

Algunos Estadísticos Sobre la Educación Primaria en México: Reporte para el Banco Mundial

Salomé Aguilar, Paola Mejía , Enrique Seira*

31 de enero de 2020

*Quisiera agradecer a Rafael de Hoyos por impulsar este tipo de proyectos de investigación. Quisiera también agradecer a un equipo de investigación excepcional: Eduardo Rivera, Emiliano Ramírez. Corresponding author: enrique.seira@gmail.com.

1. Introducción

La educación es de suma importancia para un país, no solamente porque el capital humano es un insumo indispensable en el crecimiento económico, sino porque la educación genera desarrollo personal, mayor expectativa de vida y movilidad social. La educación tiene grandes externalidades sociales positivas: reduce el crimen, aumenta la civilidad y la cohesión social. En suma, la educación es tan importante que frecuentemente equiparamos la igualdad de acceso a la educación de calidad con la igualdad de oportunidades.

¿Pero qué sabemos sobre la calidad de la educación en México? Sabemos de verdad poco. En parte esto se debe a la falta de disponibilidad de datos. Esta escasez de información ha sido en parte por diseño –para evitar la transparencia– y en otra parte por falta de interés de las autoridades educativas de publicar la información de forma desagregada. A falta de información y análisis cualquier opinión vale y la ideología se impone a la verdad.

En este documento combinamos información proporcionada por el Banco Mundial, junto con información de otras fuentes oficiales para hacer algunos diagnósticos de la educación primaria en México. Primero, documentamos que la calidad de la educación, medida por el desempeño en la (extinta) prueba Enlace es muy heterogénea en el país: hay escuelas muy buenas y otras muy malas. Encontramos dos hallazgos importantes: (a) que las escuelas de mala calidad son tanto públicas como privadas, aunque en promedio las públicas tienen peor desempeño en Enlace; (b) que la geografía explica poco de las diferencias de calidad entre las escuelas. Aun dentro de un mismo código postal encontramos escuelas buenas conviviendo con escuelas malas.

Luego examinamos si las mejores escuelas están creciendo en matrícula, y, en particular, si cuando un alumno se cambia de escuela tiende a cambiarse a escuelas con mejor desempeño en Enlace. Descubrimos que esto no es así. Tanto para escuelas públicas como para privadas, la proporción de estudiantes que se cambian a peores escuelas es casi igual al porcentaje que se cambian a mejores escuelas. Esto quiere decir que: (a) o no conocen la calidad de la escuela, (b) o no les importa porque Enlace no mide resultados que importen, (c) o hay factores de calidad que se correlacionan negativamente con Enlace, y que les importaron más estos factores.

Es difícil establecer si conocen la calificación de Enlace o no. Pero es relativamente fácil probar que Enlace está correlacionada con variables importantes y por lo tanto debería importar. Por primera vez en México podemos demostrar que las calificaciones que saca un alumno de primaria en Enlace están fuertemente relacionadas casi 10 años después con su desempeño en el mercado laboral, con la probabilidad de terminar preparatoria, con la probabilidad de ser convicto criminal, y con la probabilidad de trabajar como empleado en el sector formal. Esto significa que la prueba Enlace no es solamente un número en un papel, sino un *proxy* de resultados de vida individuales y sociales que realmente nos interesan.

2. Datos

En este documento usamos cinco bases de datos para documentar hechos sobre los resultados de Enlace entre escuela. Las cinco fuentes de información fueron anonimizadas con el fin de respetar la confidencialidad de la información, de tal forma que no es posible saber la identidad de las personas.

2.1. Panel Enlace

El panel de enlace consiste en el universo de niños que realizó la prueba enlace en los años 2007-2012. En total tenemos en esta base 26,811,134 niños de los cuales 21,287,112 son alumnos de primaria. Tenemos además 129,088 escuelas de las cuales 95,913 son primarias. 90 % de las escuelas primarias de México son públicas, y captan 91.3 % de los niños.

En este documento nos enfocamos en primarias. La prueba enlace en primaria cubre de tercero hasta sexto. En promedio observamos a cada niño en 2.3 ocasiones en primaria, por la forma en que están construidas las generaciones¹. Para cada niño observamos su calificación de Enlace en matemáticas y en español en todos los años y en que escuela está. Sin embargo para algunos años específicos se aplicaron pruebas de conocimiento en otras materias, como Ciencias, Formación Cívica y Ética Historia, Geografía.

2.2. Censo de escuelas 911

Usamos el Censo 911 para los años 2006 a 2012. Este censo se llena por parte del director de la escuela, y contiene información, por ejemplo, del número de alumnos, del número y nivel educativo de profesores, del número de docentes especiales y de la nacionalidad, discapacidad o aptitudes sobresalientes de los alumnos. El *match* entre este censo y la información de escuelas que hicieron Enlace es de 78 %.

2.3. Censo de Escuelas, Maestros y Alumnos de Educación Básica y Especial

Censo de escuelas primarias llevado a cabo por el INEGI-SEP en 2013². En este censo se visitaron 273,317 escuelas (centros de trabajo), y se censo un personal de 1,987,511, atendiendo a 23,667,973 alumnos, en 61,989 localidades. Este censo contiene además variables de infraestructura del plantel como características constructivas del inmueble educativo, espacios (aulas, laboratorios, bibliotecas, espacios recreativos, sanitarios), servicios (agua, electricidad, internet por tipo, etc), y equipamiento (e.g. número y características de los equipos de cómputo).

