados donde existe una fuerte influencia del sindicato tienen en promedio 10 puntos menos que aquellos donde la influencia es poca. [Fuchs & Woessmann \(2004\)](#) afirma que debido a las variables institucionales, como grado de autonomía en contratar profesores y autonomía en decidir los programas de estudio, los estudiantes de escuelas privadas tienen un mejor desempeño.

Por otro lado, algunos estudios se han preocupado por determinar cuáles son las causas en las diferencias en calificaciones entre diversos grupos. Por ejemplo, la literatura concuerda en el hecho de que los hombres tienen mejores puntajes en matemáticas y las mujeres en español. [Sohn \(2010\)](#) utilizando el Estudio Longitudinal de Infancia Temprana (ECLS) analiza la diferencia en calificaciones en matemáticas entre hombres y mujeres usando regresiones intercuantiles. Uno de sus principales resultados es que la brecha nace durante los primeros años de educación y se exagera durante años posteriores. [McEwan \(2004\)](#) analiza la diferencia en puntajes entre estudiantes de grupos indígenas y no indígenas en Bolivia y Chile. McEwan concluye que alrededor de un 50 % de la brecha se puede explicar por los efectos de grupo y sólo un 20 % con las características familiares. Dentro de esta línea, la literatura también se ha enfocado a estudiar las diferencias en calificaciones entre escuelas privadas y públicas bajo el argumento que las primeras tienen un mejor desempeño. [Peterson & Llaudet \(2006\)](#) usando datos de la Evaluación Nacional del Progreso Educativo (NAEP) en Estados Unidos encuentran que las escuelas privadas tienen entre 5 % y 6 % más puntaje sobre la media.

Bedi & Garg (2000) encuentra que los estudiantes de escuelas privadas alcanzan mayores salarios en Indonesia, a pesar de que las escuelas públicas son percibidas como las mejores. **Yamauchi (2005)** obtiene resultados similares a **Bedi & Garg (2000)** para Filipinas y Tailandia pero atribuye esta diferencias a que los estudiantes de escuelas privadas se seleccionan. **Vandenberghe & Robin (2004)** corrige el problema de selección usando diversas metodologías y obtiene que no existen beneficios importantes de pertenecer a las escuelas privadas, los resultados son a nivel internacional usando los datos de PISA 2000. Respecto a América Latina, gran parte de los trabajos que comparan a las escuelas públicas y privadas se centran en Chile gracias a su sistema de vouchers, concluyendo que no existe diferencia significativa entre los dos sistemas.

En México, la falta de exámenes estandarizados para evaluar al sistema educativo limitó el número de trabajos de corte cuantitativo. Sin embargo, durante los últimos años, los resultados de México en exámenes internacionales como PISA y exámenes nacionales, como la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE), ha revivido el debate de la calidad educativa y ambos exámenes proporcionan información relevante para el análisis empírico. **Backhoff, Bouzas, Contreras, Hernández, & García (2007)** usando análisis multinivel y los resultados para educación básica de los Exámenes de la Calidad para el Logro Educativo (EXCALE) estiman el peso de grupos de variables sobre la varianza de los resultados. Concluyen que las características de los estudiantes contribuyen entre 65-75 % al poder explicativo del modelo y el porcentaje restante es atribuido a variables escolares.

De Hoyos, Espino, & García (2010) explican la varianza de los resultados en la prueba ENLACE 2008 para Educación Media Superior y además utilizan análisis discriminante para obtener el modelo con mayor poder explicativo. Como resultado su modelo logra explicar el 70 % de la varianza de los resultados y la mayor aportación a este porcentaje la tienen variables a nivel estudiante (60 %) seguida de variables escolares (18 %), facto-

res institucionales (14 %) y características familiares (6 %). Ambos estudios han incluido variables dicotómicas para controlar por “sostenimiento” o tipo de escuela. De Hoyos, Espino, & García (2010) concluye que el efecto de estar en un escuela privada es casi nulo. Backhoff, Bouzas, Contreras, Hernández, & García (2007) en su modelo completo encuentra que el efecto de estar en una escuela privada se traduce en 30 puntos, es decir, ente seis y ocho puntos porcentuales más sobre la calificación de escuelas públicas.

Para ubicar el presente trabajo en la literatura económica, hay que mencionar que existen pocos estudios previos y comparables (Duncan & Sandy, 2007) y se suma a la poca literatura cuantitativa existente relacionada con la educación en México (De Hoyos, Espino, & García, 2010; Backhoff, Bouzas, Contreras, Hernández, & García, 2007). La investigación se distingue de los trabajos anteriores por dos razones. La primera es el enfoque, pues estudia de forma directa la brecha existente entre escuelas públicas y privadas incorporando la idea de desagregar esta brecha en grupos de variables observables. La segunda razón es que explora otros puntos de la distribución, como los percentiles 10 y 90, bajo la idea de que los recursos escolares o características individuales pueden contribuir de distinta forma sobre estudiantes de bajo y alto desempeño. En este sentido, son pocos los trabajos que hacen énfasis en analizar diversos puntos de la distribución (Sohn, 2010).

Capítulo 3

Metodología.

La metodología desarrollada por [DiNardo, Fortin, & Lemieux \(1996\)](#) pertenece a la familia de métodos noparamétricos y desarrollados para superar las limitantes, como análisis en la media, de los métodos paramétricos convencionales como mínimos cuadrados ordinarios. La idea principal detrás de DFL está muy relacionada con el problema fundamental de la inferencia causal, es decir, no observamos el efecto del tratamiento para el grupo de control ([Holland, 1986](#)). Para nuestro caso, atribuir el efecto de estar en la escuela privada o pública a la diferencia entre medias de las calificaciones es completamente incorrecto. Por ejemplo, consideremos:

$$\hat{\theta} = E[P_{priv}|Priv] - E[P_{pub}|Pub], \quad (3.1)$$

que resulta ser la diferencia entre el grupo de estudiantes de escuelas privadas y escuelas públicas, $\hat{\theta}$ es un estimador insesgado si los datos no tienen problemas de selección. Como no es el caso, es incorrecto tomar a $\hat{\theta}$ como el efecto total de asistir a una escuela privada. Una forma de resolver el problema de selección es asignar a los estudiantes a escuelas públicas y privadas de forma aleatoria. Llevar a cabo el experimento

es casi imposible aún sin tomar en cuenta problemas e implicaciones éticas. Por tanto, para hacer una comparación más válida es necesario introducir un término más a la Ecuación (3.1).

$$\theta = \underbrace{E[P_{priv}|Priv] - E[P_{pub}|Priv]}_1 + \underbrace{E[P_{pub}|Priv] - E[P_{pub}|Pub]}_2. \quad (3.2)$$

El término $E[P_{pub}|Priv]$ en economía es usualmente conocido como “contrafactual” e intuitivamente corresponde a la calificación promedio de los estudiantes de escuelas privadas si estuvieran en una escuela pública. Al primer conjunto de términos de la Ecuación (3) se le conoce como el efecto promedio sobre los tratados (ATT) y se puede ver como la diferencia atribuible a las formas funcionales o funciones de producción. En otras palabras, el primer conjunto está comparando los sistemas públicos y privados, es decir, toma al grupo de estudiantes de escuelas privadas, mide su desempeño tanto en escuela privada como pública y la diferencia en desempeño es una mejor aproximación de los beneficios de un sistema u otro. El segundo conjunto compara a los dos grupos de estudiantes suponiendo que ambos estudiaron en el sistema público. La diferencia obtenida se atribuye a la composición de los grupos y si los grupos son homogéneos entonces esperaríamos que su valor sea cero.

Sin embargo, los datos sólo nos proporcionan información sobre $E[P_i|i]$ con $i = Priv, Pub$ y por tanto es necesario obtener una estimación para $E[P_{pub}|Priv]$. De acuerdo con la literatura previa (De Hoyos, Espino, & García, 2010; Duncan & Sandy, 2007) los beneficios de estudiar en una escuela privada son cercanos a cero, en otras palabras, en la Ecuación 3 el conjunto marcado con uno tiene un valor cercano a cero. Por ello, uno de los objetivos centrales del trabajo es estimar el término no observado $E[P_{pub}|Priv]$ y poner énfasis en la parte 2 de la Ecuación que corresponde a la parte explicada por las características observables de los estudiantes. Aunque, una vez dada la estimación,

obtener un valor para el conjunto 1 es trivial y se incluye en el trabajo.

Es importante mencionar que la descomposición propuesta en la Ecuación (3) no es única gracias a que $E[P_{pub}|Priv]$ no es el único contrafactual que se puede construir. También se puede estimar $E[P_{priv}|Pub]$ que corresponde a las calificaciones de estudiantes de escuelas públicas si estudiaran en una escuela privada. Los dos contrafactuales son estimables y la elección de cual estimar es equivalente a qué grupo de referencia se desee. Se seguirá la literatura previa y tomaremos como grupo de referencia al grupo desfavorecido, en este caso, a las escuelas públicas y se estimaran las calificaciones de estudiantes de escuelas privadas si estuvieran en una escuela pública ($E[P_{pub}|Priv]$). Además, resultará más enriquecedor tener resultados sobre el sistema de educación pública pues se genera un margen para la elaboración de políticas públicas.

Fortin, Lemiux, & Firpo (2010) proponen que la descomposición de la Ecuación (3) es un caso especial de una descomposición más general.

$$\Delta^v = v(F_{P_{priv}|priv}) - v(F_{P_{pub}|priv}^C) + v(F_{P_{pub}|priv}) - v(F_{P_{pub}|pub}). \quad (3.3)$$

Donde v es cualquier estadístico y $F_{Y_i|i}$ es la distribución acumulada. Entre los estadísticos más usados se encuentran el índice de Gini, diferencias entre percentiles y diferencias entre medias. La Ecuación (3.3) posee el mismo problema que la ecuación (3) en el sentido de que los datos no proporcionan, de forma directa, información sobre la distribución $F_{Y_{pub}|priv}^C$. La metodología DFL permite reducir el problema de estimar la distribución del contrafactual “¿cuáles son las calificaciones de estudiantes de escuelas privadas si estuvieran en una escuela pública” ($E[P_{pub}|Priv]$) a una aplicación conjunta de métodos de Kernel y métodos paramétricos.

3.1. Estimaciones de Densidad Kernel

Para estimar funciones de densidad se hace uso de los estimadores Kernel propuestos por Rosenblatt (1956) y Parzen (1962). Ellos proponen que el estimador de la función de densidad unidimensional de una muestra P_1, P_2, \dots, P_n con pesos $\theta_1, \theta_2, \dots, \theta_n$ es:

$$\hat{f}(p) = \sum_{i=1}^n \frac{\theta_i}{h} K\left(\frac{p - P_i}{h}\right), \quad (3.4)$$

donde h es el bandwidth y $K(\cdot)$ es la función kernel. Existen ventajas y desventajas de usar este método no paramétrico. La principal ventaja es que no se asume cierta distribución de los datos; en otras palabras los datos hablan por sí solos y que la elección de K no modifica de forma radical la forma de la función de densidad. Las desventajas radican en la elección del bandwidth (h) pues se sabe que existe un intercambio, entre precisión y forma de la función de densidad, al momento de elegir un valor. Para resolver el problema seguiremos la literatura alrededor de los estimadores kernel y elegiremos el bandwidth propuesto por Silverman (1986).¹ En el caso de la función $K(\cdot)$ se usa el kernel de Epanechnikov.

3.2. Di Nardo, Fortin, Lemieux (1996)

Podemos asignar a cada alumno de la base de datos de las pruebas ENLACE a un vector (p, z, g) donde p es el puntaje o calificación obtenida en la prueba, z es un conjunto de características individuales, familiares y a nivel escuela y g el tipo de escuela (privada o pública). La función de densidad para las calificaciones de escuelas privadas y públicas se pueden escribir como:

¹El bandwidth óptimo se puede expresar como $h = 1.06 N^{-\frac{1}{5}} \min[\sigma, \frac{R}{1.34}]$ donde σ es la desviación estándar, R la diferencia intercuartil y N el número de observaciones.

$$f(p|priv) = \int m(p|z, priv)h(z|priv)dz \quad (3.5)$$

$$f(p|pub) = \int m(p|z, pub)h(z|pub)dz, \quad (3.6)$$

donde $m(p|z, g)$ es la función de densidad de las calificaciones condicionada a las características y tipo de escuela; $h(z|g)$ es la función de densidad de las características individuales sujeto al tipo de escuela. De acuerdo a [Leibbrandt, Levinsohn, & McCrary \(2010\)](#) la función $m(p|z, g)$ se puede ver como una función de producción que transforma las características individuales a una calificación, $h(z|g)$ como la parte que introduce las características y $f(p|g)$ como la variable dependiente, que en este caso son las calificaciones.

El propósito de la investigación es dar una estimación para la función de distribución de las calificaciones de los alumnos de escuelas privadas si estuvieran en escuelas públicas. El enunciado anterior es un abuso del lenguaje pues no se puede cambiar a un estudiante de escuela privada a una escuela pública. Sin embargo, es posible tomar la distribución de características de los estudiantes de escuelas privadas e imponerla a estudiantes de escuelas públicas. Formalmente, se estimará la distribución de calificaciones de estudiantes de escuelas públicas si tuvieran la distribución de características observables de estudiantes de escuelas privadas. Por tanto, gracias a que la relación entre la función de distribución y la función de densidad es natural, podemos escribir lo anterior en términos de la Ecuación 3.6 como:

$$f_h^{pub \rightarrow priv} = \int m(p|z, pub)h(z|priv)dz, \quad (3.7)$$

donde la notación $f_h^{g \rightarrow g'}$ significa que la distribución original es g y sustituimos la

parte correspondiente a la función h de la distribución g' . Por ejemplo, si se estuviera interesado en estimar la función de densidad de las calificaciones de los alumnos de escuelas públicas si tuvieran las características de los estudiantes de escuelas públicas estaríamos estimando $f_h^{priv \rightarrow pub}$ y se puede escribir como:

$$f_h^{priv \rightarrow pub} = \int m(p|z, priv)h(z|pub)dz. \quad (3.8)$$

Estimar (3.7) y (3.8) es bastante complejo y de hecho es un problema multidimensional. El valor agregado de DFL radica en reducir la estimación de (3.7) a un problema unidimensional y observado en los datos, usando ciertos “pesos”. Es decir, la ecuación (3.7) se puede escribir como:

$$f_h^{pub \rightarrow priv} = \int m(p|z, pub)h(z|priv)dz \quad (3.9)$$

$$= \int m(p|z, pub) \frac{h(z|pub)}{h(z|pub)} h(z|priv)dz \quad (3.10)$$

$$= \int m(p|z, pub)h(z|pub) \frac{h(z|priv)}{h(z|pub)} dz \quad (3.11)$$

$$= \int m(p|z, pub)h(z|pub)\Phi(z)dz. \quad (3.12)$$

La única diferencia entre las funciones de densidades (3.12) y (3.6) es la función $\Phi(z)$. Por tanto, para estimar la función de densidad del contrafactual sólo es necesario estimar $\Phi(z)$ y eso lo obtenemos gracias a la regla de Bayes.

$$\frac{h(z|priv)}{h(z|pub)} = \frac{\frac{P(priv|Z=z)f(z)}{P(priv)}}{\frac{P(pub|Z=z)f(z)}{P(pub)}} = \frac{P(priv|Z=z)}{P(pub|Z=z)} \cdot \frac{P(pub)}{P(priv)} \quad (3.13)$$

$$= \frac{P(priv|Z=z)}{1 - P(priv|Z=z)} \cdot \frac{1 - P(priv)}{P(priv)}. \quad (3.14)$$

Para estimar $P(priv|Z = z)$ se puede hacer uso de un modelo logit o probit y $P(priv)$ corresponde a las probabilidades de la muestra, es decir, número de alumnos en escuelas privadas. Usando estas ideas el estimador para la función $\Phi(z)$ resulta en:

$$\hat{\Phi}(z_i) = \frac{\hat{P}(priv|z_i)}{1 - \hat{P}(priv|z_i)} \cdot \frac{1 - \hat{P}(priv)}{\hat{P}(priv)} \quad (3.15)$$

3.3. Estrategia Empírica para DFL

El desarrollo teórico de las secciones anteriores nos permiten implementar para cada año una serie de pasos para estimar el contrafactual: “¿Cuál sería la distribución de las calificaciones de los estudiantes de escuelas públicas si tuvieran la distribución de características de los estudiantes de escuelas privadas?”

