



Armonización de la prueba censal ENLACE: Metodología de construcción de una base de datos con estructura de panel

XABER

Junio 2020

Resumen ejecutivo

La Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE) es una prueba censal estandarizada que se implementó en México entre 2006 y 2014, con participación de estudiantes entre tercero de primaria y hasta el último año del bachillerato. La prueba mide los niveles de logro en matemáticas, español y una materia alternativa. Esta nota describe el proceso de armonización de los datos para construir una base de datos con estructura de panel—siguiendo al mismo individuo en el tiempo—capaz de reconstruir las trayectorias educativas y niveles de aprendizaje para un alto número de individuos en México. Las trayectorias escolares que se infieren del panel de ENLACE son comparadas con las trayectorias que se obtienen a través de los datos administrativos (formato 911) para demostrar que el panel individual es confiable y representa, por lo tanto, una herramienta muy útil para el análisis y diseño de política educativa en México.



Contenido

1. Estructura y características de ENLACE.....	3
2. Estructura de los datos	6
3. Metodología para estructura de base panel con identificación de los alumnos en el tiempo	9
3.1. Etapas para integrar la base panel	11
4. Validación de la información	13
5. Conclusiones	15
Referencias	17
Anexo	18



1. Estructura y características de ENLACE

La prueba ENLACE nace en el ciclo escolar 2005-2006. El primer ejercicio se llevó a cabo en el fin de cursos de ese año escolar. La prueba se aplicaba anualmente y según el ciclo escolar varió la fecha específica de aplicación, la cual pudo ser desde abril hasta principios de junio. Los actores escolares eran informados sobre la evaluación y, en muchos casos, los padres de familia ayudaban a su aplicación. La prueba se basa en un instrumento psicométrico estandarizado y comparable en el tiempo, pero no estaba equiparada a través de los grados (ENLACE documento técnico, 2010). Los temas principales que evaluaba la prueba de ENLACE son matemáticas y español. Además, se aplicaba un tercer tema, el cual variaba año con año.

La información que ENLACE concentraba en una base de datos consistía en:

1) Información del alumno

- a. La Clave Única de Registro de Población (CURP) es un código alfanumérico único de identidad de 18 caracteres utilizado para identificar oficialmente tanto a residentes como a ciudadanos mexicanos de todo el país. Los primeros 16 dígitos de la CURP tienen como insumo el nombre completo, sexo, entidad y año de nacimiento, los dos últimos dígitos son asignados por el Gobierno de México para evitar duplicidades.
- b. Grado: Es el año escolar que cursa el alumno. En México, la educación básica abarca desde preescolar hasta secundaria. El nivel escolar de Primaria comprende los primeros seis grados. Mientras que el nivel escolar de Secundaria comprende de séptimo a noveno grado. La Educación Media Superior comprende de los grados décimo al doceavo, en algunos estados como Nuevo León, es nivel educativo es de dos años.
- c. Nombre: Esta variable se refiere al nombre del alumno, comprende el apellido materno, paterno y los nombres.
- d. Grupo: Esta variable define dentro de una escuela y un mismo grado en que salón el alumno fue asignado. Esto es, cada escuela en cada grado puede agrupar alumnos en salones, la clasificación a la agrupación por alumnos es el grupo escolar.

- 2) Clave del centro de Trabajo (CCT). Esta variable define un identificador único para cada escuela.



- 3) Estado: es la clasificación numérica de las 32 entidades de la República Mexicana.
- 4) Modalidad. La provisión pública o privada de la educación se clasifica con esta variable.
- 5) Resultados de la prueba
 - a. Nivel de logro por materia: insuficiente, elemental, bueno y excelente.
 - b. Calificación por materia. Por construcción, el puntaje de ENLACE tenía una media de 500 con una desviación estándar de 100.
 - c. Variable dicotómica si no contestó más del 50% de las preguntas (para algunas pruebas ENLACE).
 - d. Variable dicotómica si resultó positivo el algoritmo de copia. (para algunas pruebas ENLACE).

Cada año, se consolidó una base de datos con una estructura por observaciones y variables. La *Tabla 1* muestra el número de alumnos que fue evaluado y de los cuales se tiene información para cada ciclo escolar.

Tabla 1. Alumnos tomando ENLACE por nivel y ciclo

Ciclo	Primaria	Secundaria	Bachillerato
2005-2006	8,090,052	1,439,438	
2006-2007	8,538,047	1,610,619	
2007-2008	8,307,161	1,646,179	807,845
2008-2009	8,029,920	5,157,768	832,775
2009-2010	8,467,760	5,304,599	884,663
2010-2011	8,759,180	5,305,634	910,136
2011-2012	8,352,848	5,154,319	965,144
2012-2013	8,626,263	5,472,616	1,012,952
2013-2014			1,028,956

Fuente: Elaboración propia con datos administrativos de ENLACE 2005-2014.

Cada año, la aplicación de la prueba enfrentó retos importantes, lo que ha resultado en una limitante para el análisis del desempeño de cada uno de los alumnos a lo largo de su trayectoria académica. Uno



de los problemas principales es la falta de participación, la cual se le puede atribuir a las siguientes causas:

1. El alumno no participó
2. La escuela no participó en ENLACE
3. La entidad federativa no participó en ENLACE

Además, la estructura de ENLACE a través de los ciclos escolares no fue uniforme para todos los grados. El primer año de su aplicación, sólo participaron tercer, cuarto, quinto, y sexto grado de primaria, y su aplicación fue uniforme hasta el ciclo escolar 2012-2013. Para secundaria, los primeros tres años que la prueba fue aplicada, sólo tercero de secundaria participó. A partir del ciclo 2008-2009, primero y segundo de secundaria fueron evaluados. En educación media superior, ENLACE sólo se aplicaba en el último año del bachillerato. En la

Tabla 2 se muestran los ciclos escolares en los que se aplicó la prueba por nivel educativo.

Tabla 2. Niveles y grados en los que se aplicó la prueba ENLACE

Ciclo escolar	Primaria				Secundaria			EMS
	3	4	5	6	1	2	3	3
2005-2006	✓	✓	✓	✓			✓	
2006-2007	✓	✓	✓	✓			✓	
2007-2008	✓	✓	✓	✓			✓	✓
2008-2009	✓	✓	✓	✓	✓	✓	✓	✓
2009-2010	✓	✓	✓	✓	✓	✓	✓	✓
2010-2011	✓	✓	✓	✓	✓	✓	✓	✓
2011-2012	✓	✓	✓	✓	✓	✓	✓	✓
2012-2013	✓	✓	✓	✓	✓	✓	✓	✓
2013-2014								✓

Fuente: Elaboración propia con datos administrativos de ENLACE 2005-2014.