2.4. Información del INE

La información del INE nos permite medir tres variables de resultados: primero si el alumno Enlace votó en las elecciones del 2015. Segundo, si una persona registrada en el padrón del año 2018 fue declarada culpable por un juez en un proceso penal del año 2009 al 2018. Tercero, si la persona declara haber terminado preparatoria.

Operativamente las bases de Enlace e INE se hicieron en las oficinas del INE por el propio personal del INE. Se partió de la base de Enlace, con las 33,492,630 alumnos que hicieron Enlace de 2006 al 2012. El INE buscó a estas personas en el padrón de votantes de 2018. El 74.5 % de estos tenían más de 18 años en enero de 2018 (24,768,472), y podían por tanto potencialmente ser encontrados en el padrón del INE. Efectivamente se encontraron 19,233,531 de ellos, lo que representa al 77.6 % de match. Creemos que esta tasa de encontrados es razonable, ya que no todos los mexicanos tienen credencial para votar. Dentro de la base que hizo match 10,429,585 alumnos Enlace tenían credencial para votar en 2015 y 40 % de ellos votaron en esa fecha. Esto es un poco mas bajo que la tasa nacional de participación.

¹e.g. si una alumno esta en 6to de primaria cuando comenzó a aplicarse la prueba solo podemos verlo una vez en primaria.

²<http://www.censo.sep.gob.mx/>

Para el análisis utilizamos únicamente a los niños que observamos en primaria, por este motivo explicamos como se ven los números para el subconjunto de alumnos Enlace de primaria. Para primaria contamos con 24,688,329 alumnos en Enlace, de los cuales 16,277,768 eran mayores de 18 años en 2018.³ De estos 16 millones el INE encontró a 11,360,744 en el padrón de votantes del 2018 (69.7%). Sin embargo en el 2015 solo 4,933,669 tenía 18 años. De estos 2,065,459 asistieron a votar en 2015.

2.5. Información del IMSS

La base de datos Enlace se unió con la base de datos del IMSS –sin que los investigadores tuvieran acceso a la información personalizada– de forma que es posible saber qué alumno estaba empleado en el sector formal en diciembre del 2018, y cuál fue su salario promedio en 2018. El resultado de la unión de ambas bases es que se pudo en contar cerca de 15% de los alumnos de Enlace en el IMSS en 2018. Creemos que la calidad del match es alta, por lo que interpretamos que los no encontrados no son empleados formales. Pueden estar desempleados, trabajar en el sector informal, o tener su propio negocio. Recuérdese que si los alumnos estaban en primaria desde tercero a sexto de primaria en 2007-2012, en el 2018 tendrían entre 15 y 23, así que no todos pueden estar en el IMSS como empleados.

3. Algunos resultados de estadística descriptiva sobre la calidad de las escuelas

En esa sección del reporte hacemos un recuento de algunos estadísticos descriptivos. La intención es mostrar dos cosas: la variación en la calidad de las escuelas en México medida por la prueba Enlace y algunos determinantes de esta variabilidad en la calidad. Por otro lado, mostrar como Enlace es un predictor de resultados de vida individuales y socialmente valiosos, y por tanto argumentar que una prueba como Enlace es valiosa y debería restablecerse en México.

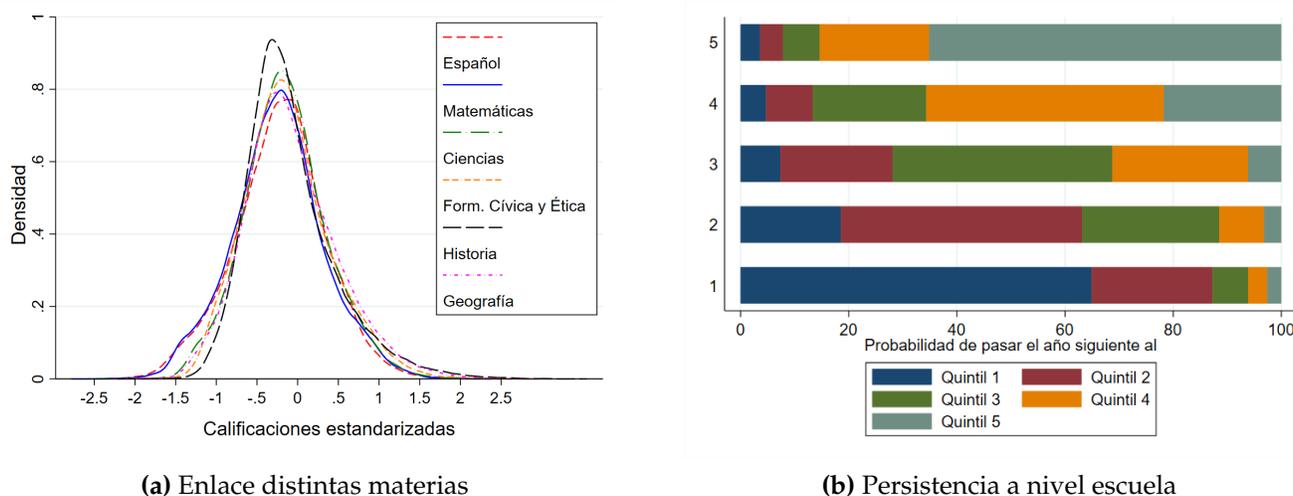
3.1. La calidad entre escuelas es muy variable