1. Calcular $\hat{P}(priv)$. Es decir, la fracción de estudiantes de escuelas privadas que se encuentran en la base de datos.
2. Usando un modelo logit estimar $\hat{P}(priv|z_i)$ para cada individuo i de la muestra.
3. Estimar la función $\hat{\Phi}(z_i)$ para cada estudiante de la muestra que se encuentra en una escuela pública.

$$\hat{\Phi}(z_i) = \frac{\hat{P}(priv|z_i)}{1 - \hat{P}(priv|z_i)} \cdot \frac{1 - \hat{P}(priv)}{\hat{P}(priv)}.$$

4. Usando el estimador kernel para la muestra de estudiantes en escuelas públicas con los pesos $\Phi_1, \Phi_2, \dots, \Phi_n$ obtenemos:

$$\hat{f}(p) = \sum_{i=1}^n \frac{\hat{\Phi}(z_i)}{h} K\left(\frac{p - P_i}{h}\right). \quad (3.16)$$

La idea intuitiva de DFL es sencilla. Por ejemplo, supongamos que para el estudiante j la probabilidad de estar en escuela privada es bastante baja ($\hat{P}(priv|z_j) \approx 0$) entonces el individuo j obtiene un peso aún más cercano a cero ($\hat{\Phi}(z_j) \approx 0$) y por tanto al momento de estimar la densidad no sería tomado en cuenta. En otras palabras DFL está haciendo un matching o apareamiento entre estudiantes de los dos grupos dependiendo de qué tan parecido son. Si se habla en términos de cálculos computacionales DFL posee una gran ventaja sobre otros métodos pues una integral multidimensional se reduce a una estimación de probabilidades. En términos de eficiencia resultados de [Hirano, Imbes, & Ridder \(2003\)](#) prueban que DFL es un método eficiente.

Existen desventajas que son importantes señalar. La primera es que, con DFL, es realmente difícil obtener una descomposición detallada (factor por factor). La segunda es que las descomposiciones que pueden lograrse no son independientes de la trayectoria que tomemos. La tercera es que el DFL original sólo nos proporciona información sobre cómo las características z_i contribuyen a la brecha existente y no nos proporciona información sobre la “eficiencia” de un grupo en usar sus recursos.

Para la primera desventaja aplicaremos un método desarrollado por [Altonji, Bharadwaj, & Lange \(2008\)](#) que detallaremos más adelante. Para la segunda desventaja tomaremos la trayectoria invertida de la descomposición y para sobrellevar la última desventaja [Leibbrandt, Levinsohn, & McCrary \(2010\)](#) estima la función de calificaciones de las escuelas privadas usando sólo información correspondiente a escuelas públicas para medir la “eficiencia” de DFL.

3.4. Retornos

En el caso de la brecha existente entre estudiantes de escuelas privadas y públicas la literatura, como hemos mencionado, ha encontrado evidencia que las característi-

cas individuales observables pueden explicar cerca del 70 % de la variación total de las calificaciones. En las subsecciones anteriores hemos desarrollado la teoría para decir cuanto de la brecha observada se le puede atribuir a las características z_i . Sin embargo, existe la posibilidad de que exista otra explicación para la brecha existente: que los retornos cambien de un grupo a otro. Por ejemplo, que el uso de la computadora genere beneficios diferentes de un grupo a otro.

En esta sección se está interesado en “¿Cuál sería la distribución de calificaciones de los estudiantes de escuelas publicas si tuvieran los retornos de las características observadas de los estudiantes de escuelas privadas?”. Y para responderla utilizaremos la elegante idea de [Leibbrandt, Levinsohn, & McCrary \(2010\)](#). Sabemos que la función de densidad de las calificaciones de estudiantes de escuelas públicas se puede escribir como:

$$f(p|pub) = \int m(p|z, pub)h(z|pub)dz, \quad (3.17)$$

donde se puede interpretar a la función $f(p|pub)$ como la variable dependiente, $m(p|z, pub)$ como la función de producción de calificaciones y $h(z|pub)$ las características observables.

Si estamos interesados en la distribución de las calificaciones de los estudiantes de escuelas públicas con los retornos de las características de las escuelas privadas entonces, podemos escribir la función de densidad de este contrafactual como:

$$f_m^{pub \rightarrow priv} = \int m(p|z, priv)h(z|pub)dz, \quad (3.18)$$

donde $f_m^{pub \rightarrow priv}$ quiere decir que tomamos la distribución original de las calificaciones de escuelas públicas y le estamos imponiendo la función de producción m de las escuelas privadas. Intuitivamente le estamos imponiendo la eficiencia de las escuelas privadas a las escuelas públicas sin cambiar los estudiantes de las escuelas públicas. Es-

timar la función de densidad descrita en (3.18) nos diría como cambiarían las calificaciones debido a diferencias en la eficiencia de los sistemas. En este sentido la ecuación (3.7),

$$f_h^{pub \rightarrow priv} = \int m(p|z, pub)h(z|priv)dz,$$

la cual corresponde al contrafactual estimado original, se puede interpretar como si estuviéramos cambiando a los estudiantes de las escuelas públicas por estudiantes de escuelas privadas dejando la eficiencia de las escuelas públicas y nos proporciona información de como cambian las calificaciones gracias a diferencias en las variables del vector z .

Antes de continuar hay que hacer algunas observaciones relacionadas con esta metodología. Consideremos la función de densidad del contrafactual “¿Cuál sería la distribución de calificaciones de los estudiantes de escuelas privada si tuvieran las características de los estudiantes de escuelas públicas?” que se puede escribir como:

$$f_h^{priv \rightarrow pub} = \int m(p|z, priv)h(z|pub)dz. \quad (3.19)$$

Matemáticamente las ecuaciones (3.18) y (3.19) no tienen diferencia alguna. Sin embargo, existen diferencias al momento de construir los contrafactuales. Por ejemplo, en la ecuación (3.19) estamos interesados en cambiar las características de los estudiantes de escuelas privadas por aquellas de escuelas públicas, es decir, partimos de información de las escuelas privadas y por tanto el término que tenemos permitido manipular es $h(z|pub)$. En la ecuación (3.18) estamos interesados en imponer los retornos de escuelas privadas a estudiantes de escuelas públicas, iniciamos con información de escuelas públicas y por tanto no podemos maniobrar con $h(z|pub)$. Así que, para obtener una forma simple para (3.18), es necesario utilizar la función de densidad m . Con esta idea en

mano podemos reescribir la ecuación (3.18):

$$\begin{aligned}
f_m^{pub \rightarrow priv} &= \int m(p|z, priv)h(z|pub)dz \\
&= \int m(p|z, priv) \frac{m(p|z, pub)}{m(p|z, pub)} h(z|pub)dz \\
&= \int \frac{m(p|z, priv)}{m(p|z, pub)} m(p|z, pub)h(z|pub)dz \\
&= \int \Gamma(p, z)m(p|z, pub)h(z|pub)dz,
\end{aligned}$$

donde $\Gamma(p, z)$ es un peso que depende de las características z . Usando la regla de Bayes podemos encontrar una forma sencilla de estimar $\Gamma(p, z)$.

$$\begin{aligned}
\Gamma(p, z) = \frac{m(p|z, priv)}{m(p|z, pub)} &= \frac{P(priv|P = p, Z = z)}{P(pub|P = p, Z = z)} \cdot \frac{P(pub|z)}{P(priv|z)} \\
&= \frac{P(priv|P = p, Z = z)}{1 - P(priv|P = p, Z = z)} \cdot \frac{1 - P(priv|z)}{P(priv|z)},
\end{aligned}$$

La parte correspondiente a $P(priv|P = p, Z = z)$ se puede estimar con un modelo probit o logit usando las características observables y la calificación de cada estudiante como variables independientes y para $P(priv|z)$ usar un modelo logit sólo con las características observables como variables independientes. El procedimiento para estimar este contrafactual queda resumido en los siguientes pasos:

1. Usando un modelo logit con la variable dependiente escuela privada y variables independientes las características observadas para todos los estudiantes $\hat{P}(priv|z_i)$
2. Estimar $\hat{P}(priv|p_i, z_i)$ para toda la muestra usando un modelo logit con la calificación del estudiante en las variables independientes.

3. Estimar el peso $\hat{\Gamma}(z)$ para cada estudiante de escuelas públicas.

$$\hat{\Gamma}(p_i, z_i) = \frac{\hat{P}(\text{priv}|P = p, Z = z)}{1 - \hat{P}(\text{priv}|P = p, Z = z)} \cdot \frac{1 - \hat{P}(\text{priv}|z)}{\hat{P}(\text{priv}|z)}. \quad (3.20)$$

4. Usar métodos de kernel para estimar la función de densidad del contrafactual:

$$f_m^{\text{pub} \rightarrow \text{priv}} = \sum_{i=1}^n \frac{\Gamma(p_i, z_i)}{\cdot} hK\left(\frac{p - P_i}{h}\right) \quad (3.21)$$

Los cuatro pasos clarifican un poco más las diferencias entre los contrafactuales que describen las ecuaciones (3.18) y (3.19) pues el último procedimiento se compone por dos estimaciones de modelos logit mientras el DFL original sólo usa una estimación.

3.5. Una descomposición detallada

En términos de política pública, es importante conocer qué factores (computadoras, profesores, idiomas, etc.) están aportando más a la brecha existente entre escuelas privadas y públicas. DiNardo, Fortin, & Lemieux (1996), además de desarrollar un método para estimar contrafactual, proponen un método para obtener la aportación de los sindicatos a la brecha existente. La metodología empleada por ellos es elegante y simple pero depende de una característica: sindicato es una variable binaria. Obtener el peso de un factor cuando este es una variable continua es, en general, complicado.

Para ello, Altonji, Bharadwaj, & Lange (2008) proponen una variante de DFL para obtener el peso de un grupo de factores a la brecha existente. Recordemos que cada estudiante está ligado a un vector (p, z, g) donde p es la calificación obtenida, z son las características observadas y g tipo de escuela. Supongamos que z se puede dividir en dos subvectores z_1, z_2 y que estamos interesados en el peso que tiene el grupo de fac-

tores que conforman a z_1 . En palabras de contrafactuales nos interesa saber ¿Cuál es la distribución de calificaciones de los estudiantes de escuelas públicas si tuvieran las características correspondientes a z_2 de los estudiantes de escuelas privadas?. Usando la regla de Bayes dos veces podemos expresar la función de densidad de los calificaciones de estudiantes de escuelas públicas como:

$$f(p|pub) = \int \int m(p|z_1, z_2, pub)h_1(z_1|z_2, pub)h_2(z_2|pub)dz_1dz_2. \quad (3.22)$$

Bajo la notación con los vectores z_1, z_2 podemos reescribir el contrafactual ¿Cuál sería la distribución de calificaciones si los estudiantes de educación pública tuvieran las características de los estudiantes de escuelas privadas?, $f_h^{pub \rightarrow priv}$, como:

$$f_h^{pub \rightarrow priv} = \int \int m(p|z_1, z_2, pub)h_1(z_1|z_2, priv)h_2(z_2|priv)dz_1dz_2. \quad (3.23)$$

Por tanto, podemos escribir el contrafactual de interés como:

$$\begin{aligned} f_{h,z_2}^{pub \rightarrow priv} &= \int \int m(p|z_1, z_2, pub)h_1(z_1|z_2, pub)h_2(z_2|priv)dz_1dz_2 \\ &= \int \int m(p|z_1, z_2, pub)h_1(z_1|z_2, pub)\frac{h_2(z_2|pub)}{h_2(z_2|pub)}h_2(z_2|priv)dz_1dz_2 \\ &= \int \int m(p|z_1, z_2, pub)h_1(z_1|z_2, pub)h_2(z_2|pub)\frac{h_2(z_2|pub)}{h_2(z_2|priv)}dz_1dz_2 \\ &= \int \int m(p|z_1, z_2, pub)h_1(z_1|z_2, pub)h_2(z_2|pub)\Phi(z_2)dz_1dz_2, \end{aligned}$$

donde $\Phi(z_2)$ se puede estimar de forma similar a (3.15). Por tanto, el estimador del contrafactual esta dado por:

$$\hat{f}_{m,h}^{pub \rightarrow priv} = \sum_{i=1}^n \frac{\hat{\Phi}(z_{2i})}{h} K\left(\frac{p - P_i}{h}\right). \quad (3.24)$$

La descomposición que se encuentra detrás de esta idea está resumida en la ecuación (3.25).

$$f_h^{pub \rightarrow priv} - f^{pub} = f_h^{pub \rightarrow priv} - f_{h,z_2}^{pub \rightarrow priv} + f_{h,z_2}^{pub \rightarrow priv} - f^{pub}, \quad (3.25)$$

la segunda parte sería la aportación de las características correspondientes a z_2 . Otra descomposición útil, para tener otra medida de la contribución de z_2 a la brecha existente, la tenemos con el contrafactual $f_{h,z_1}^{pub \rightarrow priv}$.

$$f_h^{pub \rightarrow priv} - f^{pub} = f_h^{pub \rightarrow priv} - f_{h,z_1}^{pub \rightarrow priv} + f_{h,z_1}^{pub \rightarrow priv} - f^{pub}, \quad (3.26)$$

donde la primera parte corresponde a la aportación de z_2 . Como menciona [Altonji, Bhargadwaj, & Lange \(2008\)](#) las dos medidas son, en general, distintas pero ambas dan una idea mas precisa de la aportación de z_2 .