Cabe mencionar que de aquí en adelante nos referiremos a las pruebas ENLACE y al año correspondiente a cuando se aplicaron. Esto es, la prueba ENLACE del ciclo escolar 2005-2006, la llamaremos ENLACE 2006 dado que concluyó en el fin de cursos del ciclo escolar 2005-2006.



2. Estructura de los datos

El objetivo de esta sección es explorar de manera muy general la información de la ENLACE para conocer la cobertura de la prueba, así como la proporción de registros con la CURP, ya que será usada para identificar a los alumnos en el tiempo. Al comparar los registros de la prueba ENLACE con respecto a las cifras de matrícula reportadas por la SEP se observa que la cobertura de la prueba varió con el año de aplicación y el grado, pero siempre fue superior al 80%. La prueba ENLACE no pudo alcanzar la totalidad de la cobertura debido a varias causas, entre ellas a que en algunas entidades del sur de la República no fue posible aplicar la prueba o la participación fue muy baja.

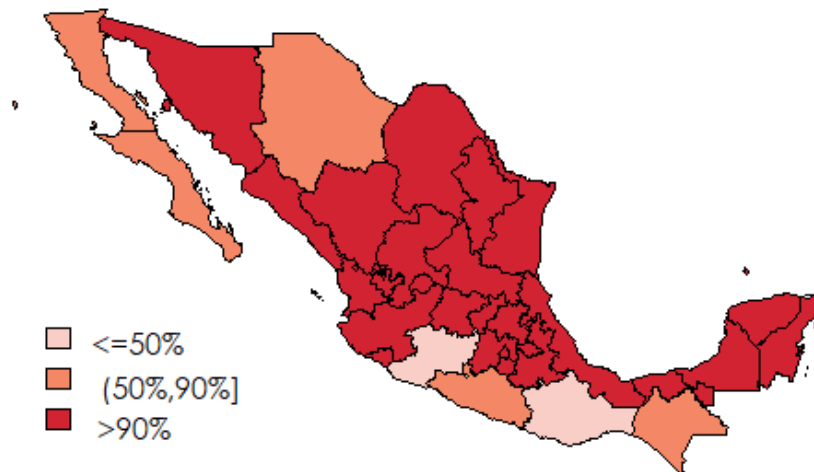
Respecto a la cobertura de ENLACE a nivel estatal se observa que en estados como Oaxaca la participación fue prácticamente nula, 1.2%, mientras que en Michoacán fue del 42% en 2006 (ver Anexo *A 1*). Durante el tiempo de aplicación de la prueba ENLACE la cobertura en educación básica en Oaxaca y Michoacán fue consistentemente baja,



Figura 1. En algunos años Guerrero y Chiapas presentan una cobertura inferior al 80%, en primaria y el comportamiento es similar para secundaria, Anexo *A 1*. El panorama es diferente para educación media superior, aunque Oaxaca y Michoacán presentan porcentajes más bajos que el resto de las entidades, registran una cobertura superior al 85% en algunos años. Por otro lado, la Ciudad de México presenta una cobertura de entre 40% y 48.6%, la causa principal se debe a que los bachilleratos de la Universidad Nacional Autónoma de México (UNAM) y los bachilleratos adscritos a la Ciudad de México no participaron en la prueba ENLACE.



Figura 1. Cobertura promedio de ENLACE en primaria, 2006-2013

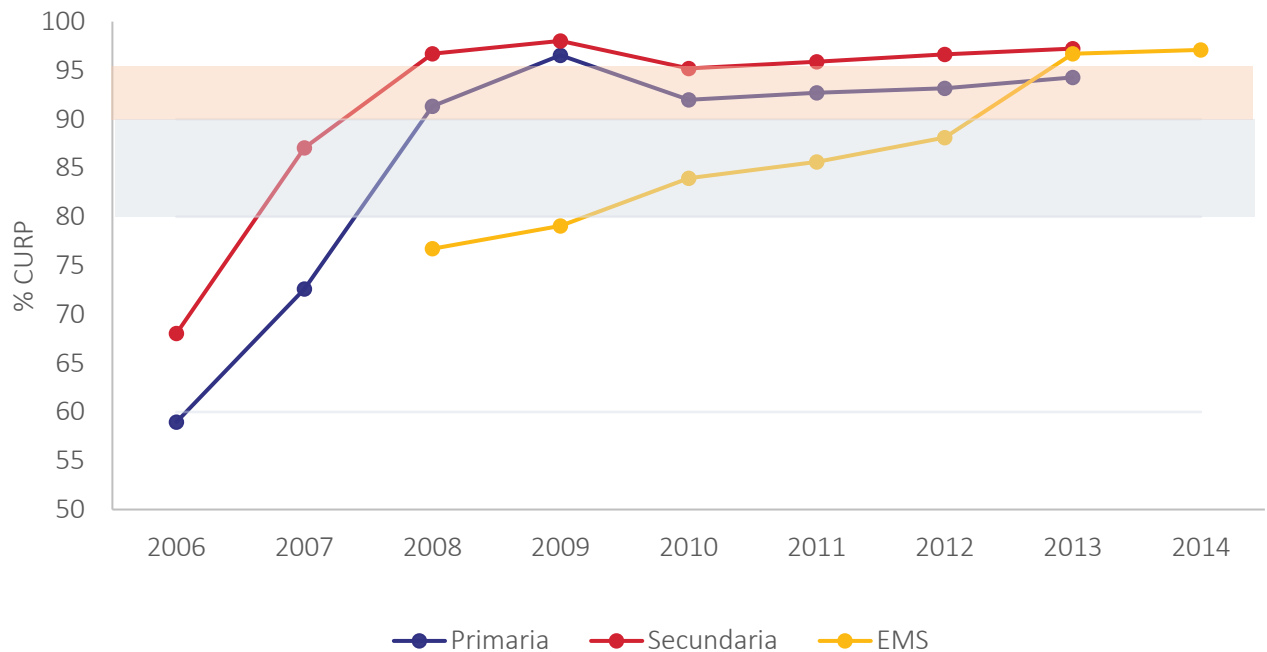


Las primeras aplicaciones de la prueba ENLACE presentan una cobertura relativamente baja con un alto número de observaciones con CURP mal capturada o con información poco confiable. En 2006 en primaria se observa que apenas el 59% de los registros cumplen con los criterios de construcción de la CURP y en secundaria el 68%. En educación media superior, el 76.7% de las observaciones en 2008 cuentan con la CURP y cumplen con los criterios de construcción. Como se puede observar en la *Gráfica 1*, a medida que pasa el tiempo, la calidad de la información¹ mejora, pues en los últimos años se logra que la proporción de registros con CURP se encuentren por encima del 90%. Por lo tanto, el número de datos faltantes se reduce, así como los registros que no cumplen con los criterios de la CURP (Anexo **Error! Reference source not found.**).