Si todas las escuelas tuvieran la misma calidad no importaría a cual escuela van los niños. Pero, en realidad, ¿cuánto varía la calidad de las escuelas? Si medimos calidad por medio del desempeño en la prueba Enlace, encontramos que existe gran heterogeneidad entre escuelas. La Figura 1a muestra la distribución de calificaciones entre escuelas, después de estandarizar (z_{igt}).⁴ Apreciamos una distribución simétrica, suave, para todas las materias que Enlace evalúa. Pero también la distribución tiene una alta varianza: en matemáticas, por ejemplo, 41% de las escuelas se encuentran a más de 0.5 desviaciones estándar de la media. Para dar una idea de la magnitud, en educación se considera como regla de dedo que intervenciones que mejoran las calificaciones en 0.2 desviaciones estándar son exitosas. Esto significa que existe un abismo entre las escuelas con buen y mal desempeño en Enlace.

³De los niños que tenían más de 18 años en 2018 y tienen INE la edad promedio es de 21 años y el 49.4% son mujeres.

⁴Generalmente en educación, para facilitar comparaciones, se reportan resultados en desviaciones estándar. En este caso el proceso de estandarización se llevó a cabo en los siguientes pasos: primero para cada grado y año calendario se calcula el promedio μ_{gt} y la desviación estándar σ_{gt} . Segundo, se estandariza la calificación del alumno i en el año t en el grado g como $z_{iegt} = (c_{igt} - \mu_{gt})/\sigma_{gt}$. Tercero, sacamos el promedio de z_{igt} dentro de cada escuela. Luego calculamos el promedio a nivel escuela e usando todos los años: $z^e := \frac{1}{\#t \times \#g \times \#i} \sum_t \sum_g \sum_i z_{iegt}$

Figura 1: Diferencias entre escuelas en Enlace, y su persistencia



Nota: a) Promedio de las calificaciones estandarizadas a nivel año y grado que obtuvieron las escuelas que presentaron la prueba ENLACE. Los promedios están a nivel escuela. Se presentan las densidades kernel de las pruebas de matemáticas, español, ciencias, formación cívica y ética, historia y geografía. b) Se calcula el promedio de las calificaciones estandarizadas de matemáticas de la escuela, luego se calcula el quintil en el que cae la escuela en cada año y la probabilidad de que la escuela cambie de quintil en años consecutivos

La calificación de Enlace es un indicador estable. Una escuela no cambia mucho en su ranking de escuelas año a año. Por ejemplo, calculamos que la probabilidad que una escuela pase del quintil más bajo al 1er, 2do, 3er, 4to y 5to quintil, es de 64.9 %, 22.3 %, 6.6 %, 3.5 % y 2.5 % respectivamente. La Figura 1b muestra estos porcentajes para cada quintil de partida.

En este documento nos concentraremos en los resultados de la prueba de Enlace de matemáticas. Esto está justificado porque la correlación entre pruebas para una misma escuela es alta: la Tabla 1 muestra que están entre 0.59 y 0.76. Es decir, las escuelas buenas en una materia son también buenas en las demás. Por ejemplo, si una escuela está en el quintil más alto de calificaciones en matemáticas –usando todos los años de esa escuela– la probabilidad de que también esté en el quintil más alto de español es de 81.8 %.

Tabla 1: Correlaciones a nivel escuela entre diferentes materias

	Matemáticas	Ciencias	Civismo	Geografía	Historia
Español	0.76	0.74	0.76	0.75	0.64
Matemáticas	-	0.70	0.68	0.73	0.62
Ciencias		-	0.63	0.64	0.60
Civismo			-	0.55	0.59
Geografía				-	0.62

Nota: Primer se obtuvo z^e para cada materia con el procedimiento descrito anteriormente. Una vez que la base de datos esta a nivel escuela se calculo la correlacion en el desempenio Enlace en las distintas materias que se miden.

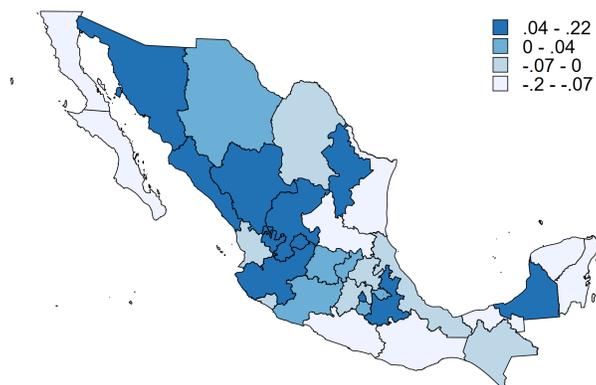
Este hecho facilita escoger escuela en base a la calificación Enlace, ya que no hay que sacrificar una materia por otras. El resto del documento se enfoca en los resultados de matemáticas solamente.