Capítulo 4

Datos y Estadísticas Descriptivas

4.1. Bases de Datos

Se utilizaron bases de datos correspondientes al año 2008 de la Evaluación Nacional del Logro Académico en Centros Escolares (*ENLACE*). ENLACE pertenece a los pocas evaluaciones, como las pruebas EXCALE, de todo el Sistema Educativo Mexicano. La primera aplicación fue en el año 2006 en Educación Básica (primaria y secundaria) y posteriormente, en el año 2008, se aplicó para Educación Media Superior (preparatoria). ENLACE se realiza cada año durante el segundo trimestre y se aplica a jóvenes de tercero a sexto grado de primaria, primero a tercer grado de secundaria y estudiantes del último semestre de bachillerato. Analizar el sistema de Educación Media Superior es interesante pues precisamente los estudiantes evaluados se encuentran a un paso de la Educación Superior (Universidad). La presente investigación utilizará la información de los resultados en Habilidad Matemática de ENLACE 2008 para Educación Media Superior.

La prueba ENLACE para Educación Media Superior se compone por 140 reactivos

distribuidos en habilidad lectora con 40 y habilidad matemática con 90 reactivos, adicionalmente se incluye un cuestionario con preguntas relacionadas con el contexto social del estudiante y en los últimos años se ha incluido un anexo para evaluar una materia distinta adicional (historia, biología) . ENLACE no se considera un examen que mide los conocimientos adquiridos sino cómo el estudiante puede aplicar a situaciones de la vida cotidiana las habilidades adquiridas a lo largo de su educación. La prueba fue desarrollada en un marco de diagnóstico y no como un instrumentos de acreditación o de repercusión sobre el estudiante y esto hace de la prueba un instrumento de evaluación más fiable. Además, la publicación de los resultados y su fácil acceso representa un paso enorme en términos de transparencia. Los padres de familia, profesores y directores pueden observar los resultados vía internet a nivel individual y del centro escolar e inclusive participar en el proceso de evaluación como observadores. Los resultados totales también se encuentran disponibles previa solicitud al Instituto Federal de Acceso a la Información (IFAI).

Las bases de datos que conforman los resultados de la prueba ENLACE se componen por *Base Principal*, *Cuestionario de Contexto* y *Cuestionario de Directores*. Por un lado, la *Base Principal* contiene los resultados de los 800,000 estudiantes que presentaron la prueba ENLACE e incluye estadísticas descriptivas a nivel escuela (asistencia de estudiantes a la prueba, porcentaje de estudiantes en nivel deficiente, bueno y excelente, entre otras). El número de estudiante en la *Base Principal* representa el 87 % de la matrícula en tercer año de Educación Media Superior.¹ Además, cada estudiante recibe un folio único de acuerdo al estado, municipio, escuela a la que pertenece el estudiante y otras variables.

Por otro lado, en la base *Cuestionario de Contexto* se encuentran todos los estudiantes, de una muestra aleatoria de escuelas y representativa a nivel nacional, que respon-

¹ *Sistema Educativo Mexicano: Principales Cifras Ciclo Escolar 2008-2009*

dieron un anexo que incluye preguntas destinadas a obtener información a nivel individual, familiar y sobre infraestructura escolar. El *Cuestionario de Contexto* contiene 214,000 observaciones con menos del 0.5 % de estas duplicadas y comparte el número de folio con la *Base Principal*. Los directores de las escuelas participantes también responden un *Cuestionario de Directores* vía internet. Este cuestionario reúne información sobre la planta de docentes, infraestructura escolar, planes de estudios, programas curriculares y otros. Los *Cuestionarios de Contexto y Directores* comparten la clave de la escuela.

La base final, resultado de unir la *Base Principal*, *Cuestionario de Contexto* y *Cuestionario de Directores*, contiene aproximadamente 195,000 observaciones y más de 600 variables.² La base final fue normalizada para poder interpretar los resultados en términos de desviaciones estándar y seguir la literatura económica alrededor de la educación.

4.2. Estadísticas Descriptivas

Nuestro objetivo es analizar de forma más detallada y encontrar los determinantes de la diferencia en calificaciones entre escuelas privadas y públicas para el año 2008 en Educación Media Superior. La Tabla 4.1 muestra las diferencias promedio en calificaciones entre escuelas privadas y escuelas públicas, por materia evaluada y por base de datos. La Base Original es aquella que contiene al 87 % de los estudiantes en Educación Media Superior. En la Tabla 4.1 se observa que la construcción de la base final no generó cambios en las diferencias del promedio en puntajes respecto a la base original.

La diferencia entre escuelas privadas y públicas se encuentra alrededor de 0.2 desviaciones estándar tanto en habilidad lectora como en habilidad matemática y las pruebas estadísticas indican que la diferencia en puntajes sobre la media es diferente de cero

² *Detalles de la construcción Apéndice A*

y significativa a niveles menores al 1%.³ La escala es producto de un cálculo donde se consideran variables como dificultad de la pregunta, probabilidad de que el alumno responda al azar, cantidad de estudiantes que respondieron correctamente una pregunta y otras características. Dado que normalizamos los datos, una comparación más directa con la literatura en educación es factible. Por ejemplo, de acuerdo con [Woesmann \(2000\)](#), la diferencia entre escuelas privadas y públicas es mayor al efecto de reducir en dos desviaciones estándar el tamaño de los grupos sobre el desempeño.

Tabla 4.1: Diferencias en Calificaciones entre Escuelas Privadas y Públicas: Base Original y Base Construida.

	Escuela Privada	Escuela Pública	Diferencia
Base Original			
Habilidad Matemática	0.1618	-0.0346	0.1964
Habilidad Lectora	0.1872	-0.0400	0.2272
Base Construida			
Habilidad Matemática	.1561	-.0501	0.2062
Habilidad Lectora	.1852	-.0299	0.2151

Fuente: elaboración propia usando datos provenientes de los Cuestionario de Contexto de ENLACE 2008 en Educación Media Superior. Las estadísticas están ajustados por los pesos muestrales correspondientes.

La Tabla 4.2 y 4.3 contienen los puntajes promedios por estado ordenados de mayor a menor para habilidad matemática y habilidad lectora respectivamente. Se puede observar que Chiapas, Guerrero, Sonora y Tamaulipas se encuentran en las últimas posiciones en ambos listados y hay que señalar que Chiapas y Guerrero son estados que sufren de los índices más altos de marginación y pobreza. Por otro lado, el Distrito Fe-

³Veáse Apéndice

deral y Querétaro son estados con mayor actividad económica y se encuentran en el cima de ambas evaluaciones.

Las tablas también nos permiten hacer una comparación de la diferencia entre estudiantes de escuelas privadas y públicas. El puntaje promedio de escuelas públicas en habilidad matemática está ubicado entre los estados de Chihuahua y San Luis Potosí mientras que el puntaje de escuelas privadas es mayor al del estado de Baja California. Esto equivale a 17 posiciones en el ranking. En el caso de habilidad lectora la diferencia entre escuelas públicas y privadas es equivalente a la diferencia existente entre Yucatán (2ª posición en el ranking) y San Luis Potosí, es decir, 19 posiciones (Tabla 4.3). Las implicaciones anteriores señalan que la diferencia existente entre el sistema público y privado es sustancial y por tanto es necesario conocer qué la está causando.

Tabla 4.2: Calificaciones Medias por Estado en Habilidad Matemática

Estado	Puntaje	Estado	Puntaje	Estado	Puntaje
1. Querétaro	0.2999	11. Colima	0.0841	21. Coahuila	-0.0205
2. Aguascalientes	0.2753	12. Durango	0.0645	22. Sinaloa	-0.0491
3. Nuevo León	0.2052	13. Morelos	0.0640	23. Nayarit	-0.0657
4. Guanajuato	0.1937	14. Puebla	0.0508	24. Michoacán	-0.0759
5. Baja California	0.1403	15. Chihuahua	0.0506	25. Quintana Roo	-0.0853
6. Yucatán	0.1398	16. San Luis Potosí	0.0406	26. Oaxaca	-0.0880
7. Distrito Federal	0.1200	17. Tlaxcala	0.0388	27. Veracruz	-0.0929
8. Jalisco	0.1115	18. Campeche	0.0202	28. Tamaulipas	-0.1808
9. Hidalgo	0.1029	19. México	-0.0126	29. Sonora	-0.2053
10. Zacatecas	0.0860	20. Baja California Sur	-0.0146	30. Chiapas	-0.2898
				31. Guerrero	-0.3554

Fuente: elaboración propia usando datos de la *Base Original* de ENLACE 2008 para Educación Media Superior. Datos provenientes de la muestra total. El estado de Tabasco no participó en esta evaluación debido a desastres naturales.

Tabla 4.3: Calificaciones Medias por Estado en Habilidad Lectora

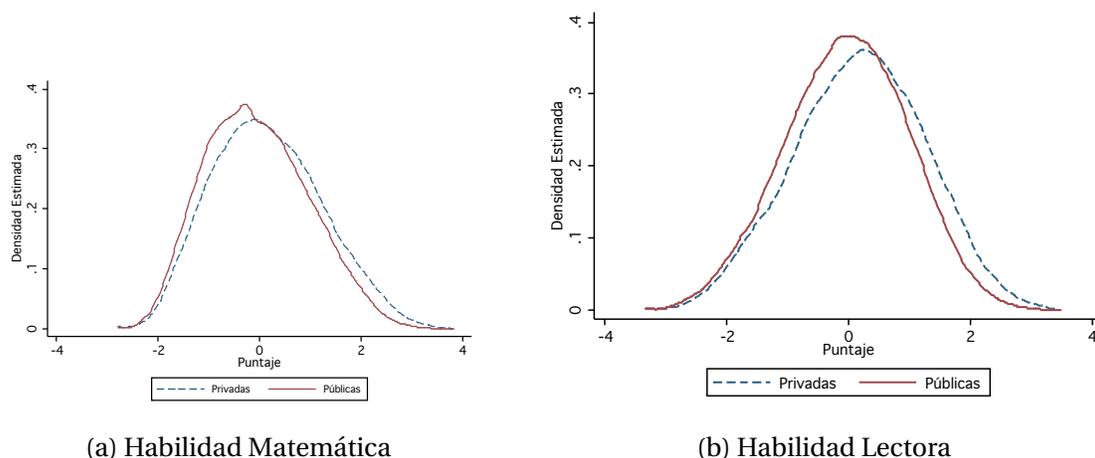
Estado	Puntaje	Estado	Puntaje	Estado	Puntaje
1. Querétaro	0.3568	11. Campeche	0.0937	21. Coahuila	-0.0819
2. Yucatán	0.1846	12. Morelos	0.0720	22. Durango	-0.0967
3. Distrito Federal	0.1601	13. Baja California	0.0551	23. Michoacán	-0.1003
4. Puebla	0.1393	14. Aguascalientes	0.0508	24. Chihuahua	-0.1012
5. Hidalgo	0.1381	15. Colima	0.0231	25. Nayarit	-0.1113
6. Guanajuato	0.1357	16. Veracruz	0.0162	26. Oaxaca	-0.1283
7. Jalisco	0.1258	17. Baja California Sur	0.0044	27. Sinaloa	-0.1445
8. Tlaxcala	0.1203	18. Zacatecas	-0.0074	28. Tamaulipas	-0.1818
9. México	0.1025	19. San Luis Potosí	-0.0343	29. Guerrero	-0.2906
10. Nuevo León	0.0959	20. Quintana Roo	-0.0747	30. Sonora	-0.3418
				31. Chiapas	-0.3826

Fuente: elaboración propia usando datos de la *Base Original* de ENLACE 2008 para Educación Media Superior. Datos provenientes de la muestra total. El estado de Tabasco no participó en esta evaluación debido a desastres naturales.

Una forma donde se puede observar con mayor detalle donde ocurre la desigualdad en calificaciones es haciendo uso de las funciones de densidad de las calificaciones. La Figura 4.1 muestra la estimación de la función de densidad de las calificaciones en habilidad matemática usando métodos de kernel. La estimación muestra que la función de densidad para las escuelas públicas se encuentra hacia la izquierda de la correspondiente a escuelas privadas, indicando una mayor media en calificaciones en los estudiantes de escuelas privadas. También se observa mayor concentración en la parte de bajas calificaciones, es decir, un estudiante de escuela privada tiene menores probabilidades de obtener una baja calificación que un estudiante de escuela pública.

Respecto a habilidad lectora, la Figura 4.1 muestra que las observaciones anteriores siguen siendo válidas. Sin embargo, ambas se encuentran un poco más a la derecha

Figura 4.1: Funciones de Densidad Originales por Sostenimiento



Fuente: elaboración propia utilizando datos de los cuestionarios de contexto de la prueba ENLACE 2008 para habilidad matemática. Grupo de comparación: escuelas privadas. Para estimar las funciones de densidad se usa el Kernel de Epanechnikov y el Bandwidth propuesto por Silvermann (1986). Las muestras para Habilidad Matemática y Habilidad Lectora contiene 194,230 observaciones.

del cero señalando que los alumnos obtuvieron mejores calificaciones en la parte de Habilidad Lectora.