¹ Calidad de la información en este documento implica que se reduce el número de datos faltantes y la CURP cumple con los criterios para su construcción.



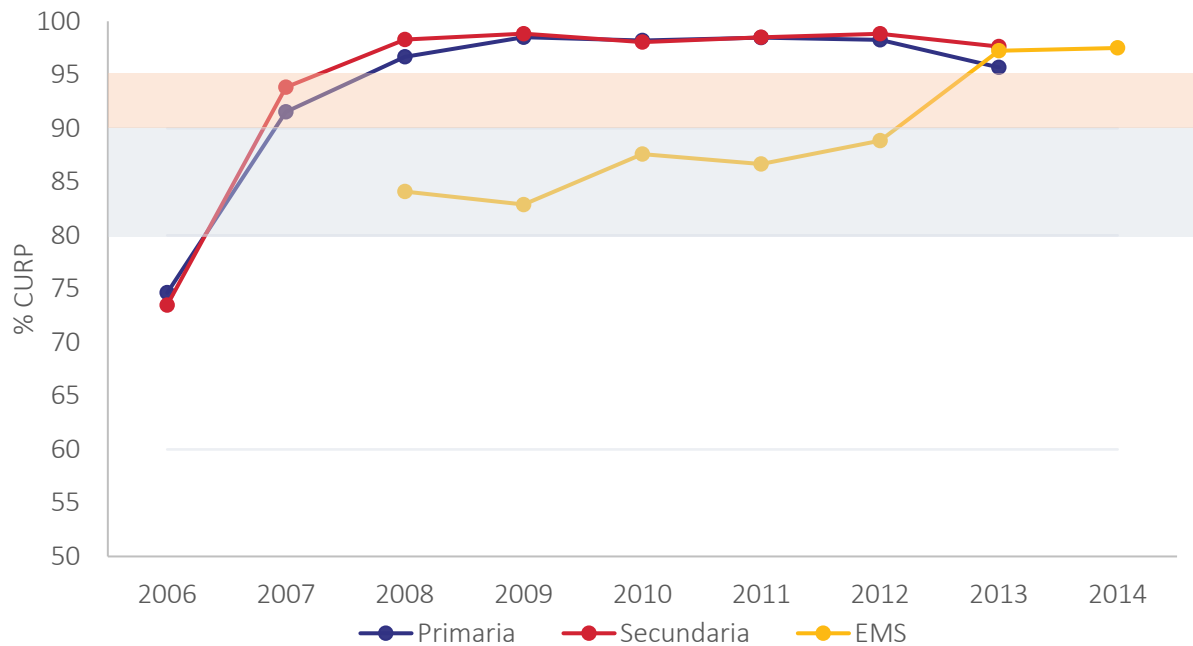
Gráfica 1. Porcentaje de observaciones que cumplen con los criterios de la construcción de la CURP



El usar la CURP a 16 caracteres (o también conocida como cuasi-CURP) en vez de 18 caracteres nos permite aumentar el número de observaciones que potencialmente podemos aparejar a través de distintos años. Por ejemplo, utilizando la cuasi-CURP se logra aumentar en más de 5% los registros en 2006 y 2007 en el nivel básico y a partir de 2007 el número de registros es superior al 90%. En educación media superior en el primer año de ENLACE, el número de observaciones con este identificador pasó de 76.7% en 2008 a 84.1%, Gráfica 3. Este enfoque produce más duplicidades, de acuerdo con los criterios establecidos para formar la CURP, las últimas dos posiciones corresponden a la homoclave, que es un candado para evitar duplicidades, sin embargo, al omitir esta homoclave no hubo un incremento sustancial respecto al número de duplicados, estas observaciones no representan ni el 1% del total de registros, ver Anexo A 3.



Gráfica 2. Porcentaje de observaciones que cumplen con los criterios de la construcción de la CURP a 16 caracteres



3. Metodología para estructura de base panel con identificación de los alumnos en el tiempo

El objetivo de esta sección es describir la metodología para agregar la información ENLACE de los estudiantes en el tiempo. Esto es, la estructura panel con forma $Y_{i,t,g}$ donde i define al individuo $i=1,\dots,N$; el año lo define $t=2006,\dots,2014$ y el grado se define $g=3,\dots,12$. La *Tabla 3* muestra 15 generaciones de alumnos que participaron durante todas las pruebas ENLACE. La base final será un panel no balanceado. Dicha tabla muestra la generación cero, la cuál es la primera generación en comenzar la escuela que se observa en ENLACE de bachillerato en 2007-2008. La generación uno, se observa en tercero de secundaria en el ciclo 2005-2006, esta generación terminaba ese año escolar secundaria, mientras la generación cero terminaba el primer año de bachillerato. Esta última se no observa en dicho ciclo dado que no se implementó ENLACE en primer grado de bachillerato (ni en ese ciclo ni durante la aplicación de ENLACE bachillerato). La generación 14 (la más joven) sólo



participo en ENLACE en tercero de primaria el ciclo escolar 2012-2013. Esta matriz muestra la trayectoria educativa por generación y su participación en ENLACE por orden secuencial de asistencia escolar.