3.2. La calidad varía fuertemente aun dentro de un mismo código postal

La subsección anterior muestra una amplia diferencia entre el desempeño de escuelas del país. Uno podría preguntarse si la variación en la calidad de las escuelas es puramente geográfica, por ejemplo con mayores calificaciones en CDMX que en Oaxaca. Esto podría resultar de diferencias en recursos u organización de la educación en el Estado donde está localizada la escuela⁵, o simplemente porque las condiciones socio-económicas de los alumnos varían mucho entre Estados. Hay algo de esto. El mapa de la Figura 2 grafica el promedio de calificaciones z^e por Estado y documenta que Guerrero, por ejemplo, tiene una calificación 0.3 desviaciones estándar menor calificación que CDMX.

Sin embargo, lo sorprendente es que la calidad varía de forma importante aun *dentro* de un municipio, más aun, dentro de un código postal (CP). De hecho, una regresión de calificaciones z_{igt} , colapsada a nivel alumno, contra efectos fijos de código postal solo explica 8.3% de la varianza, es decir el 91.7% de la variación entre escuelas se debe a diferencias en las escuelas *dentro* de un código postal.

Figura 2: Diferencias entre estados



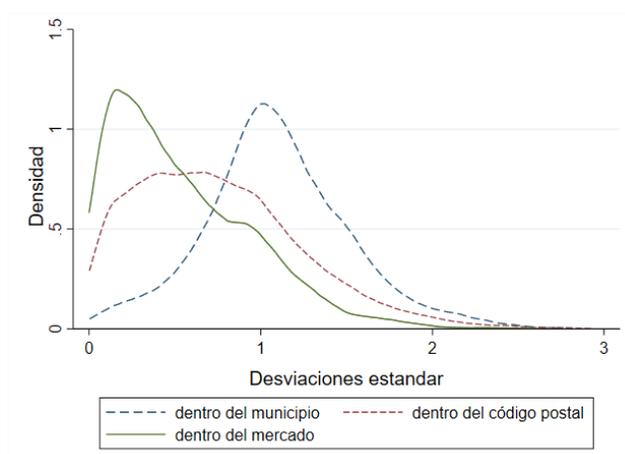
Se calcula el promedio de las calificaciones estandarizadas de matemáticas del estado para todos los años para todos los grados.

La Figura 3 es ilustrativa, y hace lo siguiente: para cada municipio busca cuál es la calificación de la escuela de ese municipio que está en el percentil 90 y le resta la calificación de la escuela de ese mismo municipio en el percentil 10. Es decir, se construye una medida de diferencia entre escuelas dentro de un mismo municipio. Hacemos lo mismo también con códigos postales y con mercados de escuelas.⁶

⁵Recuérdese que la educación primaria esta descentralizada en los Estados.

⁶Usamos la siguiente definición de mercados. Dos escuelas pertenecen a un “mercado” si en el periodo 2007-2012 compartieron al menos 2 alumnos, en el sentido de que estos alumnos estuvieron en ambas escuelas en diferentes años y ambas escuelas están a menos de 30 km de distancia. 42% de los mercados tienen mas de una escuela, y cada uno de estos tiene 12.3 escuelas en promedio. Esto de por si muestra que muchas escuelas parecen estar aisladas en México y no enfrentan competencia, ni los alumnos tienen posibilidades de elección.

Figura 3: Diferencias entre escuelas aun dentro de la misma geografía



(a) Percentil 90 menos percentil 10

Nota: a) Se calcula el promedio de las calificaciones estandarizadas de matemáticas de la escuela para todos los años y para todos los grados. Para cada grupo se encuentra el promedio que tienen las escuelas que se encuentran en el percentil 10 y 90 dentro del municipio, mercado y código postal y se calcula la resta. $Diferencia_j = percentil90_j - percentil10_j$ con el subíndice j que puede ser el mercado, el municipio o el código postal.

Como puede esperarse, por pura estadística, encontramos mayor variación al interior del municipio que del código postal, y en este que en el mercado. Lo importante es que aun dentro de un código postal la variación es brutal, 0.75 desviaciones estándar entre la escuela del 90^{mo} percentil y la del 10^{mo} en promedio, pero 28.1% de los códigos postales tienen mas de 1 desviación estándar de diferencia entre sus escuelas. Aun dentro de los mercados, la diferencia entre el 90^{mo} percentil y la del 10^{mo} es 0.54 desviaciones estándar en promedio. En análisis no reportados aquí encontramos que estos resultados se mantienen si vemos escuelas publicas solamente o privadas solamente.

Lo anterior importa por varias razones. Primero, el estatus socio-económico de la localidad (o cualquier otra variable geográfica) en donde están las escuelas explica poco. El resto se debe a diferente selección de estudiantes entre escuelas, o distintas practicas educativas entre las escuelas. Segundo, uno pensaría que no pueden convivir en una área geográfica chica escuelas buenas y malas, porque las escuelas malas cerrarían. Lo que observamos es que la competencia entre escuelas no está siendo tan fuerte como para eliminar a las escuelas con menor desempeño en Enlace aun dentro de un código postal.⁷ Tercero, el hecho de que aun dentro de mercados haya estas diferencias sugiere que el mismo tipo de alumnos están yendo a escuelas buenas y malas. En la Tabla 2 mostraremos que muchos alumnos se cambian de escuela, y que gran parte de las veces es a peores escuelas.