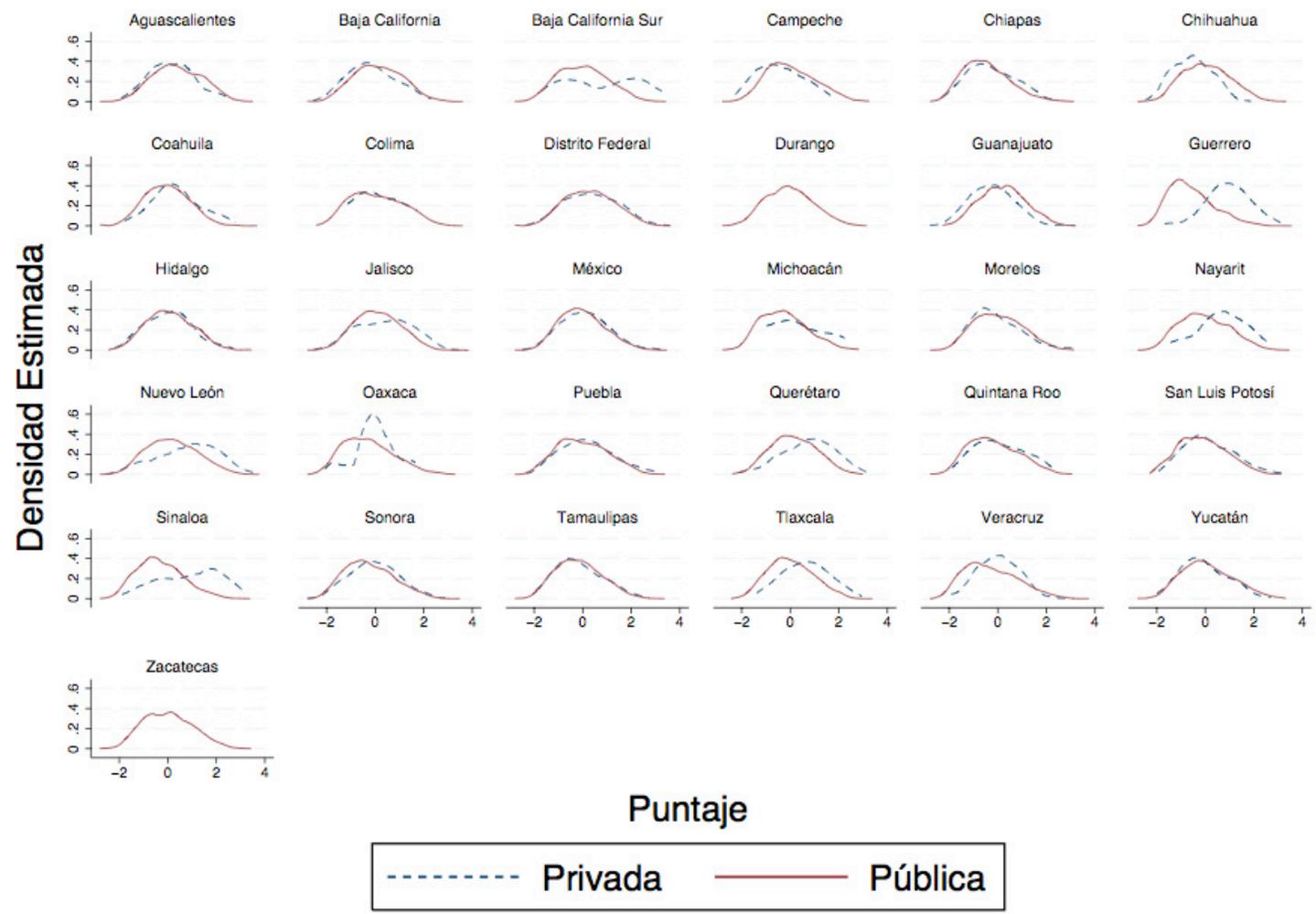
Para ilustrar el porqué usar las funciones de densidad para analizar la diferencia en calificaciones se presenta la Figura 4.2 donde se desglosa las calificaciones por tipo de subsistema y estados. Las estimaciones muestran que en estados como Hidalgo, Morelos, Yucatán y San Luis Potosí no se puede apreciar diferencia alguna entre las calificaciones de escuelas privadas y públicas. También muestran que existen algunos estados, como Chihuahua y Guanajuato, donde la función de densidad de las escuelas privadas se encuentra a la izquierda de la correspondiente a escuelas públicas, en otras palabras, las escuelas privadas tienen un peor desempeño en estos estados. Sin embargo, en Sinaloa, Veracruz, Querétaro y Nuevo León sucede lo contrario, la estimación de las escuelas privadas se localiza a la derecha de la estimación de las escuelas públicas y por tanto que

las escuelas privadas tienen un mejor desempeño en la prueba. También es importante mencionar que algunos estados como Zacatecas y Durango no tienen observaciones de estudiantes de escuelas privadas y por ello no existe una estimación.

La Figura 4.2 no proporciona alguna clave directa sobre qué está causando la diferencia en puntajes. Por ejemplo, por un lado Chiapas es uno de los estados más marginados del país mientras que el Distrito Federal aporta más del 20 % del Producto Interno Bruto Nacional⁴ y en ambos no se observan grandes diferencias entre las escuelas privadas y públicas. Por otro lado, Oaxaca y Nuevo León son estados con características económicas y sociales distintas y en ambos existen diferencias obvias entre escuelas privadas y públicas. Por tanto, un análisis más profundo y detallado es necesario.

⁴Instituto Nacional de Estadística y Geografía 2010

Figura 4.2: Funciones de Densidad de Puntajes por Estado y Sostentamiento



La literatura económica en educación presenta resultados ambiguos y contradictorios cuando se habla de los factores que afectan al desempeño de los estudiantes en pruebas estandarizadas. Estudios como Backhoff, Bouzas, Contreras, Hernández, & García (2007), De Hoyos, Espino, & García (2010) y Hanushek (2003) indican que las características socioculturales de los estudiantes tienen un mayor peso que los insumos escolares mientras estudios como Alvaréz, García-Moreno, & Patrinos (2007) y Graddy & Stevens (2005) apuntan en la dirección opuesta. La Tabla 4.4 contiene el puntaje promedio en habilidad matemática desagregado por los subgrupos que componen a algunas características individuales. En ella se puede observar que existe una relación positiva entre el promedio obtenido en educación primaria y el puntaje obtenido en la prueba, lo cual puede ser un indicador de que los primeros años de educación son los más importantes para el desarrollo cognoscitivo de las personas. Respecto a la variable sexo, la literatura (Sohn, 2010) sostiene que los hombres obtienen mejores puntajes en matemáticas y este caso no es la excepción. Inclusive las aspiraciones de los estudiantes influyen de forma positiva y esto puede ser consecuencia de una mayor motivación: estudiantes que anhelan alcanzar grados de estudios más altos están conscientes del esfuerzo que conlleva. Además, la relación de fumar y trabajar son las esperadas. Un resultado inesperado, y posiblemente debido a respuestas falsas, es que el hecho de beber ayuda a mejorar el puntaje.

Tabla 4.4: Media en Puntaje de Habilidad Matemática sobre Características Individuales

Variable	Media	Variable	Media
Sexo		Tipo de Escuela Secundaria	
Hombre	0.1418	Pública	-0.0835
Mujer	-0.1505	Privada	0.5353
Fumar		Aspiraciones de estudios	
No	-0.0060	Bachillerato	-0.7282
Si	-0.0474	Técnico Universitario	-0.4032
		Licenciatura	-0.2083
Beber		Posgrado	0.1916
No	-0.0790		
Si	0.0470	Horas dedicadas a Matemáticas	
		1	-0.5322
Promedio en Primaria		2	-0.4354
6.0-6.4	-0.7314	3	-0.0438
6.5-6.9	-0.7492	4	-0.0142
7.0-7.4	-0.6983	5	0.2036
7.5-7.9	-0.5356	6	0.1772
8.0-8.4	-0.3934	7	0.3846
8.5-8.9	-0.1200		
9.1-9.4	0.1986	Trabajos	
9.5-9.9	0.6032	No	0.0214
10	0.8746	Sí	-0.1290

Fuente: elaboración propia usando información de los Cuestionario de Contexto ENLACE 2008 para Educación Media Superior. Datos ajustados por pesos muestrales.

La Tabla 4.5 describe el puntaje desagregado respecto a algunas variables a nivel hogar. Por un lado, en ella se observa que tanto la educación de la madre como del padre exhibe una correlación positiva con el promedio de los puntajes y puede reflejar herencia de conocimientos de los padres hacia los hijos. En una primera instancia es posible afirmar que la educación de la madre tiene un efecto mayor que la educación del padre. El resultado anterior no es sorprendente pues usualmente la madre se encarga del cuidado de los hijos. Sin embargo, inusual es el hecho que un nivel de posgrado tiene un

promedio menor que estudios de licenciatura. Por otro lado, las variables piso, computadoras y libros en casa son variables relacionadas con la infraestructura del hogar y su relación con el puntaje es positiva.

Tabla 4.5: Diferencias en Puntaje sobre Características del Hogar

Variable	Media	Variable	Media
Educación de la Madre		Educación del Padre	
Sin estudios	-0.4329	Sin estudios	-0.3283
Primaria	-0.1993	Primaria	-0.2297
Secundaria	-0.0499	Secundaria	-0.1041
Bachillerato	0.0267	Bachillerato	0.0238
Carrera Técnica	0.2580	Carrera Técnica	0.1891
Licenciatura	0.4969	Licenciatura	0.4273
Posgrado	0.3307	Posgrado	0.3841
Piso del Hogar		Cuartos en el Hogar	
Tierra	-0.6326	1	-0.3797
Cemento	-0.1487	2	-0.1475
Mosaico	0.1708	3	-0.0071
		4	0.0679
		5	0.1546
Libros		6	0.1946
Ninguno	-0.4766	7	0.3579
1 a 10	-0.1945		
11 a 25	-0.0381		
26 a 50	0.0381	Computadoras en el Hogar	
51 a 100	0.1407	0	-0.2541
101 a 200	0.2363	1	0.0975
201 a 500	0.3811	2	0.3251
Más de 500	0.3619	3	0.5340
		4	0.7941

Fuente: elaboración propia usando información de los Cuestionario de Contexto ENLACE 2008 para Educación Media Superior. Datos ajustados por pesos muestrales.

La Tabla 4.6 muestra variables relacionadas con la infraestructura escolar, docentes y otras. Por un lado, la falta de electricidad, agua potable y un buen número de laboratorios presentan signos de un peor desempeño en la prueba. No se observa una correlación directa con el nivel de estudios de los docentes y a pesar de ello las escuelas donde existen profesores con posgrado obtienen las mejores calificaciones. Además, consideramos que la variable “*estado de las aulas*” es una buena aproximación para la infraestructura escolar total pues usualmente es prioridad en proyectos como el Programa de Escuelas de Calidad (PEC).

Tabla 4.6: Diferencias en Puntajes sobre Características de la Escuela

Variable	Media	Variable	Media
Número de Laboratorios		Nivel de Estudios de Docentes	
Sí son suficientes	0.1228	Bachillerato	0.3508
No son suficientes	-0.1226	Normal Básica	-0.4029
No hay	-0.3499	Normal Superior	0.0850
		Lic. en Pedagogía	-0.2703
		Licenciatura	-0.0410
		Posgrado	0.2545
Internet		Asociaciones de Estudiantes	
Sí	0.0317	Sí	0.0418
No	-0.4212	No	-0.0588
Director por Examen de Oposición		Vigilancia	
Sí	-0.0434	Sí	0.0570
No	-0.0066	No	-0.1888
Tutorías Académicas		Estados de las Aulas	
Sí	0.0398	No hay	-0.8219812
No	-0.1766	Malo	-0.2499278
		Regular	-0.168692
		Bueno	-0.0177502
		Excelente	0.3497249
Agua Potable			
Sí	0.0238		
No	-0.4964218		
Luz Eléctrica			
Sí	-0.0141404		
No	-0.5739765		

Fuente: elaboración propia usando información de los Cuestionario de Contexto ENLACE 2008 para Educación Media Superior. Datos ajustados por pesos muestrales.

La variable *libertad del profesor* en la Tabla 4.7 está ligada al estudio de [Alvaréz, García-Moreno, & Patrinos \(2007\)](#) donde encuentran que variables del tipo institucional afectan de forma positiva el desempeño. La existencia de huelgas e irrupción de porros podrían señalar problemas en el contexto social entre profesores y alumnos, respectivamente, y su relación con el puntaje es negativa. Las variables *relación entre estudiantes y docentes* están relacionadas con el ambiente social de la escuela. Además, que la escuela participe en pruebas estandarizadas no centralizadas como PISA no exhibe diferencias significativas y sustenta el proceso de aleatorización del examen.

Tabla 4.7: Diferencias en Puntajes sobre Características Institucionales

Variable	Media	Variable	Media
Irrupción de Porros		Nivel de Exigencia	
Sí	0.0356	Nada Exigente	-0.4318
No	-0.0145	Poco Exigente	-0.3370
		Exigente	-0.0476
Huelgas		Muy Exigente	0.3196
Sí	-0.1332		
No	-0.0099	Estudios de Directores	
		Bachillerato	-0.2068
Relación entre estudiantes y docentes		Normal Básica	0.0458
Regular	-0.2479	Normal Superior	-0.0886
Buena	-0.0141	Licenciatura en	-0.0846
Excelente	0.0489	Licenciatura	-0.0611
		Especialidad	0.0851
PISA		Maestría	0.0785
Sí	-0.0258	Doctorado	-0.2568
No	-0.0075		

Fuente: elaboración propia usando información de los Cuestionario de Contexto ENLACE 2008 para Educación Media Superior. Datos ajustados por pesos muestrales.

4.3. Diferencia en Características Observables.

La hipótesis implícita a la metodología desarrollada por DiNardo, Fortin, & Lemieux (1996) es que las características observables que componen el vector Z son distintas entre los grupos que se pretenden analizar. La Tabla 4.8 resume las diferencias en proporciones de algunas características observables entre estudiantes de escuelas públicas y privadas, y se obtienen resultados interesantes. Primero, características como *Sexo* y *Trabaja* no presentan diferencias entre estudiantes de escuelas públicas y privadas. Segundo, a nivel hogar la diferencia en variables como *educación de la madre* es importante y con ventaja para las escuelas privadas, es decir, existe una mayor proporción de estudiantes cuyas madres tienen estudios por arriba de la Educación Media Superior. Tercero, se observa que estudiantes de escuelas privadas tienen mayores aspiraciones que los estudiantes de escuelas privadas y probablemente significaría que los estudiantes de escuelas privadas se encuentran más motivados. Cuarto, estudiantes de escuelas privadas tienen ventaja en acceso a internet y número de laboratorios. Por último, los estudiantes de escuelas privadas tienden a fumar más y con ello repercutir en su desempeño escolar. Dadas las observaciones anteriores se considera que la explicación de la brecha, vía diferencias en las características, es totalmente factible.

Tabla 4.8: Diferencias en Proporciones sobre Características Observables

Variable	Pública	Privada	Variable	Pública	Privada
Aspiraciones de Estudios			Educación de la Madre		
Bachillerato	4.28	1.97	No estudió	6.53	2.24
Técnico Universitario	11.25	5.83	Primaria	34.11	15.66
Licenciatura	28.28	22.35	Secundaria	28.5	20.33
Posgrado	56.19	69.85	Nivel Medio Superior	12.21	18.11
			Carrera Técnica	9.56	16.61
			Licenciatura	7.18	20.46
Sexo			Posgrado	1.91	6.58
Hombre	46.61	46.36			
Mujer	53.39	53.64			
			Número de Laboratorios		
Trabajas			Sí son suficientes	52.54	80.42
Sí	23.61	23.38	Suficientes	33.43	8.86
No	76.39	76.62	No hay	14.04	10.72
Fumas			Servicio de Internet		
Sí	19.16	29.2	Sí	89.09	96.09
No	80.84	70.8	No	10.91	3.91
Relación Estudiantes, Docentes			Nivel de Exigencia		
Regular	4.29	0.54	Nada Exigente	0.01	0
Buena	84.97	66.34	Poco Exigente	8.98	3.48
Excelente	10.74	33.12	Exigente	78.58	66.1
			Muy Exigente	12.43	30.42
PISA					
Sí	22.22	18.88	Asociaciones de Estudiantes		
No	77.78	81.12	Sí	46.67	44.38
			No	53.33	55.62

Fuente: elaboración propia usando información de los Cuestionario de Contexto ENLACE 2008 para Educación Media Superior. Datos ajustados por pesos muestrales.