Tabla 3. Generaciones de ENLACE

Ciclo	Primaria				Secundaria			Bachillerato		
	Tercero	Cuarto	Quinto	Sexto	Primero	Segundo	Tercero	Primero	Segundo	Tercero
2005-2006	7	6	5	4			1			
2006-2007	8	7	6	5			2			
2007-2008	9	8	7	6			3			0
2008-2009	10	9	8	7	6	5	4			1
2009-2010	11	10	9	8	7	6	5			2
2010-2011	12	11	10	9	8	7	6			3
2011-2012	13	12	11	10	9	8	7			4
2012-2013	14	13	12	11	10	9	8			5
2013-2014										6

Fuente: Elaboración propia

La *Tabla 4* muestra el número de generaciones por años de participación en la prueba. La primera y la última generación en participar sólo se observan un año. Cuatro generaciones estuvieron dos años, mientras que sólo dos generaciones participaron tres años. Podremos ver, tres generaciones durante toda la educación básica. Finalmente, el promedio simple de participación es de 3.8 años.

Tabla 4. Participación por generación

Generación	Años de participación
0	1
1	2
2	2
3	2
4	3
5	5
6	7
7	7
8	7
9	6



10	5
11	4
12	3
13	2
14	1

Fuente: Elaboración propia con datos de ENLACE 2005-2014.

3.1. Etapas para integrar la base panel

La variable más importante para realizar la agrupación de las ENLACE es la CURP del estudiante (para más información ver Anexo A 4).² El primer paso fue la organización de las bases y renombre para hacer más eficiente la consolidación de las pruebas ENLACE. El segundo paso fue limpieza de la CURP. El tercer paso se refiere a la depuración de la CURP si no cumple los criterios propuestos. El cuarto paso fue agrupar las ENLACE por nivel educativo y crear variables de identificación de año y nivel. Finalmente, se juntaron todas las bases para identificar los individuos a través del tiempo.³

3.1.1. Limpieza de CURP

En primer lugar, se llevó a cabo la limpieza a la CURP, la cual consistió en:

- Limpiar caracteres en conflicto. Es decir, eliminación o sustitución de caracteres especiales, signos de puntuación (&_';`*3/4'~%#@/?!.<>,#\$-*'|¥Ð}_ÀÁÒÇ'"'+-=:~·) o espacios en blanco. En algunos casos fue posible identificar algún símbolo, por ejemplo, el signo yen (¥) y sustituirlos por su caracter correspondiente “X”, en otros casos se eliminaron los caracteres, pues no fue posible establecer algún patrón.
- Homologar caracteres de palabras altisonantes/prohibidas conforme a las excepciones generales de la CURP (para más información ver el anexo). Por ejemplo, “CACA”.

² La Clave Única de Registro de Población (CURP) es un código alfanumérico único de identidad de 18 caracteres utilizado para identificar oficialmente tanto a residentes como ciudadanos mexicanos de todo el país.

³ Existe un ejercicio similar para estructurar la información de ENLACE en el tiempo, elaborada por Aguilar, S., Mejía, P., Seira, E. (2020). Para más detalles: <http://www.enriqueseira.com/uploads/3/1/5/9/31599787/reporte.pdf>



3.1.2. Depuración de la variable llave – CURP –

Para hacer una depuración de las CURP se retoma el método analizado por Aguilar, S. et al. (2020) y se propone una nueva metodología.

Metodología Xaber

1. Identificar número de alumnos que no tienen CURP
 - a. Esto es que su CURP no tenga ningún carácter alfanumérico menos de 16 caracteres
2. Cuantificar número de alumnos que no tienen información del nombre
3. Borrar alumnos sin nombre y sin CURP
4. Borrar a los que no tienen CURP
5. Revisar que las CURPs de la base de datos estén en el formato típico de 16 a 18 dígitos y con las descripciones normales que denotan los dígitos y su orden. Esto es posible con el comando 'regexm'.

Debilidades

1. Existen CURPS rescatables con sólo 16 caracteres que contengan la fecha de nacimiento y también esté el nombre para recrear la CURP
2. No verifica la CURP con la información de nombre y apellidos.

Tabla 5. Generaciones de ENLACE – metodología Xaber

Ciclo	Primaria				Secundaria			Bachillerato		
	Tercero	Cuarto	Quinto	Sexto	Primero	Segundo	Tercero	Primero	Segundo	Tercero
2005-2006	1,503,499	1,509,658	1,516,398	1,508,128			1,057,752			
2006-2007	1,962,796	1,936,456	1,959,501	1,956,065			1,511,436			
2007-2008	2,012,731	2,014,638	2,001,308	2,003,357			1,617,814			678,735
2008-2009	2,109,407	1,974,010	1,929,636	1,895,941	1,749,969	1,725,142	1,622,378			689,911
2009-2010	2,220,408	2,129,916	1,993,181	1,972,902	1,775,354	1,762,516	1,663,213			773,856
2010-2011	2,265,290	2,227,970	2,134,398	1,995,382	1,830,345	1,740,374	1,655,102			788,574
2011-2012	2,006,289	2,121,456	2,077,725	2,001,348	1,799,187	1,707,382	1,586,801			853,221
2012-2013	1,982,373	2,024,531	2,135,347	2,111,261	1,941,643	1,751,870	1,650,124			984,991
2013-2014										1,003,176

Fuente: elaboración propia con datos administrativos de ENLACE 2005-2014.



3.1.3. Limitaciones

La metodología no tiene 100% certeza de que se cumplan completamente con las reglas de la CURP. Por ejemplo, dicho procedimiento verifica que las posiciones 5 y la 10 sean números, pero no verifican el orden. No es lo mismo 991030 que 301099; la CURP acepta el primero no el segundo. De la misma manera, no lo hacen con los nombres.

3.1.4. Consolidación de bases de datos

El ejercicio de consolidación de las bases de datos identifica dos armonizaciones, el panel_18 (exacto) es armonización del panel identificado con las CURPS de 18 caracteres, mientras que el panel_16 se construyó con base en la CURP sólo con 16 caracteres. Se agruparon las ENLACES por nivel educativo y se crearon variables de identificación de año y nivel. Finalmente, se reunieron todas las bases para identificar los individuos a través del tiempo. En la *Tabla 6* se muestran los resultados.

Tabla 6. Resultados para todas las generaciones

Base	Número de CURP único	Número de observaciones	Personas con los 4 años de primaria completos	Personas con los 3 años de secundaria completos	Años alumno promedio
panel_16	31,998,712	98,867,887	7,095,737	4,243,423	3.1
panel_18	31,047,264	93,202,004	6,075,957	4,132,639	3

Fuente: elaboración propia con datos administrativos de ENLACE 2005-2014.

