

INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO



**ANALFABETISMO Y EDUCACIÓN INDÍGENA EN
MÉXICO:
EVALUACIÓN DE IMPACTO DEL PROGRAMA DE
INCLUSIÓN Y EQUIDAD EDUCATIVA**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE
MAESTRO EN ECONOMÍA APLICADA
PRESENTA**

ALEJANDRO GRIMALDI FERREIRA

ASESOR: DR. ARTURO ALBERTO AGUILAR ESTEVA

CIUDAD DE MÉXICO

2021

“Con fundamento en los artículos 21 y 27 de la Ley Federal del Derecho de Autor y como titular de los derechos moral y patrimonial de la obra titulada **"ANALFABETISMO Y EDUCACIÓN INDÍGENA EN MÉXICO: EVALUACIÓN DE IMPACTO DEL PROGRAMA DE INCLUSIÓN Y EQUIDAD EDUCATIVA"**, otorgo de manera gratuita y permanente al Instituto Tecnológico Autónomo de México y a la Biblioteca Raúl Bailleres Jr. autorización para que fijen la obra en cualquier medio, incluido el electrónico y la divulgación entre sus usuarios, profesores, estudiantes o terceras personas, sin que pueda percibir por la divulgación una contraprestación”.

ALEJANDRO GRIMALDI FERREIRA

लोकाः समस्ताः सुखिनो भवन्तु

Lokah Samastah Sukhino Bhavantu

“Que todos los seres en todas las dimensiones sean prósperos, libres y felices y que mis pensamientos, palabras y acciones contribuyan a ello”.

Agradecimientos

Me siento infinitamente agradecido con la vida por haberme dado la oportunidad de llegar hasta este punto: amo a mi familia, amo a mis amigos, amo mi carrera y amo mi trabajo. Gracias a todas y todos mis amigos que me acompañaron en este y en tantos procesos; saben quienes son y lo que significa cada uno de ustedes para mí.

Gracias a Rocío, mi maestra de yoga, a Alicia, mi psicóloga, y a Carina, mi nutrióloga, porque también fueron parte importante de este proceso. No solo como maestras, sino como guías de un camino de aprendizaje que me ha llevado a la felicidad de espíritu. Que este sea un recordatorio para no olvidar que no valemos por lo que hacemos ni lo que tenemos, sino por lo que somos.

Gracias a todas y todos mis maestros de esta etapa: Claudia Aburto, Arturo Aguilar, Isidoro García, Rafa de Hoyos, Sumiko Hayasaka, Manett Vargas, Emilio Gutiérrez y Shaun McRae. Los admiro mucho y les agradezco por su pasión, humanismo y compromiso para la enseñanza. En especial, le agradezco a Claudia por haber confiado en mí y por haberme alentado a tomar esta experiencia. Gracias por tu gran capacidad docente, pero, sobre todo, por tu preocupación genuina hacia los alumnos.

También, quisiera agradecerle a Arturo no solo por sus excelentes cursos de econometría, sino por orientarme en la construcción de

esta tesis. Siempre encuentras más interrogantes que abren la investigación hacia nuevas rutas de análisis para profundizar en lo que hay detrás de cualquier argumento e inferencia. Gracias, además, por haberme dado la oportunidad de ayudarte con el curso de Microeconometría Aplicada; aprendí y disfruté mucho la experiencia.

Gracias a mis sinodales por sus valiosos comentarios: Claudia, Manett y Mauricio Romero. Les agradezco por haber leído mi trabajo y por su apertura para dialogarlo y mejorarlo.

Por último, gracias a mi familia que siempre me ha apoyado en todos mis proyectos. Rubén, gracias por caminar junto a mí siempre; ciertamente, creo que no hubiera podido soñar con un mejor hermano, pero la vida me sonrió contigo. Mamá, gracias por ser siempre una luz que guía mi camino, eres un ejemplo de amor y compasión, esas fuerzas que realmente cambian al mundo. Papá, gracias por tu pasión y compromiso, eres mi ejemplo de lucha, trabajo y solidaridad; te debo lo que sé y lo que he logrado, aunque nunca se acerque a la manera en que tú salvas vidas cada día.