4.4. Especificación Econométrica

La metodología desarrollado por DiNardo, Fortin, & Lemieux (1996) es una combinación de métodos no paramétricos y paramétricos y uno de sus beneficios es no asumir una forma reducida para la función de producción. Lo anterior imposibilita dar una es-

pecificación econométrica. La Tabla 4.9 resume todas las variables utilizadas en el vector Z_i divididas en características individuales, familiares, escolares e institucionales. La distinción entre variables relacionadas con el estudiante y aquellas relacionadas con la escuela es bastante clara. Sin embargo, hacer una diferencia entre variables institucionales y escolares es más complicado. Es este sentido, seguimos la distinción de autores como Backhoff, Bouzas, Contreras, Hernández, & García (2007); De Hoyos, Espino, & García (2010); Fuchs & Woessmann (2004).

Tabla 4.9: Variables Incluidas en el Modelo Logit

Variables Individuales	Variables Escolares
Sexo	Educación de los Docentes
Trabaja	Número de Laboratorios
Tipo de Primaria	Internet
Tipo de Secundaria	Examen de Oposición
Fumar	Vigilancia
Bebidas Alcoholicas	
Aspiraciones del Estudiante	
Promedio Primaria	Variables Institucionales
Horas de Clase en Matemáticas	Asociaciones estudiantiles
Promedio del último año de secundaria	Participación en PISA
	Irrupción de Porros
Variables Familiares	Huelgas
Piso del Hogar	Relación Estudiantes Profesores
Estudios de la Madre	Exigencia
Estudios del Padre	Tutorías
Número de Libros	Estudios Director
Número de Cuartos	
Computadoras	

Fuente: elaboración propia usando información de los Cuestionario de Contexto ENLACE 2008 para Educación Media Superior.

Capítulo 5

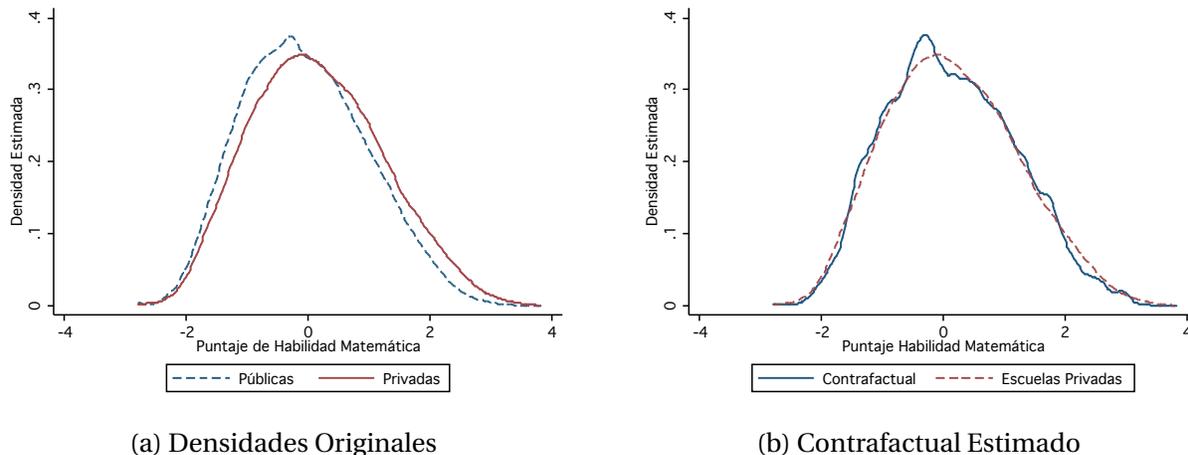
Resultados

5.1. Di Nardo, Fortin & Lemieux

La Figura 5.1 contiene la estimación del contrafactual en habilidad matemática usando la metodología original de DiNardo, Fortin, & Lemieux (1996). El contrafactual estimado es: ¿cuál es la distribución de calificaciones de los estudiantes escuelas públicas, si tuvieran la distribución de características de los estudiantes de escuelas privadas y la distribución de puntajes de las escuelas públicas? y está descrito en la Ecuación 3.9. De la figura se pueden enunciar dos observaciones. La primera es que la función de densidad estimada se encuentra a la derecha de la función de densidad de calificaciones de escuelas públicas, indicando que los estudiantes de escuelas públicas mejorarían su desempeño si tuvieran las características de los estudiantes de escuelas privadas. El segundo es que la estimación replica casi en su totalidad a la función de densidad de escuelas privadas, es decir, las características observables estarían explicando gran parte de la brecha existente. Más adelante resumiremos cómo cambian las magnitudes de diversos estadísticos y qué tanto se puede explicar de la brecha existente vía característi-

cas observables.

Figura 5.1: Función de Densidad Estimada para el Contrafactual.



Fuente: elaboración propia utilizando datos de los cuestionarios de contexto de la prueba ENLACE 2008 para Habilidad Matemática. El contrafactual estimado es: ¿cómo se distribuyen las calificaciones de los estudiantes de escuelas públicas si tuvieran la distribución de características de escuelas privadas?. Grupo de comparación: escuelas privadas. Las cuatro funciones de densidad estimadas usan el Kernel de Epanechnikov y el bandwidth propuesto por Silvermann (1986). Las muestras para las funciones de densidad originales y el contrafactual contienen 194,230 y 162,423, respectivamente.

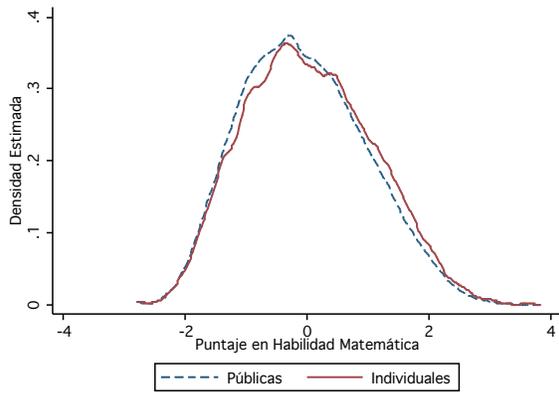
5.2. Descomposición Detallada

En la Figura 5.2 se encuentran las estimaciones resultantes de la metodología propuesta por Altonji, Bharadwaj, & Lange (2008) descrita en la Ecuación 3.18. Las figuras muestran la evolución de los contrafactuales cuando se cambia la distribución de los cuatro grupos de variables. Por ejemplo, la Figura 5.2a contiene dos funciones de densidad estimada. La primera estimación, marcada por la etiqueta “Públicas”, corresponde a la distribución original de calificaciones de estudiantes de escuelas públicas. La segunda estimación (*Individual*) es la distribución de calificaciones de estudiantes de

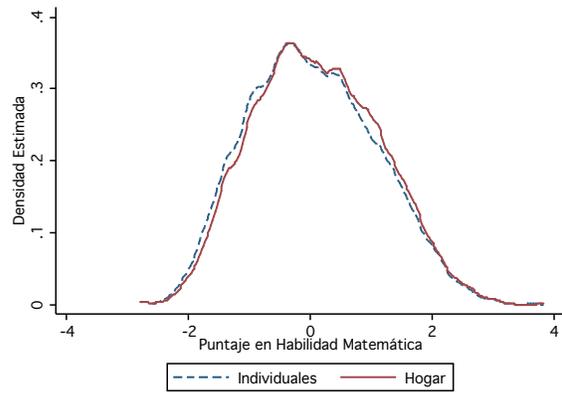
escuelas públicas si solamente tuvieran la distribución de las características a nivel individual de estudiantes de escuelas privadas. En la Figura 5.2b, la primera estimación “*Individuales*” es la segunda estimación de la Figura 5.2a. La segunda estimación (*Hogar*) corresponde a la distribución de calificaciones de los estudiantes de escuelas públicas, si tuvieran la distribución de características a nivel individual y hogar de los estudiantes de escuelas privadas y así sucesivamente.

Los efectos de las características a nivel individual y hogar son claros. Cambiar estas características a estudiantes de escuelas públicas mueve la función de densidad hacia la derecha e implica una mejora en los resultados de la prueba. Al incluir las variables escolares la función de densidad no presenta cambios importantes y por tanto, el efecto de estas variables es casi nulo. Respecto a las variables institucionales, se observa un ligero movimiento de la función de densidad estimada hacia la izquierda al momento de incluirlas, implicando que este tipo de variables pueden afectar de forma negativa al desempeño de cierto tipo de estudiantes. Las gráficas son ilustrativas para observar el efecto general de los grupos de variables sobre la brecha en calificaciones. Sin embargo, parte de nuestro objetivo es aprovechar la metodología para explorar qué está pasando en diversos puntos de la distribución.

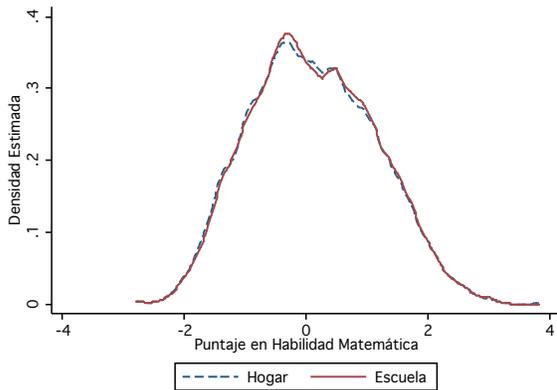
Figura 5.2: Descomposición Detallada Estimada



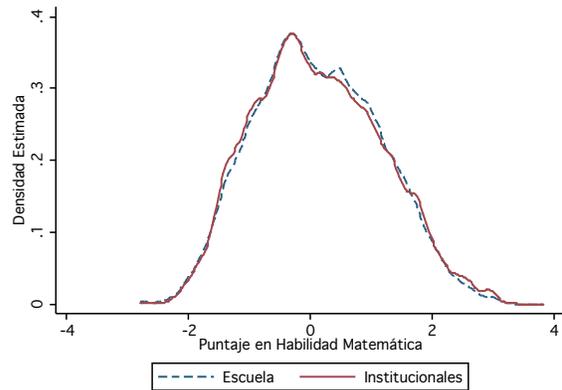
(a) Contrafactual a Nivel Individual



(b) Contrafactual a Nivel Hogar



(c) Contrafactual a Nivel Escuela



(d) Contrafactual a Nivel Institucional

Fuente: elaboración propia utilizando datos de los cuestionarios de contexto de la prueba ENLACE 2008 para habilidad matemática. El contrafactual estimado es: ¿cómo se distribuyen las calificaciones de los estudiantes de escuelas públicas si tuvieran la distribución de características hasta el nivel “ x ” de estudiantes de escuelas privadas?, donde x es individual, hogar, escolares e institucionales. Las funciones de densidad estimadas usan el Kernel de Epanechnikov y el bandwidth propuesto por Silvermann (1986)

Tabla 5.1: Estadísticos de la Descomposición Detallada Original

	Públicas	Privadas	Diferencia	Estudiantes	Hogar	Escuela	Institucionales	Residual
P10	-1.3344 [0.006]	-1.1998 [0.015]	0.1346 [0.017] [100 %]	0.0223 [0.020] [17 %]	0.0838 [0.020] [62 %]	0.0243 [0.023] [18 %]	-0.0116 [0.028] [-9 %]	0.0158 [0.028] [12 %]
P25	-0.8135 [0.005]	-0.6420 [0.016]	0.1714 [0.017] [100 %]	0.0679 [0.015] [40 %]	0.0937 [0.020] [55 %]	0.0217 [0.021] [13 %]	-0.0202 [0.026] [-12 %]	0.0084 [0.026] [5 %]
P50	-0.1195 [0.005]	0.0933 [0.019]	0.2128 [0.020] [100 %]	0.1003 [0.017] [47 %]	0.0820 [0.022] [39 %]	0.0078 [0.022] [4 %]	-0.0101 [0.030] [-5 %]	0.0328 [0.028] [15 %]
P75	0.6520 [0.005]	0.8880 [0.019]	0.2360 [0.019] [100 %]	0.1311 [0.019] [56 %]	0.0750 [0.026] [32 %]	0.0117 [0.026] [5 %]	0.0103 [0.032] [4 %]	0.0079 [0.029] [3 %]
P90	1.3405 [0.005]	1.6222 [0.021]	0.2817 [0.022] [100 %]	0.1229 [0.020] [44 %]	0.0404 [0.028] [14 %]	0.0066 [0.030] [2 %]	0.0565 [0.036] [20 %]	0.0553 [0.035] [20 %]
Media	-0.0501 [0.004]	0.1561 [0.012]	0.2062 [0.013] [100 %]	0.0892 [.014] [43 %]	0.0739 [0.019] [36 %]	0.0134 [0.020] [7 %]	0.0057 [0.022] [3 %]	0.0240 [0.018] [12 %]

Fuente: elaboración propia usando base de datos del diseño muestral del cuestionario de contexto 2008. Todos los resultados son ajustados por los pesos de la muestra. Errores estándar en corchetes y fueron calculados con el método no paramétrico Bootstrap con 250 repeticiones.

La Tabla 5.1 presenta la aportación de los cuatro grupos de variables sobre la brecha existente en diversos puntos de la distribución. La parte de *Públicas* y *Privadas* corresponde a las distribuciones originales de los puntajes. En *Diferencia* resumimos la brecha original sobre distintos puntos de las distribuciones originales. Por ejemplo, en el percentil 10 la brecha existente entre escuelas públicas y privadas es de 0.1346 desviaciones estándar, favoreciendo a los estudiantes de escuelas privadas. Las siguientes columnas contienen la contribución de cada grupo usando los contrafactuales estimados en la descomposición propuesta por [Altonji, Bharadwaj, & Lange \(2008\)](#). En *Escuela* el con-

trafactual estimado es: ¿cuál es la distribución de calificaciones de escuelas públicas si tuvieran la distribución de calificaciones de características a nivel individual, hogar y escolar de los estudiantes de escuelas privadas?. Para obtener una medida de la contribución de las variables a nivel escolar se compara con el contrafactual *Hogar*, que cambia la distribución de las variables hogar e individual, y se divide entre la brecha original. El *residual* corresponde a la parte no explicada por el modelo.