4. Validación de la información

¿Qué tan confiable es la información contenida en el panel de ENLACE? Para contestar esta pregunta se llevaron a cabo dos ejercicios, el primero consiste en comparar las tasas de supervivencia de la población mediante la información administrativa de la SEP (el formato 911) y el segundo compara la proporción de alumnos en el panel con otra base de datos longitudinal creada por el Instituto Nacional para la Evaluación de la Educación, INEE, con datos del Registro Nacional de Alumnos (RNA).



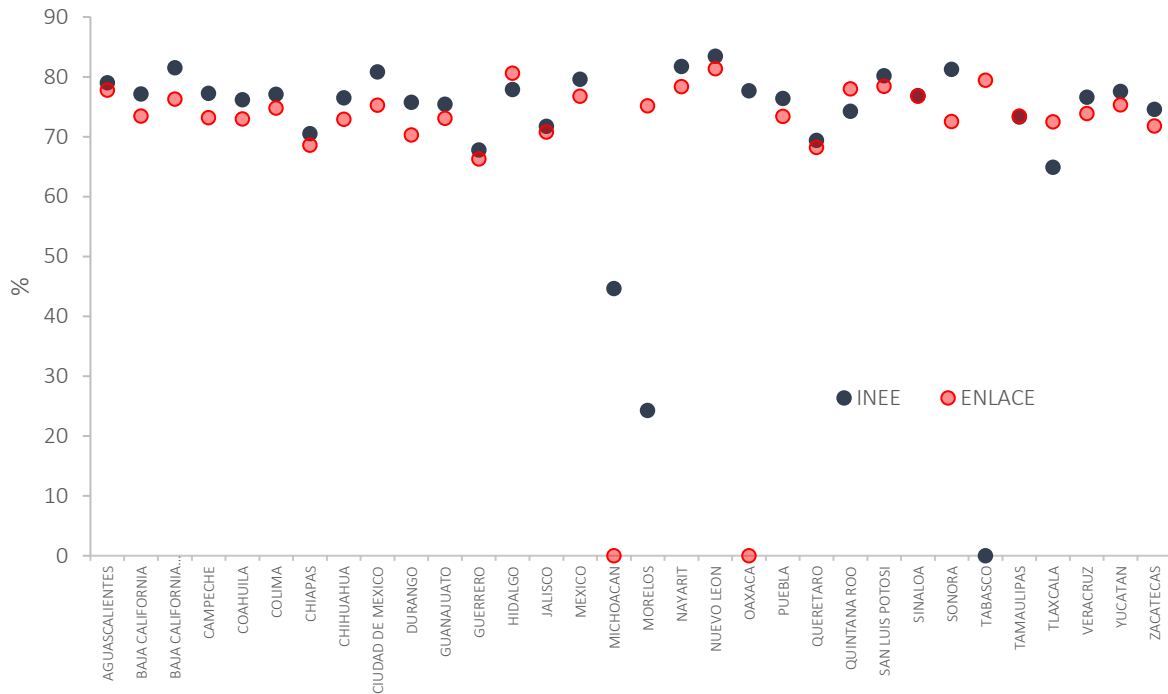
Tal como lo hace de Hoyos, Estrada y Vargas (2018), para cuantificar en qué medida los factores como la absorción, deserción escolar o a la reprobación, desgastan a una generación, tomamos como medida de mayor confianza los datos administrativos del censo anual a escuelas, Formato 911. Al considerar la trayectoria escolar del total de alumnos en sexto de primaria en 2007-2008, se esperaba que el 76 por ciento llegará al noveno grado en 2010 y el 41 por ciento llegará al último grado de educación media superior en 2014. De acuerdo con este primer ejercicio, el panel ENLACE tiene una tasa de supervivencia de 5 puntos porcentuales menos en la transición a noveno grado y de 9 puntos porcentuales menos en la transición a educación media superior.

Con el propósito de construir indicadores de cohorte longitudinal, el INEE realizó el seguimiento de alumnos con datos del RNA,⁴ el cual incluye datos personales de los alumnos de primaria y secundaria para inicio y fin de cursos de los ciclos de 2007-2008 a 2010-2011 (Videgain, K. 2015). Mediante este panel es posible comparar, la proporción de alumnos en sexto de primaria en el ciclo 2007-2008 que transitaron a novena grado en el ciclo 2010-2011, cohorte 6, en este caso la proporción es del 74 por ciento muy similar a la tasa obtenida mediante datos administrativos y a la tasa obtenida a través de los datos de ENLACE que fue de 71.6% o 2.4 puntos porcentuales menos. La *Gráfica 3* muestra el porcentaje de alumnos de tercero de secundaria que se pudo aparejar con sexto, con respecto al total de alumnos que concluyeron la primaria. Tanto en el panel del INEE como en ENLACE se observa que los patrones son similares, además en ambas bases se observan algunos estados en los que no fue posible contar con información disponible (como Michoacán y Tabasco en el caso de ENLACE y Tabasco en el caso de la base de datos del INEE).

⁴ <https://www.inee.edu.mx/wp-content/uploads/2019/08/P3B103-1.pdf>



Gráfica 3. Porcentaje de alumnos tomando como base sexto de primaria 2007, que se observan en tercero de secundaria en 2010



5. Conclusiones

La base de ENLACE provee una oportunidad única para seguir el desempeño educativo y monitorear los aprendizajes de varias generaciones de estudiantes en México. La metodología empleada para aparejar la información a nivel individual y construir una base de datos longitudinal, consiste en limpiar y depurar la variable de identificación del alumno de una manera sencilla a través de la Clave Única de Registro Poblacional (CURP). Al unir múltiples evaluaciones en diferentes grados durante el periodo de 2006 a 2014 es posible crear una base de datos longitudinal en donde observamos al mismo estudiante en varios puntos durante su trayectoria educativa. Los cuestionarios de contexto de la prueba también permiten identificar los factores asociados al desempeño académico de los alumnos, incluyendo las características de las instituciones educativas en las que participan.



Es importante también resaltar las limitaciones de la base de datos longitudinal. Esta base de datos no resulta en un panel balanceado, ya que existen alumnos que no participaron en todas las aplicaciones de la prueba ENLACE. Adicionalmente, en algunos estados del país la cobertura de la prueba ENLACE fue muy baja, particularmente en sus primeros años de aplicación, lo que limita la posibilidad de analizar las trayectorias educativas de todos los alumnos en estos estados.