3.3. Escuelas públicas versus privadas

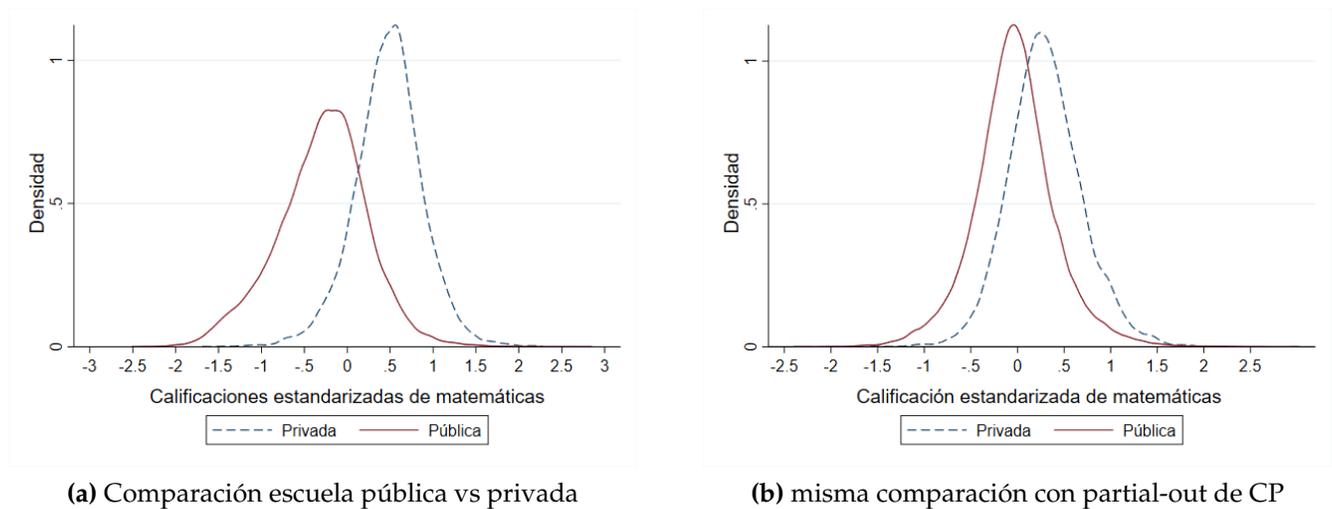
¿Podrá ser que la heterogeneidad que observamos se deba a que estamos comparando escuelas públicas versus privadas? ¿Cuáles son mejores? Una forma de comparar escuelas publicas y privadas es mostrar un histograma o densidad de las calificaciones, es decir que porcentaje de escuelas sacan cierta

⁷En un paper académico encontramos que las escuelas privadas con menor desempeño cierran con mayor probabilidad subsecuentemente, pero que esto no es verdad para las escuelas publicas. Las publicas malas se dan el lujo de seguir operando.

calificación, y hacerlo por separado para escuelas publicas y privadas. La Figura 4a muestra justo esto.

Hay varias lecciones de la Figura. Primero, que las escuelas privadas y publicas comparten soporte, en especial si comparamos a las del mismo código postal (Figura 4a(b)). Es decir, que por cada escuela privada buena encuentras también una publica buena, y lo mismo para las malas. Segundo, que aun dentro de las escuelas públicas y de las privadas hay gran heterogeneidad. Hay escuelas privadas buenas y otras patito. Tercero, que en promedio las escuelas privadas tienen mejor desempeño en Enlace que las públicas; de hecho, las escuelas privadas tienen puntajes de 0.6 desviaciones estándar por encima de las publicas (sin controles) y 0.3 controlando por códigos postales y características de la escuela y de los estudiantes⁸ Esta ventaja de las escuelas privadas sobre las publicas se da en todas las materias, incluyendo por ejemplo educación cívica. Esto es importante porque la teoría es que un papel importante de la educación provista por el Estado es inculcar valores cívicos y generar cohesión e identidad social.

Figura 4: Distribución de calificaciones: Publicas vs. Privadas



El panel (a) muestra densidades kernel de z^e , por separado para escuelas publicas y privadas. El panel (b) primero corre una regresión OLS de z^e versus efectos fijos de códigos postales incluyendo escuelas publicas y privadas en la misma regresión, para quitar su influencia, y luego muestra densidades kernel de los residuos de esa regresión. De tal forma que se comparan escuelas publicas y privadas en el mismo código postal.

⁸Se controló por las siguientes características de la escuela: razón maestro alumno, razón de maestras mujeres y maestros hombres, tamaño de los salones, computadoras por alumnos, y costo anual de la escuela. Asimismo se controló por las siguientes características de los alumnos usando la prueba de contexto: nivel de escolaridad del padre, nivel de escolaridad de la madre, tiene refrigerador, estufa, horno de micro-ondas, televisión, internet y número de autos en el hogar.

3.4. Cambios entre escuelas

Quisiéramos entender la actividad del mercado en el sentido de cuánto se cambian de escuela los estudiantes, y en particular, si se cambian a escuelas con mejor desempeño de Enlace. La Tabla 2 muestra que un poco más de 1.5 millones de alumnos se cambian de escuela en nuestro periodo muestral (3.17% si dividimos por el número de estudiantes-año). El mayor numero de cambios es de una escuela pública a otra pública, seguida por cambios privada-privada, mostrando que estos mercados están segmentados.