De la Tabla 5.1 podemos hacer varias observaciones. Primero, la brecha entre escuelas privadas y públicas no es un fenómeno único de la media y se presenta tanto en puntos bajos como altos de la distribución. Además, la brecha observada en otros percentiles (Columna 3) se incrementa conforme avanzamos a puntos más altos de la distribución, es decir, si comparamos a los mejores estudiantes de escuelas privadas y escuelas públicas (percentil 90), la diferencia en calificaciones es mayor en aproximadamente un 40 % respecto a la diferencia original en medias (Tabla 4.1). Segundo, la explicación de las diferencias, vía las características observables, es exitosa pues logra explicar en promedio un 85 % de la brecha en calificaciones entre escuela públicas y privadas, encontrando su punto más bajo en el percentil 50 con un residual del 15 % y su punto más alto en el percentil 75 con un 3 % de la brecha no explicada. Tercero, de acuerdo a la Ecuación 3.3, la diferencia en medias o en cualquier percentil puede descomponerse en una parte consecuencia de las diferencias en características observables y en una segunda parte correspondiente al verdadero beneficio de estar en un grupo, en este caso, el hecho de estudiar en un colegio privado. En la Tabla 5.1, la columna *residual* resulta ser el verdadero beneficio de estar en una escuela privada y la suma de las columnas *Individual*, *Hogar*, *Escuela e Institucionales* es la parte explicada por las diferencias en características observables. En la Tabla 5.1 se observa que los verdaderos beneficios de estar en una escuela privada son más importantes para alumnos de alto desempeño (20 %). Sin embargo, también se observa que los beneficios son importantes para estudiantes ubicados en la parte media y baja de la distribución (12 %-15 %). En

este sentido, se puede afirmar que se obtienen beneficios del hecho de estudiar en una escuela privada aunque estos sólo representen una fracción de la diferencia en calificaciones promedio original.

Tercero, los resultados obtenidos para estudiantes ubicados en la parte media de la distribución se encuentran en línea con los reportados por la literatura, en el sentido de que los factores escolares tienen poco peso sobre el desempeño (Hanushek, 2003). Sobre la media, las características a nivel estudiante y hogar explican casi 80 % de la diferencia y sólo un 10 % de la brecha se puede explicar con factores relacionados con la escuela. Al analizar otros puntos de la distribución se obtienen resultados interesantes. Respecto a estudiantes de bajo rendimiento (percentil 10), las características individuales no tienen tanto peso (17 %) como en otro tipo de estudiantes con mayor rendimiento (40 % - 56 %). En cambio, las características a nivel hogar son las que contribuyen de forma importante (62 %) en este tipo de estudiantes. En total, para estudiantes de bajo rendimiento las características individuales y familiares explican casi un 80 % de la brecha existente. Este resultado podría reflejar un contexto familiar complicado o que a nivel individual los estudiantes de escuelas públicas y privadas son similares (percentil 10). La enorme contribución de factores individuales y familiares en la diferencia en calificaciones decrece conforme avanzamos en la distribución (Columnas 2 y 3). Por ejemplo, para estudiantes de alto rendimiento las características individuales y familiares explican sólo el 58 % de la diferencia en puntajes. Una posible explicación es que los estudiantes de escuelas privadas y públicas son más parecidos conforme se avanza en la distribución, en otras palabras, los estudiantes de alto rendimiento están inmersos en condiciones familiares más favorables.

Respecto a los factores escolares, por un lado, la mayor contribución se encuentra en estudiantes de bajo rendimiento (13 %-18 %) y la contribución de estos se reduce en estudiantes localizados en la parte alta de la distribución (3 % - 5 %). El resultado an-

terior apunta a que los retornos de los factores escolares están variando por el tipo de estudiante de una forma importante y resalta la importancia de analizar el efecto de los factores escolares en otros puntos de la distribución distintos a la media. Sin embargo, la gran contribución de los factores escolares en estudiantes de bajo desempeño se reduce por la aportación negativa de los factores institucionales (-9 %). Para ilustrar el porqué de la contribución negativa se presenta la Tabla 5.2 que contiene las proporciones respecto al tipo de sostenimiento de los factores institucionales para estudiantes que se encuentran por debajo del percentil 10.

De la Tabla 5.2 tres observaciones son importantes: 1) una ligera mayor proporción de estudiantes de escuelas públicas reciben tutorías y dado que estamos considerando a los estudiantes de bajo rendimiento existe un claro impacto de tutorías sobre su desempeño. 2) La variable *asociaciones estudiantiles* refleja el ambiente social entre los compañeros y existe una clara ventaja para estudiantes de escuelas públicas. 3) Los estudiantes de bajo rendimiento de escuelas públicas se someten en mayor proporción a pruebas estandarizadas como el examen PISA. De acuerdo con [Woesmann \(2000\)](#), el hecho de someterse a evaluaciones externas influye de forma positiva sobre el desempeño de los estudiantes. Por tanto, se considera que los estudiantes en el percentil 10 de escuelas públicas están inmersos en un mejor clima académico. Por tanto, los estudiantes de escuelas privadas están en desventaja en comparación con estudiantes de escuelas públicas en cuanto a factores institucionales se refiere y por ello, se considera que al cambiar las características institucionales de los estudiantes de escuelas públicas por las características institucionales de los estudiantes de escuelas privadas, aunado a la descomposición propuesta, la contribución es negativa. En la siguiente sección se proponen dos enfoques para resolver la dependencia en el orden de la descomposición y obtener estimadores de la contribución en cada grupo.

Tabla 5.2: Diferencia en Factores Institucionales en el Percentil 10

	Privadas	Públicas		Privadas	Públicas
Asociaciones Estudiantiles	33.69 %	47.21 %	Estudios del Director		
PISA	13.18 %	23.57 %	Bachillerato	1.06 %	0.12 %
Tutorías	69.58 %	71.07 %	Normal Básica	0.90 %	0.32 %
Paros	1.50 %	2.61 %	Normal Superior	2.01 %	1.80 %
Irrupción de Porros	0.14 %	4.03 %	Pedagogo	4.39 %	4.64 %
			Licenciatura	47.30 %	57.84 %
Exigencia de la Institución			Especialidad	5.93 %	2.11 %
Nada Exigente	0.00 %	0.01 %	Maestría	35.88 %	30.01 %
Poco Exigente	4.24 %	14.10 %	Doctorado	2.53 %	3.15 %
Exigente	79.25 %	77.43 %			
Muy Exigente	16.51 %	8.47 %	Relaciones Estudiantes-Profesores		
			Regular	0.46 %	5.47 %
			Buenas	75.48 %	83.65 %
			Excelentes	24.06 %	10.88 %

Fuente: elaboración propia usando base de datos del diseño muestral de los Cuestionarios de Contexto 2008. Todos los resultados son ajustados por los pesos de la muestra.

En cambio, para estudiantes de alto rendimiento los factores institucionales son muy importantes (20 %) pues contribuyen más que los factores familiares y escolares juntos (16%). Respecto a estudiantes en la parte media de la distribución la aportación en valor absoluta es pequeña (4 %-5 %). De acuerdo a los resultados anteriores, los factores institucionales juegan un rol importante en estudiantes que se encuentran en las colas de la distribución.

En resumen, los factores individuales y del hogar contribuyen a explicar buena parte de la brecha pero la proporción varía conforme al rendimiento de los estudiantes. Además, por un lado, variables institucionales importan mucho más en estudiantes con mejor desempeño y afectan de forma negativa a los estudiantes en la parte baja de la distribución. Por otro lado, la aportación de la diferencia en factores escolares es mayor

para estudiantes de bajo rendimiento y decrece conforme avanzamos en la distribución. Todos los resultados anteriores muestran que explorar otros puntos de la distribución aporta información relevante en temas de educación y proporcionan una mejor perspectiva de cómo se encuentra conformado el Sistema Educativo Mexicano.

5.3. Descomposición Independiente

Uno de las limitaciones de la metodología propuesta por [Altonji, Bharadwaj, & Lange \(2008\)](#) es que es dependiente del orden que elegimos. En su artículo sugieren que la elección del orden de la descomposición se debe elegir con base en argumentos económicos y la descomposición de la sección anterior se eligió siguiendo la literatura de los modelos jerárquicos en educación ([Backhoff, Bouzas, Contreras, Hernández, & García, 2007](#)). El orden propuesto es: características individuales, características del hogar, factores escolares y factores institucionales. La propuesta clásica, para resolver el problema de dependencia en el orden, es tomar el orden inverso de la descomposición elegida. En este caso, la descomposición inversa sería: Institucionales, Escuela, Hogar y Estudiantes. Sin embargo, existe un problema con la descomposición: el impacto de las variables a nivel hogar sobre las características individuales impide cambiar primero la distribución de características del hogar y después las características individuales. Por ello, el presente trabajo propone dos enfoques para resolver la dependencia de la descomposición.

En la primera etapa, el primer enfoque estima todas las descomposiciones posibles donde las características individuales se cambian antes que las características del hogar. Existen 12 descomposiciones válidas bajo este enfoque.¹ El segundo enfoque considera

¹(Escuela, Individual, Hogar, Institucional) es un ejemplo de descomposición válida mientras que (Hogar, Individual, Institucional, Escuela) es una descomposición no válida.

las características individuales y del hogar como un sólo grupo, reduciendo los grupos de variables a tres (Individuales, Escuela e Institucionales) y estimando las seis combinaciones posibles para la descomposición. En una segunda etapa, se calcula el promedio de las contribuciones en cada descomposición y el resultado provee una mejor idea de la magnitud de la contribución independiente del orden.

La Tabla 5.3 resume las contribuciones de cada grupo sobre la brecha de calificaciones entre estudiantes de escuelas privadas y públicas bajo el primer enfoque. Si comparamos estos resultados con los obtenidos en la descomposición original (Individual, Hogar, Escuela, Institucional) la mayoría de los resultados se mantienen con la diferencia de que los factores escolares tienen un peso mayor. El primero es que la contribución de los factores escolares y características del hogar decrece respecto al desempeño de los estudiantes, para estudiantes de bajo rendimiento (percentil 10) la contribución de los factores escolares y características del hogar es de 41.50 % y 33.80 %, respectivamente, mientras que para estudiantes ubicados en la parte alta de la distribución (percentil 90) el efecto se reduce a 22 % y 14.81 %. El segundo está relacionado con la magnitud de la contribución de las características *individuales* de estudiantes de bajo rendimiento (percentil 10) en comparación con otro tipo de estudiantes. En la Tabla 5.1 las diferencias del percentil 10 con el percentil 25, 50, 75 y 90 son 23 %, 30 %, 39 % y 27 %, respectivamente.² En la Tabla 5.3 las diferencias son 20 %, 28 %, 37 % y 30 %, respectivamente. Por tanto, a pesar de que la contribución de características individuales en el percentil 10 cae a 0.02 % los demás percentiles se reducen en una magnitud similar.

²Las diferencias se estimaron como IndividualesP25-IndividualesP10 y así sucesivamente.

Tabla 5.3: Estadísticos de la Descomposición
Detallada Bajo el Primer Enfoque

	Individuales	Hogar	Escuela	Institucionales
10	0.02 %	33.80 %	41.50 %	12.95 %
25	20.10 %	42.37 %	34.16 %	-0.66 %
50	28.24 %	31.32 %	23.86 %	2.36 %
75	37.22 %	29.75 %	26.04 %	5.20 %
90	30.63 %	14.81 %	22.37 %	13.60 %
Media	26.94 %	28.26 %	26.14 %	8.02 %

Fuente: elaboración propia usando base de datos del diseño muestral de los Cuestionarios de Contexto 2008. Todos los resultados son ajustados por los pesos de la muestra.

El tercer resultado es la contribución de los factores institucionales en estudiantes ubicados en la cola de la distribución (12.95 %-13.60 %), aunque, el signo negativo en estudiantes de bajo desempeño obtenido en la Tabla 5.1 se pierde. El signo negativo de la contribución en estudiantes de bajo rendimiento reflejaba la combinación de un clima académico más complicado en escuelas privadas y el orden propuesto y, precisamente, la idea de estimar todos los posibles ordenamientos es eliminar esa dependencia en el orden, resultando en una contribución positiva para los factores institucionales. Respecto a los percentiles de la parte media de la distribución, se observa que las características institucionales no son relevantes.

La Tabla 5.4 contiene los resultados bajo el segundo enfoque. La Tabla 5.5 presenta los resultados de la Tabla 5.3 si se tomará a las características del estudiantes y del hogar como una sola. Las tablas no presentan diferencias importantes y por tanto, las observaciones de los párrafos anteriores, como el decaimiento de la contribución de los factores escolares respecto al desempeño, se mantienen. Es de señalar que la contribución de las características escolares, en la Tabla 5.4, a la brecha existente es ligeramente

mayor que los resultados obtenidos en las tablas anteriores.

Tabla 5.4: Estadísticos de la Descomposición Detallada Segundo Enfoque

	Individuales	Escuela	Institucionales
P10	34.71 %	39.28 %	14.28 %
P25	56.54 %	35.20 %	3.37 %
P50	53.11 %	25.99 %	5.48 %
P75	59.31 %	27.17 %	10.18 %
P90	40.26 %	23.49 %	16.61 %
Media	49.79 %	27.64 %	10.92 %

Fuente: elaboración propia usando base de datos del diseño muestral de los Cuestionarios de Contexto 2008. Todos los resultados son ajustados por los pesos de la muestra.

Tabla 5.5: Estadísticos de la Descomposición Detallada Agrupando Resultados de la Tabla 5.3

	Individuales	Escuela	Institucionales
P10	33.83 %	41.50 %	12.95 %
P25	62.47 %	34.16 %	-0.66 %
P50	59.55 %	23.86 %	2.36 %
P75	66.97 %	26.04 %	5.20 %
P90	45.44 %	22.37 %	13.60 %
Media	55.20 %	26.14 %	8.02 %

Fuente: elaboración propia usando base de datos del diseño muestral de los Cuestionarios de Contexto 2008. Todos los resultados son ajustados por los pesos de la muestra.