Escribí esta tesis con el deseo de contribuir a la investigación empírica sobre el impacto de la educación en México, porque la educación y los derechos, en general, cambian vidas y construyen sociedades más libres, justas y prósperas. Espero que, algún día, todas y todos en el mundo tengan acceso a los derechos sociales con dignidad y justicia.

Contents

Agradecimientos	1
1. Introducción	4
2. Marco Teórico	10
2.1. Alfabetización y diversidad lingüística	10
2.2. Desarrollo de habilidades cognitivas y educación básica	14
2.3. Un programa para la educación indígena	16
3. Presentación de Variables y Análisis Exploratorio	23
3.1. Analfabetismo, población indígena y características sociodemográficas	23
3.2. Cobertura del Programa de Inclusión y Equidad Educativa	32
4. Estrategia Empírica	41
4.1. Efectos Fijos y Aleatorios	41
4.2. Diferencias en Diferencias	44
5. Resultados	48
5.1. El efecto de hablar una lengua indígena sobre el analfabetismo	48

5.2. El impacto del Programa de Inclusión y Equidad Educativa	52
5.2.1. Diferencias en Diferencias	52
5.2.2. Impacto por edades	56
5.3. Mecanismo Causal y Efectos Indirectos	66
5.3.1. Mecanismo causal: Asistencia escolar	67
5.3.2. Mecanismo causal: Matrícula escolar	71
5.3.3. Mecanismo causal: Número de docentes . . .	74
5.3.4. Mecanismo causal: Docentes con educación superior	76
5.3.5. Mecanismo causal: Número de aulas	78
5.3.6. Efectos indirectos: Alfabetización de madres y abuelas	80
6. Conclusiones	82
7. Bibliografía	86
8. Replicabilidad	93

1. Introducción

“Las instituciones básicas de la sociedad, según Rawls, no deben distinguirse simplemente por ser ordenadas y eficientes: ellas deben ser, sobre todo, justas. Y si no lo son, entonces, deben ser reformadas o abolidas.” (Gargarella 1999).

El proceso de alfabetización tiene muchas implicaciones sobre el desarrollo de la persona. Es el principal canal de comunicación con el mundo: cuando te levantas todas las mañanas, sabes que son las 8 a.m. porque puedes leer los números del reloj, sales a tomar el transporte público y sabes hacia dónde dirigirte porque puedes leer las señales e indicaciones del lugar, paras en el supermercado para comprar algunos bienes que te hacen falta y puedes comparar precios y calidad de productos y marcas para tomar la mejor decisión de consumo.

Usualmente, no le damos importancia al papel que juegan la lectura y escritura básicas en nuestras vidas, porque son cosas que damos por hecho, como si hubiéramos nacido con ellas por naturaleza. Sin embargo, no es así, son aptitudes que se desarrollan, principalmente durante la infancia, y que son condiciones necesarias para acceder a un nivel de vida digno.

Sobre esta misma línea, el analfabetismo se traduce no solo en la privación del derecho a la educación, sino también en un canal importante para la marginación y la pobreza. Cuando una persona

no tiene acceso a la educación (especialmente, en este caso, al proceso de alfabetización) es muy probable que esta problemática tenga implicaciones sobre el bienestar económico de la persona y sobre su acceso al resto de los derechos sociales.

“When children are first taught in a language that they speak and understand well, they learn more, are better placed to learn other languages, are more likely to stay in school and enjoy a school experience appropriate to their culture and local circumstances”. (World Bank 2021, p. 8).

Existe una amplia literatura al respecto del impacto de enseñar en un idioma conocido por las y los estudiantes, lo cual no solo tiene un impacto sobre su logro académico (Bühmann y Trudell 2007; Miscevic-Kadijevic 2015), sino sobre otras variables relevantes como la asistencia escolar (Smits, Huisman y Kruijff 2008).