Lo impresionante es que los cambios hacia escuelas de menor calidad en Enlace (0.2 desviaciones estándar menos) son casi tan frecuentes que movimientos hacia escuelas de mayor calidad (0.2 desviaciones estándar más), y esto se da tanto dentro de publicas como de privadas, aunque para las publicas es un poco mayor el problema 34% alumnos que se cambian de publica a publica van a peores escuelas, mientras que 23% de los alumnos de privada lo hacen.

Las demás columnas de la tabla exploran si este comportamiento de cambio de escuela es diferente para alumnos que están por arriba o por debajo de su escuela origen (uno podría pensar que los de arriba de la media están mas informados), pero no encontramos grandes diferencias.

Tabla 2: Cambios de escuelas, por calidad Enlace y por tipo de escuela

cambio	todos los alumnos		buenos alumnos		malos alumnos		casos
	a mejor	a peor	a mejor	a peor	a mejor	a peor	
publ-publ	0.27	0.30	0.31	0.27	0.24	0.33	1,265,899
publ-priv	0.59	0.12	0.63	0.10	0.54	0.15	58,218
priv-publ	0.09	0.65	0.10	0.61	0.07	0.68	101,971
priv-priv	0.26	0.33	0.29	0.29	0.23	0.37	147,681

Una escuela de destino es mejor que la de origen si la media de sus calificaciones es 0.2 desviaciones estándar mayor, es peor si la media es 0.2 desviaciones estándar menor. Un estudiante es bueno si su calificación es mejor que la mediana de sus escuelas de origen y destino y es malo si su calificación es menor a la mediana tanto en su escuela de origen como de destino.

¿Por qué los papas están escogiendo cambiar a sus hijos a escuelas de menor calidad? Una hipótesis es que las buenas escuelas suben la colegiatura y los papas al no poder pagar se van a una mas barata que a su vez tiene peor desempeño en enlace. Esta hipótesis no explicaría porque el fenómeno se da en cambios de publicas a publicas dado que estas son gratuitas. Una segunda hipótesis es que los padres desconocen la calidad de la escuela medida por Enlace. Una tercera es que sí conocen la calificación Enlace pero que ésta no les importa. Una cuarta hipótesis es que sí la conocen y sí les importa, pero también les importan otras cosas, como infraestructura de la escuela, o cercanía a la casa, y que están dispuestos a sacrificar calidad medida por Enlace por estas otras cosas.

En este reporte descriptivo no nos corresponde evaluar estas hipótesis. Pero podemos decir dos cosas. Conjeturamos que la cuarta hipótesis no es verdadera. Enlace esta correlacionada positivamente con varios otros proxies de calidad de la escuela, como tener internet, computadoras por alumno, y el monto de colegiatura anual (para el caso de las privadas), como se muestra la Tabla 3. Desafortunadamente no

tenemos coordenadas de los hogares y no podemos calcular distancia a las distintas escuelas. Tampoco creemos que la tercera hipótesis sea verdadera, es decir que Enlace no le importe si la supieran. O al menos creemos que si les debería importar. La razón es que Enlace es un buen *proxy* de variables como crimen, terminar preparatoria, y salario, como se muestra en la Sección 4.

Tabla 3: Correlaciones de la infraestructura y profesores de la escuela con resultados de ENLACE

Características de la escuela	
Colegiatura anual	0.334
Alumnos por salón	-0.129
Escuela pública	-0.457
Tiene internet	0.143
Tiene drenaje	0.358
Tiene biblioteca	0.149
computadoras alumnos salón	0.260
Tiene juegos	0.191
Horas que labora	0.028
Tiene consejo escolar	0.014
Tiene sociedad de padres	-0.065
Variables docentes	
Años de educación de profesores	0.062
Profesores por estudiante	0.030
Salario anual de los profesores	0.190
Otras variables	
% de alumnos extranjeros	0.102
% de alumnos atrasados	-0.297
% de alumnos repetidores	-0.370

Las correlaciones se calculan con respecto al promedio de la calificación estandarizada de matemáticas para los años 2006 hasta 2012, los datos de la infraestructura de la escuela se sacaron del cuestionario de CEMABE realizado en 2013 a las escuelas.