5.4. Diferencias en Retornos: McCrary *et al.* (2010)

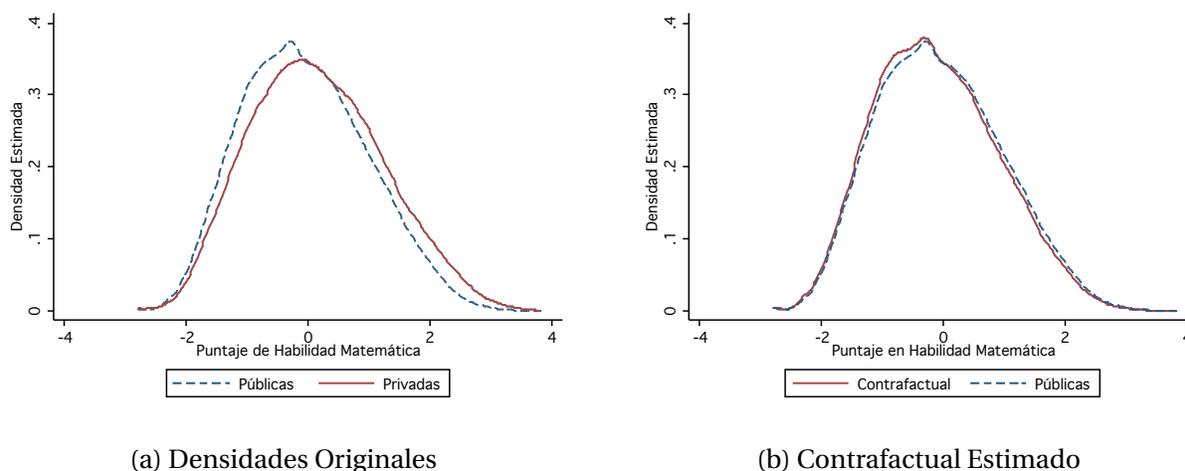
En la sección 3.4 se discute una segunda posible explicación para la brecha existente entre las escuelas privadas y públicas. De acuerdo con Leibbrandt, Levinsohn, & McCrary (2010) diferentes retornos sobre las características observables podría causar la diferencia en calificaciones entre los grupos. Por ejemplo, se puede pensar que las escuelas privadas, al estar sujetas a competencia, serían más eficientes al utilizar los recursos escolares. En términos matemáticos estamos interesados en comprobar si $m = l$, donde m, l son las funciones de producción educativa para escuelas públicas y privadas,

respectivamente. Es decir:

$$P_{pub} = m(Z_{pub}, \xi_{pub}) \quad P_{priv} = l(Z_{priv}, \xi_{priv}).$$

La Figura 5.3 resume los resultados obtenidos. La Figura 5.3a muestra la distribución original de puntajes tanto de escuelas públicas como de escuelas privadas. La Figura 5.3b contiene la estimación del contrafactual: ¿cuál sería la distribución de calificaciones de escuelas públicas si tuvieran los retornos de las características observables de las escuelas privadas?. El contrafactual se encuentra descrito en la Ecuación 3.18, y se compara con la distribución original de calificaciones de escuelas públicas. De acuerdo con la Figura 5.3, se observa que el contrafactual de escuelas públicas no presenta cambios importantes cuando le imponemos los retornos de las escuelas privadas. Esto sugiere que la hipótesis de distintas formas de producción no es suficientemente válida para nuestro caso y apoya la explicación de la diferencia en calificaciones vía diferencia en características observables. Una implicación directa de este resultado es que el sistema de educación privada y pública transforman los insumos en formas muy similares y el hecho de estar sujetos a competencia no repercute, de forma importante, en las calificaciones.

Figura 5.3: Función de Densidad Estimada.



Fuente: elaboración propia utilizando datos de los cuestionarios de contexto de la prueba ENLACE 2008 para habilidad matemática. El contrafactual estimado es: ¿cómo se distribuyen las calificaciones de los estudiantes de escuelas públicas si tuvieran la distribución de puntajes de los estudiantes de escuelas privadas?. Grupo de comparación: escuelas públicas. Las funciones de densidad estimadas usan el Kernel de Epanechnikov y el bandwidth propuesto por Silvermann (1986)

En la Figura 5.3b también se puede observar que el contrafactual se encuentra ligeramente ubicado a la izquierda de la distribución de escuelas públicas. El contrafactual estimado tiene una media de -0.1072 mientras que la distribución de escuelas públicas tiene una media de -0.0501 corroborando la última afirmación. Esto indica que, a pesar de no presentar cambios importantes, los estudiantes de escuelas públicas son un poco más eficientes al transformar los insumos disponibles. El resultado concuerda con las observaciones hechas por Duncan & Sandy (2007).

5.5. El Impacto del Pasado

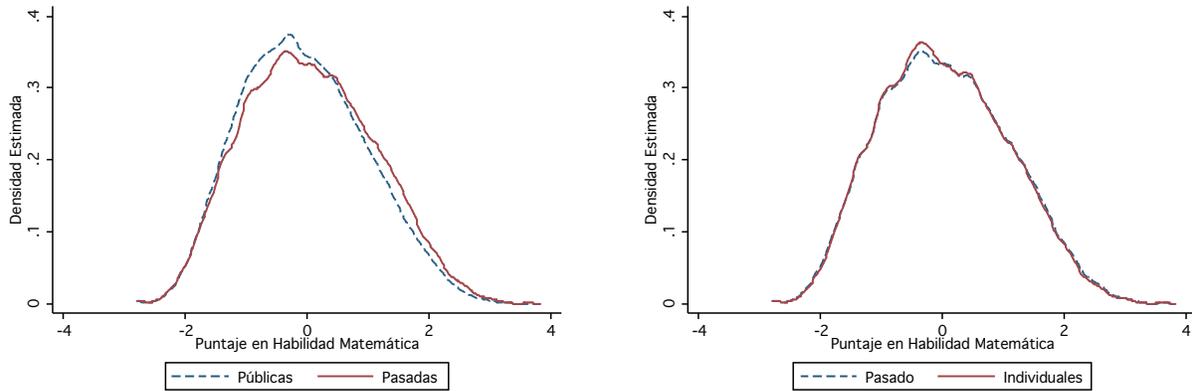
En una serie de artículos, Heckman y coautores describen el rol de las habilidades cognitivas y no cognitivas, en el desarrollo futuro de la persona, y explican cómo diferencias que se forjan a una edad temprana en este tipo de habilidades son más difíciles de cerrar (Heckman & Masterov, 2007; Heckman, 2008; Cunha, Heckman, Lochner, & Masterov, 2006). También señalan la importancia de las inversiones a una temprana edad para erradicar estas diferencias y los problemas a resolver en su diseño e implementación. En este sentido, la metodología desarrollada en el presente trabajo nos permite hacer una extensión para analizar la propuesta de Heckman & Masterov (2007); Heckman (2008); Cunha, Heckman, Lochner, & Masterov (2006) . Reclasificaremos las características de los estudiantes en dos tipos: características presentes y características pasadas.³ Posteriormente se aplicó la descomposición de Altonji, Bharadwaj, & Lange (2008) para medir la contribución de las variables pasadas sobre la brecha en calificaciones.

La Figura 5.4 presenta los resultados obtenidos de la descomposición. La Figura 5.4a compara el contrafactual “¿cuál sería la distribución de calificaciones de escuelas públicas si tuvieran la distribución de características pasadas de estudiantes de escuelas privadas?” y la distribución original de calificaciones de estudiantes de escuelas públicas. Se observa que el contrafactual estimado se localiza a la derecha de la distribución original indicando que existe una mejora en las calificaciones de los estudiantes de escuelas públicas. En la Figura 5.4b, se presenta el contrafactual: ¿cuál es la distribución de calificaciones de estudiantes en escuelas públicas si tuvieran la distribución de características pasadas y presentes de estudiantes de escuelas privadas? (línea continua) y el contrafactual cambiando sólo características pasadas. De la Figura 5.4b se infiere que

³Características pasadas: Promedio en Educación Primaria, Tipo de Escuela Primaria, Promedio en Educación Secundaria, Tipo de Escuela Secundaria. Características Presentes: Número de Horas en Clases de Matemáticas, Sexo, Fumar, Beber, Trabaja y Aspiraciones del Estudiante.

las características relacionadas con el presente de los estudiantes no contribuyen de forma importante a la brecha existente en calificaciones.

Figura 5.4: Función de Densidad Estimada.



(b) Contribución de Características Presentes

Fuente: elaboración propia utilizando datos de los cuestionarios de contexto de la prueba ENLACE 2008 para habilidad matemática. El contrafactual estimado es: ¿cómo se distribuyen las calificaciones de los estudiantes de escuelas públicas si tuvieran la distribución de características relacionadas con el pasado de los estudiantes de escuelas privadas?

En números, cambiar sólo las características relacionadas con el pasado del estudiante implica un aumento de 0.1 desviaciones estándar sobre la media de calificaciones de estudiantes de escuelas públicas⁴ mientras que incluir las características presentes reduce el número en 0.01 desviaciones estándar sobre la media, comprobando la pequeña contribución de las características presentes a la diferencia existente. La gran contribución del pasado de los estudiantes sobre la brecha en calificaciones comprueba las ideas propuestas por Heckman & Masterov (2007); Heckman (2008); Cunha, Heck-

⁴0.09 desviaciones estándar la contribución de las características individuales en la descomposición detallada con el modelo jerárquico

man, Lochner, & Masterov (2006).

En el 2009, el gasto por alumno para Educación Primaria, Educación Secundaria y Bachillerato fue de 12.2, 18.7 y 25.7 miles de pesos, respectivamente.⁵ Los resultados de esta sección sugieren una mayor atención durante los primeros años de educación y con el objetivo de, que un gasto por alumno más equitativo en Educación Primaria y Educación Secundaria sin descuidar los niveles posteriores, ayudaría a cerrar la brecha en calificaciones, en un futuro, entre estudiantes de escuelas privadas y públicas.

5.6. Implicaciones de Política Pública

La Evaluación Nacional del Logro Académico en Centros Escolares es una fuente valiosa de información y provee un panorama general del Sistema Educativo Mexicano. Los resultados presentados en las secciones anteriores complementan la información disponible y generan un margen para el diseño de políticas públicas.

El hecho de que las diferencias en recursos escolares contribuyan de forma importante a la diferencia en calificaciones en estudiantes de bajo rendimiento (18 %-40 %) es evidencia de la importancia de este grupo sobre el desempeño escolar para este tipo de estudiantes. Lo anterior sugiere que los programas con el objetivo de mejorar la infraestructura escolar, como el Programa Escuelas de Calidad (PEC) en Educación Básica, deben destinarse a escuelas donde la proporción de estudiantes con bajo rendimiento es mayor. Las políticas públicas para mejorar las condiciones familiares de los estudiantes se topa con problemas tanto el diseño como su implementación pues implica modificar el comportamiento de los familias. Sin embargo, debido a la importancia de estos factores sobre el desempeño de los estudiantes ubicados en el percentil 10 (40 %) no se

⁵SEP (2010): *Principales Cifras Ciclo Escolar 2008-2009*

pueden ignorar. El número de libros en el hogar se utiliza para aproximar la importancia de la educación en el hogar. Una mayor participación de los padres en actividades escolares sería un buen comienzo para mejorar las condiciones familiares en el corto plazo.

Para estudiantes de alto desempeño, los recursos escolares decrecen en importancia y son los factores institucionales los que adquieren relevancia. Dado el conjunto de variables incluidas, un programa educativo enfocado a mejorar el clima escolar es el más adecuado para escuelas con un buen desempeño escolar. Para lograr ese objetivo es necesario mejorar las relaciones entre los profesores y estudiantes y lograr una mayor integración de los estudiantes a actividades académicas. Como un segundo objetivo, la implementación de programas educativos más ambiciosos ayudaría a mejorar el factor de exigencia en las escuelas y por tanto afectando el desempeño escolar de los estudiantes en una forma positiva.

Respecto a la contribución de los factores relacionados con el pasado académico del estudiante, las estimaciones señalan la importancia de inversiones tempranas en la educación. En este sentido, el desarrollo de programas como Escuelas de Calidad (PEC) y Escuela Segura (PES) es adecuado e indispensable para mejorar el desempeño escolar. Así mismo, es necesario revisar si los programas no sufren de disyuntiva entre sus objetivos teóricos y su implementación pues, de acuerdo a [Hanushek \(2003\)](#), esto podría ser la causa de los bajos impactos que presentan sobre los puntajes obtenidos por los estudiantes. También es necesario analizar el Sistema de Educación Básica con mayor detalle, en vista de obtener resultados que permitan mejorar la focalización de los programas a este nivel, mejorar su eficiencia y con ello mejorar los resultados de los estudiantes a largo plazo.

Capítulo 6

Conclusiones

El objetivo es analizar la diferencia en calificaciones, entre estudiantes de escuelas públicas y escuelas privadas, utilizando los resultados de la Evaluación Nacional del Logro Académico en Centros Escolares para Educación Media Superior. La metodología utilizada permite estimar la distribución de calificaciones de estudiantes de escuelas públicas si tuvieran la distribución de características observables de estudiantes de escuelas privadas y, al estimar funciones de densidad, permite analizar de forma natural otros puntos en la distribución distintos a la media.

El primer resultado de la investigación es que las características observables logran explicar el 88 % de la diferencia media en calificaciones y tal porcentaje es mayor en los percentiles 25 y 75 (97 %-95 %). La parte de la diferencia no explicada gracias a las características no observables corresponde a una aproximación del beneficio real de estudiar en las escuelas privadas. Gracias al tipo de variables que incluimos el beneficio se debe a factores como la competencia en el mercado o mayor grado de libertad para resolver problemas. Los estudiantes de los percentiles 10, 50 y 90 son aquellos en donde el beneficio de estudiar en una escuela privada es mayor (12 %-20 %). Para estudiantes del percentil 25 y 75 el beneficio es menor (3 %-5 %).

Posteriormente, la parte que se logra explicar por las diferencias en características observables se desagregó, utilizando la metodología desarrollada por [Altonji, Bharadwaj, & Lange \(2008\)](#), en las contribuciones de cuatro grupos que la literatura económica considera relevante para explicar el desempeño de los estudiantes en pruebas estandarizadas: características individuales, características del hogar, recursos escolares y factores institucionales. En la media, las características del estudiantes explican el 43 % y las características del hogar 36 % de la diferencia mientras los recursos escolares y los factores institucionales sólo contribuyen en 7 % y 3 %, respectivamente y se encuentran en línea con los resultados reportados por la literatura. No obstante, la contribución de cada grupo varía respecto al rendimiento del estudiante. En estudiantes de bajo rendimiento (percentil 10), los grupos relevantes son las características del hogar y recursos escolares que contribuyen en un 62 % y 18 % , respectivamente, a la diferencia en calificaciones. En cuanto a los factores institucionales, su contribución es negativa (-9 %) y se argumenta que podría refleja un clima académico más complicado de las escuelas privadas en ciertos aspectos. A pesar del signo negativo, es necesario notar que en valor absoluto la contribución es importante. En cambio, las diferencias en factores institucionales y características del hogar son más importantes para explicar la diferencia en calificaciones en estudiantes de alto rendimiento (percentil 90).

Dado que la descomposición propuesta por [Altonji, Bharadwaj, & Lange \(2008\)](#) depende del orden, en la Sección 5.3 se presentan dos enfoques para obtener contribuciones independientes del orden. Bajo estos enfoques, la importancia de los factores institucionales en estudiantes de alto y bajo rendimiento se mantiene junto con otros resultados como la relación negativa entre los factores escolares y el puntaje. Por otro lado, con excepción de los estudiantes de bajo rendimiento, la importancia de las características individuales en la diferencia en calificaciones es innegable en la primera descomposición y bajo los dos enfoques. Por ello, siguiendo las ideas propuestas por [Heckman & Masterov \(2007\)](#), en la Sección 5.4 se analiza la contribución de los factores

relacionados con el pasado de los estudiantes. Los resultados indican que la diferencia en factores relacionados con el pasado académico logran explicar prácticamente toda la contribución de las características individuales a la diferencia en calificaciones (110 %) y por tanto, señalan la importancia de una mejor calidad educativa durante los primeros años de estudios.