Evaluar la confiabilidad del panel creado es complicado ya que existen pocos ejercicios que sigan la trayectoria de los alumnos en México. Sin embargo, mediante el panel longitudinal creado por el INEE y los datos administrativos del Formato 911, es posible comparar la proporción de alumnos que concluyen la educación básica y comparar este estadístico con la transición que arroja el panel ENLACE. Esta comparación refleja patrones de transición similares entre las tres fuentes de información, lo que sugiere que el panel de ENLACE es una fuente de información confiable para analizar y monitorear las trayectorias de los alumnos en 30 de los 32 estados donde la tasa de participación en ENLACE superó consistentemente el 75% de los alumnos inscritos (siendo mayor a 90% en la mayoría de los años / estados).

A pesar de sus limitaciones—como cualquier otra base de datos longitudinal creada a partir de datos administrativos—el panel de ENLACE tiene un enorme potencial para contribuir al análisis de los determinantes de las trayectorias educativas y la efectividad de las políticas públicas para mejorar esas trayectorias. Sin información sobre la importancia que juegan las condiciones de inicio en las trayectorias educativas futuras, es imposible diseñar políticas educativas capaz de reducir las desigualdades educativas y mejorar, con ello, las oportunidades de vida de las niñas, niños y jóvenes menos favorecidos.



Referencias

Aguilar, S., Mejía, P., Seira, E. (2020). Algunos Estadísticos Sobre la Educación Primaria en México. Instituto Tecnológico Autónomo de México. <http://www.enriqueseira.com/uploads/3/1/5/9/31599787/reporte.pdf>

de Hoyos, Rafael & Estrada, Ricardo & Vargas, Maria Jose, 2018. "Predicting individual wellbeing through test scores: evidence from a national assessment in Mexico," *Policy Research Working Paper Series 8459*, The World Bank.

Videgain, K. (2015). Análisis longitudinal del Registro Nacional de Alumnos sobre trayectorias. Educativas. Instituto Nacional para la Evaluación de la Educación. <https://www.inee.edu.mx/wp-content/uploads/2019/08/P3B103-1.pdf>



Anexo

A 1. Cobertura de ENLACE por nivel educativo

Tabla 7. Cobertura de ENLACE en primaria

Entidad	2006	2007	2008	2009	2010	2011	2012	2013
Aguascalientes	96.0	97.2	95.1	94.4	97.7	98.2	98.1	97.8
Baja California	82.4	90.4	91.7	89.6	89.6	91.5	91.6	91.1
Baja California Sur	73.4	94.0	84.3	92.4	92.8	93.1	92.5	92.2
Campeche	92.0	89.0	83.2	92.1	93.1	92.6	92.4	92.6
Coahuila	93.3	93.9	93.8	90.7	95.4	93.6	92.8	94.8
Colima	86.9	95.5	94.6	93.8	94.4	93.7	94.9	95.2
Chiapas	89.3	84.7	84.8	90.8	91.0	94.4	62.4	75.9
Chihuahua	80.3	88.8	88.9	89.7	90.5	89.6	89.7	90.0
Ciudad de México	89.6	96.1	97.0	97.8	97.3	96.8	94.3	95.7
Durango	90.7	95.0	86.9	95.6	94.1	95.7	94.9	94.6
Guanajuato	92.4	95.9	95.6	95.5	95.9	95.5	94.8	94.8
Guerrero	88.2	86.8	90.8	56.5	78.3	84.6	12.2	73.0
Hidalgo	95.1	96.9	95.6	91.9	96.2	94.1	94.7	96.1
Jalisco	93.9	94.7	95.3	95.2	94.6	95.3	94.9	94.5
Estado de México	92.8	95.1	95.2	91.0	95.4	94.8	95.0	94.7
Michoacán	42.1	49.1	0.8	33.7	37.9	46.7	25.8	29.2
Morelos	89.8	95.5	91.8	81.2	93.1	93.0	91.4	93.4
Nayarit	92.3	92.0	92.2	92.2	92.8	92.1	92.1	91.8
Nuevo León	89.5	92.9	94.2	93.9	95.0	94.7	95.7	94.9
Oaxaca	1.2	68.1	77.8	10.0	1.2	1.2	1.4	1.3
Puebla	86.8	87.4	90.5	78.2	93.9	95.9	94.6	96.3
Querétaro	95.7	96.8	96.2	96.3	96.9	96.2	94.9	95.4
Quintana Roo	91.3	94.3	93.0	90.2	92.1	93.9	92.6	87.3
San Luis Potosí	95.2	95.6	90.1	94.7	95.4	96.7	96.0	97.2
Sinaloa	90.9	93.0	93.5	90.9	92.0	94.1	94.0	94.3
Sonora	88.0	93.1	94.2	89.0	91.6	87.1	92.7	92.4
Tabasco	86.3	94.1	92.9	94.0	94.6	93.3	93.0	94.1
Tamaulipas	91.0	93.9	95.4	94.4	94.1	95.7	95.1	95.6
Tlaxcala	88.9	95.4	96.4	93.1	87.4	93.8	90.2	96.5
Veracruz	92.1	92.9	95.0	95.3	95.5	95.0	96.6	96.5
Yucatán	95.3	95.7	95.7	93.9	95.9	95.8	96.1	94.9
Zacatecas	89.4	92.9	93.5	92.1	94.6	94.7	89.1	94.7