4. Enlace es predictivo de resultados en el largo plazo

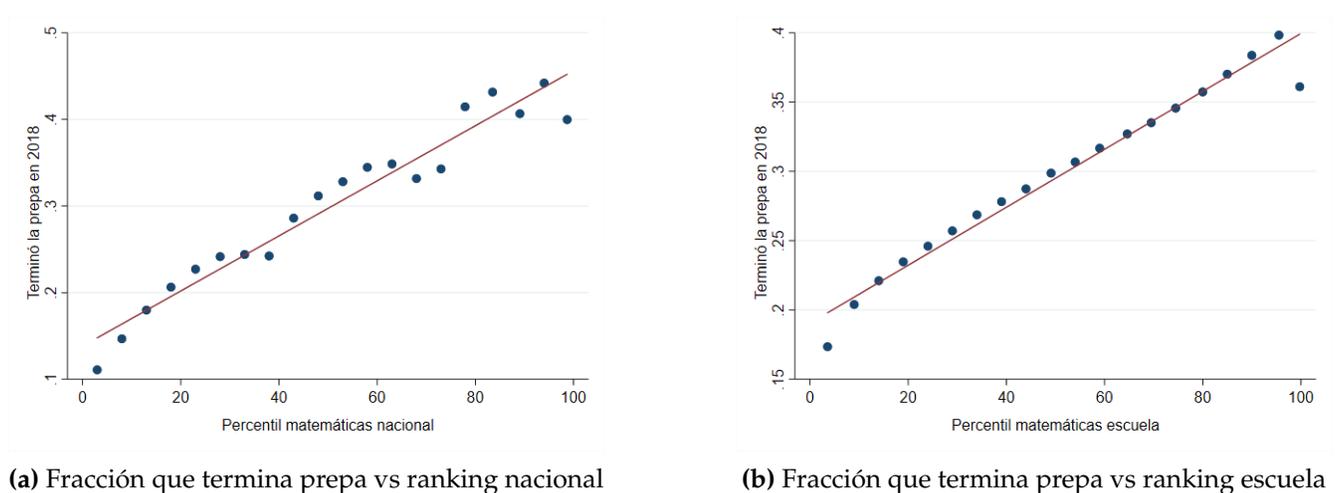
Una pregunta valida es ¿por qué tanto énfasis en pruebas como Enlace aquí y en la literatura sobre educación? La respuesta es que estas pruebas estandarizadas están diseñadas para medir un conjunto de conocimientos y habilidades que consideramos deseables. Esto ya de por sí es importante. Afortunadamente podemos ir mas allá de eso. En este documento demostramos que el desempeño en la prueba Enlace es un *proxy* de resultados personales y sociales importantes. El desempeño de un alumno en Enlace es un predictor de cómo le va a ir en la vida. Predice la probabilidad de cometer un crimen, terminar preparatoria, y votar en elecciones. Esto es importante. Pasemos a explicar los resultados.

El primer paso es tener una forma fácil de comparar *alumnos* (ya no escuelas) que sea intuitivo. Decidimos mostrar los resultados en términos del ranking nacional del alumno, agrupándolos en centiles: desde el lugar 100 que agrupa al 1% de los mejores alumnos⁹, hasta el lugar 1, que agrupa al 1% de las peores escuelas del país. Calculamos dos rankings: el nacional mide en qué lugar quedó un alumno comparado con todos los alumnos de México; y el segundo mide en que lugar salió el alumno con respecto a los otros alumnos de *la misma escuela*. Esto es importante porque no compara alumnos de zonas ricas contra aquellos en zonas pobres, ni alumnos en escuelas buenas versus malas. En la medida en que los niños que van a una escuela sean socio-económicamente parecidos, este tipo de comparación *dentro* de escuelas controla por nivel socio-económico.

El segundo paso es establecer cómo el lugar en el ranking de un alumno se relaciona con su vida futura. Buscamos una forma didáctica de mostrar estas relaciones sin imponer formas funcionales a los datos. Lo que decidimos fue hacer gráficas de *binscatter* dividiendo el eje X en veinte intervalos iguales y, para cada intervalo, se calcula el promedio de la variable Y en ese intervalo y se grafica como un punto. En este sentido, graficamos la esperanza condicional de Y dado X de forma no paramétrica.

Primero nos concentramos en resultados “privados” como terminar preparatoria y después en “sociales” como crimen y voto. Las Figuras 5a y 5b muestran la fracción que terminó la preparatoria en función de su ranking nacional (panel a) o en su escuela (panel b). Los resultados muestran que hay una relación casi lineal y muy fuerte: pasar del percentil 50 al 100 está asociado con un incremento de 10 puntos porcentuales en la probabilidad de terminar preparatoria, de 30% a 40%. Esta relación es similar si comparamos niños de la misma escuela.

Figura 5: Terminó Preparatoria



Nota: xx

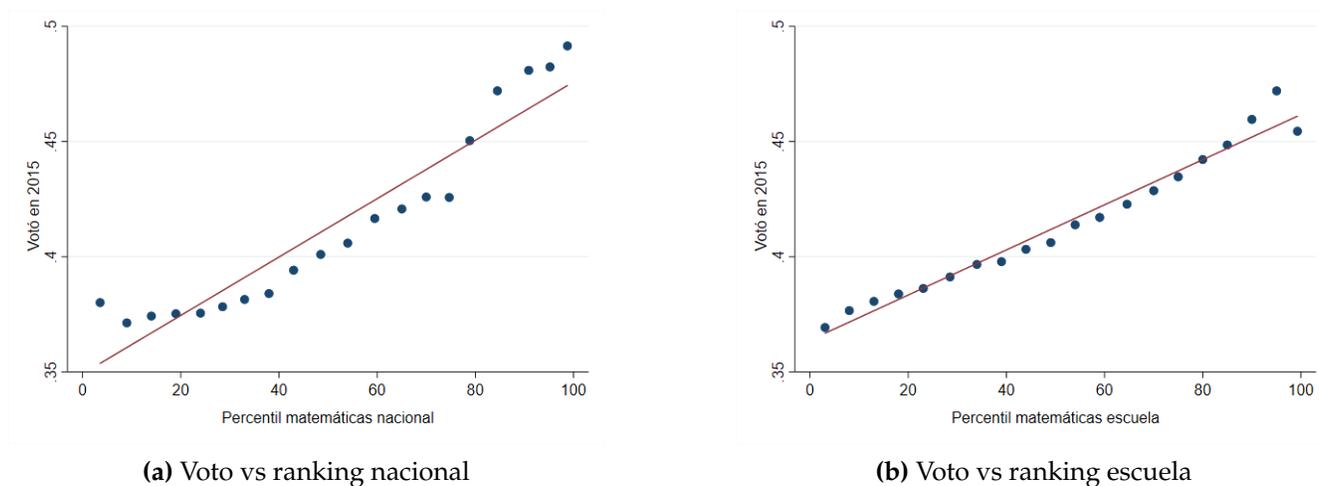
Entre las variables con externalidades sociales, consideramos dos: que el alumno vote cuando esté en edad de votar, y si el alumno ha sido condenado de haber cometido un crimen por un juez. Ambas