En términos de políticas públicas, los resultados inducen un margen para la implementación de programas más focalizados. Un programa destinado a mejorar los recursos escolares sería más adecuado para escuelas con una mayor proporción de estudiantes de bajo rendimiento. Para escuelas con un buen desempeño lo indicado es un programa encargado de mejorar los factores institucionales, es decir, mejorar la relación profesor-estudiante, estar sujetos a mayores evaluaciones o programas de estudios más avanzados (mayor exigencia).

En conclusión, se exhibe evidencia de que las escuelas privadas son mejores que las escuelas públicas. Sin embargo, las estimaciones indican que sólo un 5 %-20 % de las diferencias en calificaciones se puede interpretar como un verdadero beneficio. El resto de las estimaciones señalan la importancia de estudiar los determinantes del desempeño escolar bajo la idea de que no todos los estudiantes son iguales. Omitir este hecho podría causar la pérdida de una fuente valiosa de información e importante para el diseño de políticas públicas. En temas de educación, la falta de información no está permitida, pues el costo futuro tanto social como económico puede ser más que considerable.

Bibliografía

- ALTONJI, J. G., P. BHARADWAJ, & F. LANGE (2008): “Changes in the Characteristics of American Youth: Implications for Adult Outcomes,” Discussion Paper 13883, National Bureau of Economic Research.
- ALVARÉZ, J., V. GARCÍA-MORENO, & H. PATRINOS (2007): “Institutional Effects as Determinants of Learning Outcome: Exploring State Variations in Mexico,” Working Paper 4286, The World Bank.
- ANGRIST, J., & V. LAVY (1999): “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement,” *Quarterly Journal of Economics*, 114, 533–575.
- BACKHOFF, E., A. BOUZAS, C. CONTRERAS, E. HERNÁNDEZ, & M. GARCÍA (2007): “Factores Escolares y Aprendizaje en México: El Caso de la Educación Básica,” Discussion paper, Instituto Nacional para la Evaluación de la Educación.
- BEDI, A. S., & A. GARG (2000): “The Effectiveness of Private Versus Public Schools: the Case of Indonesia,” *Journal of Development Economics*, 61, 463–494.
- BETTINGER, E. P. (2005): “The Effect of Charter Schools on Charter Students and Public Schools,” *Economics of Education Review*, 24, 133–147.
- CARD, D. (1993): “Using Geographic Variation in College Proximity to Estimate the Return to Schooling,” Working Paper 4483, National Bureau of Economic Research.

- CARD, D., & A. KRUEGER (1990): "Does School Quality Matter? Returns to Education and the Characteristics of Public School in the United States," Discussion Paper 3358, National Bureau of Economic Research.
- CUNHA, F., J. J. HECKMAN, L. LOCHNER, & D. V. MASTEROV (2006): "Chapter 12 Interpreting the Evidence on Life Cycle Skill Formation," vol. 1 of *Handbook of the Economics of Education*, pp. 697 – 812. Elsevier.
- DE GREGORIO, J., & J. LEE (2003): "Growth and Adjustment in East Asia and Latin America," Working Papers 245, Central Bank of Chile.
- DE HOYOS, R., J. ESPINO, & V. GARCÍA (2010): "Determinantes del Logro Escolar en México: Primeros Resultados Utilizando la Prueba ENLACE Media Superior," Discussion paper, Secretaría de Educación Media Superior.
- DEARDEN, L., J. FERRI, & C. MEGHIR (2002): "The Effect of School Quality on Educational Attainment and Wages," *The Review of Economics and Statistics*, 84(1), 1–20.
- DINARDO, J., N. M. FORTIN, & T. LEMIEUX (1996): "Labor Market Institutions and the Distributions of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–1044.
- DINARDO, J., & J. L. TOBIAS (2001): "Nonparametric Density and Regression Estimation," *Journal of Economic Perspectives*, 15(4), 11–28.
- DUNCAN, K., & J. SANDY (2007): "Explaining the Performance Gap Between Public and Private School Students," *Eastern Economic Journal*, 33(2), 177–191.
- EVERS, M., R. D. MOOIJ, & D. V. VUUREN (2008): "The Wage Elasticity of Labour Supply: A Synthesis of Empirical Estimates," *The Economist*, 156(1), 25–43.

- FIRPO, S., N. M. FORTIN, & T. LEMIUX (2009): "Unconditional Quantile Regressions," *Econometrica*, 77(Unconditional Quantile Regressions), 953–973.
- FORTIN, N., T. LEMIUX, & S. FIRPO (2010): "Decomposition Methods in Economics," Working Paper 16045, National Bureau of Economic Research.
- FUCHS, T., & L. WOESMANN (2004): "What Accounts for International Differences in Student Performance? A Re-examination Using PISA Data," Working Paper 1235, Ifo Institute for Economic Research.
- GRADDY, K., & M. STEVENS (2005): "The Impact of School Resources on Student Performance: A Study of Private Schools in the United Kingdom," *Industrial and Labor Relations Review*, 58(2), 435–451.
- GUICHARD, S. (2005): "The Education Challenge in Mexico: Delivering Good Quality Education to All," Working Paper 447, Organisation for Economic Co-operation and Development.
- HANUSHEK, E. (2003): "The Failure of Input-Based Schooling Policies," *The Economic Journal*, 113, 64–98.
- HECKMAN, J. J. (1974): "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 47(1), 153–161.
- (2008): "Schools, Skills, and Synapses," *Economic Inquiry*, 46(3), 289–324.
- HECKMAN, J. J., & D. V. MASTEROV (2007): "The Productivity Argument for Investing in Young Children," Working Paper 13016, National Bureau of Economic Research.
- HIRANO, K., G. W. IMBES, & G. RIDDER (2003): "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrie*, 71(4), 1161–1189.

- HOLLAND, P. (1986): "Statistics and Causal Inference," *Journal of the American Statistical Association*, 81(396), 945–960.
- KANG, C. (2007): "Does Money Matter? The Effect of Private Educational Expenditures on Academic Performance," Departmental working papers, National University of Singapore, Department of Economics.
- KRAMARZ, F., S. MACHIN, & A. OUAZAD (2009): "What Makes a Test Score? The Respective Contributions of Pupils, Schools and Peers in Achievement in English Primary Education," Discussion paper, Centre for the Economics of Education.
- LEIBBRANDT, M., J. A. LEVINSOHN, & J. MCCRARY (2010): "Income in South Africa after the Fall of Apartheid," *Journal of Globalization and Development*, 1(1), Artículo 2.
- LEMIUX, T. (2010): "Decomposing Changes in Wage Distributions: A Unified Approach," *Canadian Journal of Economics*, 35(4), 646–688.
- LLERAS-MUNEY, A. (2005): "The Relationship Between Education and Adult Mortality in the United States," *Review of Economic Studies*, 72(1), 189–221.
- LOCHNER, L., & E. MORETTI (2004): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *The American Economic Review*, 94(1), 155–189.
- MACHIN, S. (2008): "The New Economics of Education: Methods, Evidence and Policy," *Journal of Population Economics*, 21, 1–19.
- MCEWAN, P. J. (2004): "Indigenous Test Score Gap in Bolivia and Chile," *Economic Development and Cultural Change*, 53(2), 157–190.
- MROZ, T. A. (1987): "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economics and Statical Assumptions," *Econometrica*, 55(4), 765–799.

- OECD (2010): *Pisa 2009 Results: Executive Summary*. Organisation for Economic Co-operation and Development.
- PARZEN, E. (1962): "On Estimation of a Probabilidty Density Function and Mode," *The Annals of Mathematical Statistics*, 33, 1065–1076.
- PETERSON, P., & E. LLAUDET (2006): "On the Public-Private School Achievement Debate," Discussion paper, American Political Science Association.
- RIVKIN, S. G., E. HANUSHEK, & J. F. KAIN (2005): "Teachers, Schools and Academic Achievement," *Econometrica*, 73(2), 417–458.
- ROSENBLATT, M. (1956): "Remarks on Some Non-parametric Estimates of a Density Function," *The Annals of Mathematical Statistics*, 27, 832–837.
- SEP (2010): *Principales Cifras Ciclo Escolar 2008-2009*. Secretaría de Educación Pública.
- SILVERMAN, B. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.
- SOHN, K. (2010): "A New Insight into the Gender Gap in Math," *Bulletin of Economic Research*, 63.
- VANDENBERGHE, V., & S. ROBIN (2004): "Evaluating the effectiveness of private education across countries: a comparison of methods," *Labour Economics*, 11(4), 487 – 506, European Association of Labour Economists 15th Annual Conference, Universidad Pablo de Olavide, Seville, 18-21 September 2003.
- WEBBINK, D. (2005): "Causal Effects in Education," *Journal of Economic Surveys*, 19, 536–560.

WOESMANN, L. (2000): “Schooling Resources, Educational Institutions, and Student Performance: The International Evidence,” Working Paper 983, Kiel Institute of World Economics.

WORLD BANK (2010): *Education Data* <http://data.world.org/topic/education>.

YAMAUCHI, F. (2005): “Why Do Schooling Returns Differ? Screening, Private Schools, and Labor Markets in the Philippines and Thailand,” *Economic Development and Cultural Change*, 53(4), 959–981.

Apéndice A

Diversos

Para el trabajo la mayor parte de las variables que se utilizaron en la implementación empírica son las respuestas originales. Algunas variables se construyeron para obtener una respuesta menos detallada y a continuación se describe su construcción:

- *Fumar*. En el cuestionario de contexto se les pregunta a los estudiantes “¿Cuántos días a la semana fumas?”. Se considero que el estudiante fumaba si al menos un día de la semana realizaba la actividad.
- *Beber*. La pregunta original es “¿Cuántos días a la semana ingieres bebidas alcohólicas?”. Se considero que un estudiante bebía si al menos un día a la semana ingería bebidas alcohólicas.

En el sección 3.2 se describe la estrategia empírica para implementar la metodología desarrollada por [DiNardo, Fortin, & Lemieux \(1996\)](#). En la Tabla [A.1](#) se presentan los coeficientes del modelo logit utilizado para estimar $P(priv|z_i)$.

Tabla A.1: Modelo Logit Estimado para Calcular Pesos $\Phi(z)$

Sexo	0.2650***	Estudios Madre		Estudios de Profesores	
Fuma	0.2778***	No estudió	-0.4494***	Bachillerato	-1.1876
Bebe	0.2928***	Primaria	-0.3993***	Normal Superior	-3.6948***
Trabaja	-0.2989***	Secundaria	-0.2628***	Pedagogía	-1.8558*
Secundaria Privada	2.3194	Preparatoria	-0.1993**	Licenciatura	-3.4879***
Aspiraciones		Carrera Técnica	-0.1392*	Posgrado	-3.1782***
Técnico Universitario	-0.0368	Licenciatura	-0.0453	Estado de los Laboratorios	
Licenciatura	0.1744	Estudios Padre		Sí son suficientes	-0.0454
Posgrado	0.2633**	No estudió	-0.3777	No son suficientes	-1.6510***
Primaria Privada	0.0336*	Primaria	-0.3873**	Internet	-0.8427***
Promedio Primaria		Secundaria	-0.2776***	Examen Oposición	0.2159***
6.5-6.9	-0.0946	Preparatoria	0.715x10 ⁻⁶	Asociaciones Estudiantiles	0.7562***
7.0-7.4	-0.0449	Carrera Técnica	-0.0758	PISA	0.3524***
7.5-7.9	-0.1470	Licenciatura	-0.0336	Porros	2.4553***
8.0-8.4	-0.2244	Computadora	0.3005***	Huelgas	-0.5314***
8.5-8.9	-0.1904	Piso	0.7399***	Tutorías	0.9664***
9.0-9.4	-0.3467	Libros en Casa		Estudiantes-Profesores	
9.5-9.9	-0.5066	Ninguno	-0.1296	Buena	2.0665***
10	-0.3151	1 a 10 0	0.2431**	Excelente	3.3257***
Horas de Matemáticas		11 a 25	0.0217	Estudios Director	
1	0.1845**	26 a 50	0.0947	Bachillerato	2.8328***
2	0.2014*	51 a 100	0.0680	Normal Superior	1.4277***
3	-0.7102***	101 a 200	0.1202	Pedagogía	1.1624***
4	0.0967	201 a 500	0.1180	Licenciatura	0.7343***
5	-1.0476***	Número de Cuartos		Especialidad	1.3933***
6	-1.0436***	0	-0.8761***	Maestría	0.9167***
7	-0.0775	1	-0.6224***	Doctorado	1.6031***
Promedio Secundaria		2	-0.5021***	Exigencia Escolar	
6.5-6.9	-0.3619***	3	-0.3223***	Exigente	0.7141***
7.0-7.4	-0.4359***	4	-0.2282***	Muy Exigente	1.2438***
7.5-7.9	-0.5036***	5	-0.1488*	Vigilancia	-0.5837***
8.0-8.4	-0.6239***	6	-0.1219		
8.5-8.9	-0.7342***				
9.0-9.4	-0.8125***				
9.5-9.9	-0.9730***				
10	-0.9278***				

Fuente: elaboración propia usando base de datos del diseño muestral de los Cuestionarios de Contexto 2008. Todos los resultados son ajustados por los pesos de la muestra.

Índice de cuadros

4.1. Diferencias en Calificaciones entre Escuelas Privadas y Públicas: Base Original y Base Construida.	27
4.2. Calificaciones Medias por Estado en Habilidad Matemática	28
4.3. Calificaciones Medias por Estado en Habilidad Lectora	29
4.4. Diferencias en Puntajes sobre Características Individuales	34
4.5. Diferencias en Puntaje sobre Características del Hogar	35
4.6. Diferencias en Puntajes sobre Características de la Escuela	36
4.7. Diferencias en Puntajes sobre Características Institucionales	37
4.8. Diferencias en Características Observables entre Privada y Públicas	39
4.9. Variables Incluidas en el Modelo Logit	40
5.1. Descomposición Detallada Original	45
5.2. Diferencia entre Privadas y Públicas en Factores Institucionales en el Percentil 10	49
5.3. Descomposición Detallada con Características Individuales antes que Familiares	52
5.4. Descomposición Detallada Agrupando Características Individuales y Familiares .	53

5.5. Descomposición Detallada Agrupando Características Individuales y Familiares de la Tabla 5.3	53
A.1. Modelo Logit Estimado para Calcula Pesos $\Phi(z)$	70

Índice de figuras

4.1. Funciones de Densidad Originales por Sostenimiento	30
4.2. Funciones de Densidad de Puntajes por Estado y Sostenimiento	32
5.1. Contribución de las Características Observables	42
5.2. Descomposición Detallada en Características Individuales, Características Familiares, Recursos Escolares y Factores Institucionales	44
5.3. Diferencias en Funciones de Producción McCrary <i>et al.</i> (2010)	55
5.4. Descomposición de Características Individuales en Factores del Pasado y Factores del Presente	57