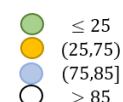




Tabla 8. Cobertura de Enlace en secundaria

Entidad	2006	2007	2008	2009	2010	2011	2012	2013
Aguascalientes	92.6	93.3	93.3	92.3	94.6	94.9	94.5	94.1
Baja California	80.9	89.4	92.7	89.1	86.4	90.1	89.8	89.7
Baja California Sur	90.7	93.3	83.8	90.5	91.5	91.0	90.7	91.3
Campeche	89.2	92.5	91.3	90.5	91.2	90.7	90.2	89.5
Coahuila	90.2	92.3	92.0	89.2	90.4	89.3	88.9	90.5
Colima	92.1	94.0	92.8	91.5	92.0	90.9	91.5	92.1
Chiapas	89.1	86.8	86.2	71.6	91.2	92.1	86.6	91.4
Chihuahua	50.2	91.0	92.4	90.8	91.7	90.7	90.8	90.8
Ciudad de México	81.1	91.2	92.1	91.8	92.9	91.7	91.0	91.5
Durango	86.2	92.3	91.6	89.4	89.8	92.1	90.2	90.1
Guanajuato	93.2	94.9	93.7	92.8	93.4	92.7	92.3	92.1
Guerrero	85.2	87.8	91.7	70.4	86.4	88.9	19.2	82.6
Hidalgo	94.1	95.5	94.7	93.2	94.8	93.7	94.1	94.7
Jalisco	91.8	92.8	93.1	90.6	92.5	92.3	92.2	92.5
Estado de México	89.2	93.5	93.7	91.7	92.9	90.8	91.9	92.7
Michoacán	2.8	4.2	0.2	31.2	35.0	42.8	25.3	26.8
Morelos	91.4	92.6	92.2	80.9	90.4	90.5	91.1	90.5
Nayarit	90.9	91.7	92.1	89.9	91.5	90.2	90.7	90.4
Nuevo León	61.6	93.1	93.4	92.0	93.5	93.0	93.5	93.9
Oaxaca	0.2	76.8	79.8	12.0	0.7	0.7	1.3	1.4
Puebla	89.9	85.4	93.1	81.8	92.5	93.1	92.9	93.6
Querétaro	86.9	94.8	94.7	93.4	94.3	93.0	92.3	92.6
Quintana Roo	88.4	90.7	87.9	87.8	91.2	92.2	92.1	89.3
San Luis Potosí	92.9	92.7	94.1	90.4	93.9	93.8	93.8	95.2
Sinaloa	91.0	93.5	93.1	89.0	89.1	92.5	92.0	92.1
Sonora	86.7	91.6	92.8	90.3	91.2	84.1	89.6	88.4
Tabasco	87.4	91.9	90.3	90.5	92.8	91.9	90.9	91.7
Tamaulipas	89.4	91.8	93.0	91.5	92.7	93.6	93.2	93.6
Tlaxcala	88.7	89.6	93.9	93.1	93.6	93.0	90.9	94.7
Veracruz	91.0	92.6	94.0	93.0	93.4	91.9	93.9	94.0
Yucatán	92.1	93.1	93.1	91.1	93.2	92.3	92.5	91.5
Zacatecas	88.3	90.9	92.0	90.6	92.8	92.2	92.2	92.3

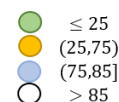
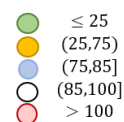




Tabla 9. Cobertura de Enlace en educación media superior

Entidad	2008	2009	2010	2011	2012	2013	2014
Aguascalientes	90.1	93.1	94.1	95.7	94.1	94.7	93.7
Baja California	78.8	81.4	82.3	84.8	85.6	85.8	86.2
Baja California Sur	79.2	77.8	84.7	90.0	89.9	89.4	79.3
Campeche	82.3	89.0	87.0	86.0	88.5	88.3	84.3
Coahuila	125.1	117.6	123.5	116.8	120.6	125.5	119.5
Colima	93.3	93.3	93.6	90.8	92.8	93.3	96.5
Chiapas	71.3	80.7	84.9	77.7	78.3	78.3	81.3
Chihuahua	87.5	84.5	85.6	85.5	86.5	89.8	81.6
Ciudad de México	48.6	47.5	46.3	40.8	45.1	47.9	48.3
Durango	71.8	75.6	83.2	85.6	84.9	83.7	81.5
Guanajuato	91.7	90.3	88.9	90.2	92.9	91.3	91.4
Guerrero	85.2	86.0	84.2	83.8	84.5	86.7	87.8
Hidalgo	82.6	86.8	86.1	87.1	88.9	87.8	86.1
Jalisco	91.2	88.1	93.1	84.2	97.7	97.9	93.0
Estado de México	83.8	80.2	84.3	84.3	84.3	83.3	82.5
Michoacán	59.3	72.9	85.4	84.5	86.8	81.3	77.7
Morelos	87.2	81.5	83.9	84.2	83.5	83.6	84.3
Nayarit	84.8	87.5	91.2	90.6	88.8	87.2	83.3
Nuevo León	175.6	185.6	182.8	176.7	190.1	189.8	197.2
Oaxaca	70.4	67.7	85.4	85.9	90.6	89.7	86.1
Puebla	86.5	87.3	88.7	89.9	90.0	90.5	90.3
Querétaro	90.7	90.2	91.4	91.7	93.2	92.3	91.3
Quintana Roo	79.6	80.6	91.3	88.8	92.4	94.1	86.7
San Luis Potosí	126.4	127.6	125.8	126.6	125.8	126.7	121.4
Sinaloa	86.3	89.5	84.2	89.7	89.3	91.5	90.0
Sonora	83.8	95.6	97.7	90.5	92.4	89.2	89.3
Tabasco	79.0	82.7	88.3	87.7	89.8	87.7	88.8
Tamaulipas	102.6	103.5	88.1	94.7	96.1	95.2	95.6
Tlaxcala	86.8	88.7	89.0	79.0	86.7	91.9	88.4
Veracruz	82.3	84.7	87.8	88.6	89.8	89.7	79.6
Yucatán	90.3	90.0	89.3	88.9	88.0	90.8	82.6
Zacatecas	78.2	85.2	90.6	87.1	89.6	89.7	85.4





A 2. Estructura de las bases de datos por nivel educativo a 18 caracteres

Tabla 10. Estructura de los datos en primaria, CURP a 18 caracteres (porcentaje)

Año	CURP cumple con criterios	Duplicados	Faltantes	No cumplen con criterios	Total
2006	59.0	0.1	21.3	19.7	100
2007	72.6	0.1	6.5	20.8	100
2008	91.3	0.1	2.5	6.0	100
2009	96.6	0.1	1.3	2.0	100
2010	92.0	0.1	1.6	6.3	100
2011	92.7	0.1	1.3	5.9	100
2012	93.2	0.1	1.6	5.2	100
2013	94.3	0.0	4.3	1.4	100

Tabla 11. Estructura de datos en secundaria, CURP a 18 caracteres (porcentaje)