⁹La calidad de un alumno se define como el promedio de sus calificaciones estandarizadas $z_i^e := \sum_t \sum_g z_{iegt}$

variables son cruciales para el buen funcionamiento de una sociedad democrática.

No es claro que el desempeño en Enlace sea un buen predictor del comportamiento cívico de un estudiante cerca de una década posterior a tomar el examen. Sin embargo encontramos que si es así! La Figura 6 muestra que pasar del lugar 1 al lugar 100 en Enlace esta asociado a un aumento de 9 puntos porcentuales en la probabilidad de votar, pasando de 37 % a 46 %. Es una magnitud enorme en cuestión de participación en el voto.

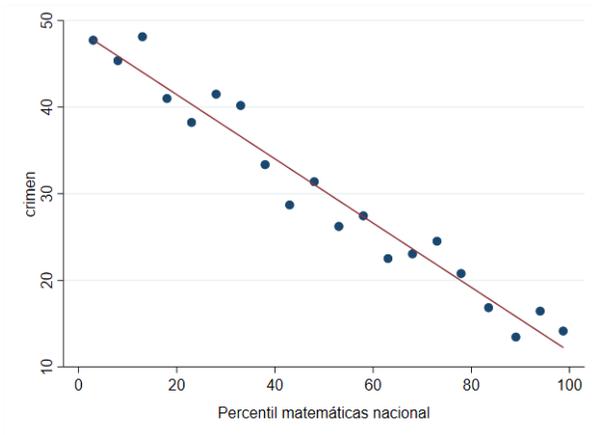
Figura 6: Voto en 2015



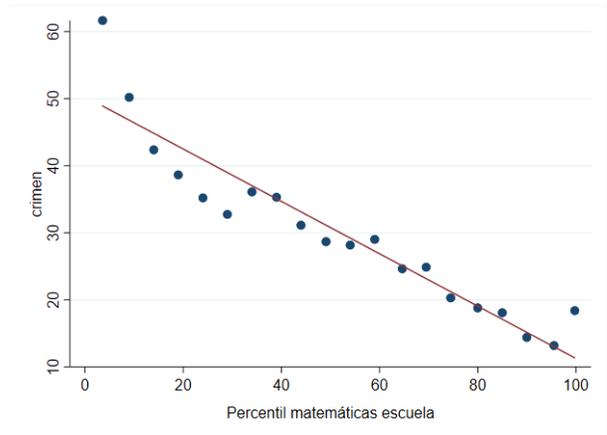
El poder judicial le reporta al INE si una persona ha sido declarada culpable en un proceso penal, debido a que en este caso la persona no puede votar. En este sentido el INE tiene posiblemente la base de datos mas confiable a nivel nacional en cuanto al numero de personas que son convictos criminales. El tipo de crimen incluye desde homicidios hasta robo. Conforme al acuerdo en la literatura, presentamos el nivel de crimen (eje Y de las gráficas) en términos de criminales por cada 100,000 votantes registrados.

La Figura muestra la relación entre la probabilidad de haber sido declarado criminal en algún momento antes del 2018 y el desempeño en la prueba Enlace. Puede observarse que el nivel es razonable en México, cerca de 30 personas por cada 100 mil electores. Pero lo impresionante es que aun dentro de una misma escuela, salir en los primeros lugares de Enlace está asociado a 5 veces menos crimen! De 50 por cada 100 mil a 10 por cada 100 mil.

Figura 7: Condenado por proceso penal (hasta 2018)



(a) Crimen vs ranking nacional



(b) Crimen vs ranking escuela

Nota: Crimen por 100,000

Enlace no es solamente un numerito en un papel, sino que mide comportamientos y resultados de vida de 6 a 12 años después de tomarla! El hecho de que pueda medir tan fielmente este tipo de indicadores quiere decir que no es ruido, quiere decir también que no es cierto que está completamente manipulada y que sus resultados sean infiables.

5. Conclusión

En vista de estos resultados, es indispensable recuperar una prueba como la enlace, que se le aplica al universo de estudiantes. Es la mejor forma de dar seguimiento a la calidad de las escuelas y al futuro de los niños y, por tanto, de México. Y puede usarse para identificar a estudiantes talentosos y dirigir becas especiales a niños talentosos y ayuda especial a niños o escuelas que están teniendo problemas.