Año	CURP cumple con criterios	Duplicados	Faltantes	No cumplen con criterios	Total
2006	68.0	0.1	19.3	12.6	100
2007	87.1	0.0	4.0	8.9	100
2008	96.7	0.1	1.5	1.7	100
2009	98.0	0.0	1.1	0.9	100
2010	95.2	0.1	1.8	2.9	100
2011	95.9	0.0	1.4	2.7	100
2012	96.6	0.0	1.1	2.3	100
2013	97.2	0.0	2.3	0.4	100

Tabla 12. Estructura de datos en educación media superior, CURP a 18 caracteres (porcentaje)

Año	CURP cumple con criterios	Duplicados	Faltantes	No cumplen con criterios	Total
2008	76.7	0.4	9.9	13.0	100
2009	79.1	0.2	14.2	6.6	100
2010	84.0	0.2	10.8	5.1	100
2011	85.6	0.2	8.7	5.5	100
2012	88.1	0.2	7.1	4.6	100
2013	96.7	0.2	1.3	1.7	100
2014	97.1	0.1	1.6	1.1	100



A 3. Estructura de las bases de datos por nivel educativo a 18 caracteres

Tabla 13. Estructura de los datos en primaria, CURP a 16 caracteres (porcentaje)

Año	CURP cumple con criterios	Duplicados	Faltantes	No cumplen con criterios	Total
2006	74.5	0.3	21.3	3.9	100
2007	91.3	0.4	6.5	1.9	100
2008	96.4	0.4	2.5	0.6	100
2009	98.2	0.4	1.3	0.1	100
2010	97.9	0.4	1.6	0.1	100
2011	98.2	0.4	1.3	0.1	100
2012	98.0	0.4	1.6	0.1	100
2013	95.4	0.3	4.3	0.0	100

Tabla 14. Estructura de los datos en primaria, CURP a 16 caracteres (porcentaje)

Año	CURP cumple con criterios	Duplicados	Faltantes	No cumplen con criterios	Total
2006	73.4	0.2	22.0	4.5	100
2007	93.7	0.2	4.0	2.1	100
2008	98.1	0.2	1.5	0.2	100
2009	98.6	0.3	1.1	0.0	100
2010	97.8	0.3	1.8	0.1	100
2011	98.2	0.3	1.4	0.1	100
2012	98.6	0.3	1.1	0.1	100
2013	97.4	0.3	2.3	0.0	100

Tabla 15. Estructura de los datos en educación media superior, CURP a 16 caracteres (porcentaje)

Año	CURP cumple con criterios	Duplicados	Faltantes	No cumplen con criterios	Total
2008	83.9	0.6	9.9	5.6	100
2009	82.8	0.3	14.2	2.8	100
2010	87.4	0.3	10.8	1.5	100
2011	86.5	0.3	8.7	4.4	100
2012	88.7	0.3	7.1	3.9	100
2013	97.1	0.3	1.3	1.2	100
2014	97.3	0.3	1.6	0.7	100



A 4. Metodología para construcción de la CURP

- Primera letra y la primera vocal del primer apellido.
- Primera letra del segundo apellido.
- Primera letra del nombre de pila: se tomará en cuenta el primer nombre. ("exceptuando los nombres compuestos cuando estos se antepongan los nombres de MARÍA y JOSÉ") Ejemplo: Juan Francisco se tomaría la letra J. Ejemplo 2: Si se llama José Eduardo tomará E.
- Fecha de nacimiento sin espacios en orden de año, mes y día. Ejemplo: 960917 (1996, septiembre 17).
- Letra del sexo (H o M).
- Dos letras correspondientes a la entidad de nacimiento (en caso de haber nacido fuera del país, se marca como NE, "Nacido en el Extranjero"; véase el Catálogo de Claves de Entidades Federativas).
- Primera consonante interna () del primer apellido.
- Primera consonante interna () del segundo apellido.
- Primera consonante interna () del nombre.
- Dígito del 0-9 para fechas de nacimiento hasta el año 1999 y A-Z para fechas de nacimiento a partir del 2000.
- Dígito para evitar duplicaciones.

Por ejemplo, si una persona se llamara Luis Raúl Bello Mena, de sexo masculino y hubiera nacido el 13 de marzo de 1992 en el Estado de México, su CURP podría ser BEML920313HMCLNS09. Habitualmente, los primeros dos caracteres de la clave son las dos primeras letras del primer apellido, pero en caso de que un primer apellido tenga una consonante como segunda letra, se asignará entonces la primera vocal que se encuentre en el apellido. Por ejemplo, si una persona se llamara Silverio Ortiz Castellanos, sus primeros cuatro caracteres de la CURP serían: OICS, ya que la segunda letra de "Ortiz" es consonante.

Tampoco la letra Ñ se asigna en la CURP. Un claro ejemplo es el apellido Núñez. En casos como éste es necesario reemplazar, en la clave, la letra "Ñ" por el carácter "X".

Cuando una persona tenga dos nombres y su primer nombre sea María, en el caso de mujeres, o José, en el caso de hombres, el cuarto carácter se tomará de la primera letra del segundo nombre, en vez del



primero. Esto debido a que los nombres "María" y "José" son muy comunes y generarían muchos duplicados. Por ejemplo: si la persona se llama María Fernanda Escamilla Arroyo, los primeros cuatro caracteres serán EAAF (María no cuenta para formar el cuarto carácter).

Otra excepción en la cual sería necesario utilizar "X" sería el caso de personas que no cuentan con segundo apellido, o algunos extranjeros en cuyos países el segundo apellido es inexistente. Por ejemplo, un hombre proveniente de Irlanda cuyos nombres son Brian Donall y su único apellido es Plunkett nació el 21 de octubre de 1957 y, como no nació en ninguna entidad mexicana, las dos claves correspondientes serán "NE" (Nacido en el Extranjero), y su CURP podría ser: PUXB571021HNELXR00. Se usarían, aquí, la "X" por no tener segundo apellido, ni mucho menos primera consonante del mismo, "NE" porque no nació en México y "U" porque la segunda letra de su apellido, "Plunkett", es una consonante.

Además de las excepciones por motivo de ser ofensivo en el idioma español ejemplo: Pulido Torres Amanda su curp sería ofensiva lo cual se cambia la segunda letra por una X quedando así: PXTA

En la estructura de la CURP (posiciones 1-4) que en ocasiones forma una palabra cuya pronunciación se considera ofensiva para los patrones socialmente establecidos, en cuyo caso la letra de la segunda posición se sustituye por una 'X'.