En todo caso, esta evidencia tiende a señalar que la enseñanza básica se tendría que basar en la lengua materna (Bühmann y Trudell 2007; Ganuza y Hedman 2017; Nkonde, et al. 2018), mientras que el español se aprenda solo como segunda lengua durante los primeros años. Si, eventualmente, se llega a secundaria o preparatoria sin haber aprendido español, el desempeño escolar será menor en promedio que el de aquellos estudiantes bilingües o cuya lengua materna es el español.

En el caso de México, la población indígena es una de las más afectadas por esta problemática: en 2010, el porcentaje de la población de 15 años o más con analfabetismo se situó en 6.9%; sin embargo, entre la población que habla lengua indígena, la cifra se eleva a un 27.3% (Narro & Moctezuma, 2012). La problemática es aún más grave entre adultos mayores y mujeres: en el mismo año, el 72.7% de las mujeres de 60 años o más que hablaban lengua indígena eran analfabetas (Narro & Moctezuma, 2012).

Este resultado tiene múltiples causas que pasan por la pobreza, marginación, discriminación, desigualdad, falta de movilidad social, entre otros elementos. En un primer momento, este ensayo buscará aislar el efecto de estas causas y probar que el propio hecho de hablar una lengua indígena tiene un impacto positivo sobre el analfabetismo en México.

Esto se debe a que las lenguas indígenas provienen de múltiples familias lingüísticas diferentes, con estructuras sintácticas y morfológicas completamente diferentes al español. Esto dificultaría los procesos de alfabetización hasta tener un impacto positivo sobre el analfabetismo.

Para verificar este resultado, se utilizó información a nivel municipal proveniente del Censo de Población 2010 y la Encuesta Intercensal 2015 del Instituto Nacional de Geografía y Estadística (INEGI) y se estimaron modelos de regresión para datos panel de efectos fijos y efectos aleatorios. Se concluye que un incremento

de un punto en el porcentaje de hablantes de lengua indígena incrementa en 0.052 puntos el porcentaje de analfabetismo en los municipios de México, controlando por el nivel de pobreza, desigualdad, género, urbanización, población, rezago educativo, nivel académico, asistencia escolar y bienestar económico.

Esta evidencia señala la necesidad del sistema educativo mexicano para adaptarse para permitir una mayor inclusión de la población indígena en los procesos de alfabetización. Se requiere fortalecer a las escuelas de educación indígena: por ejemplo, a través de políticas públicas que fortalezcan la infraestructura educativa de estas escuelas, que mejore la calidad docente y que mejore la asistencia escolar, eficiencia terminal y los aprendizajes de sus estudiantes.

En segundo lugar, se buscará hacer una evaluación de impacto al componente de educación indígena del Programa de Inclusión y Equidad Educativa de la Secretaría de Educación Pública (SEP), el cual brindaba apoyo económico a las escuelas primarias de educación indígena en México entre 2014 y 2019. Para ello, se utilizaron modelos de Diferencias en Diferencias y se estimó que la presencia del programa redujo el porcentaje de analfabetismo entre la población de 6 a 11 años de los municipios con escuelas de educación indígena en 3.424 puntos porcentuales de 2010 a 2020. El mayor impacto se parece concentrar entre las niñas y niños de 6 a 8 años (entre 4.220 y 5.741 puntos porcentuales), pero la evidencia más robusta se refiere al efecto para la población de 8 a 11 años:

4.494, 2.660, 1.600 y 1.083 puntos porcentuales, respectivamente.

Para cada especificación de modelos, se incorporaron efectos placebo entre 2000 y 2010 con el objetivo de aportar evidencia sobre el supuesto de tendencias paralelas. En general, el supuesto pareciera creíble a la luz de que ninguno de los coeficientes del placebo mostraron significancia estadística (a excepción de las regresiones para 6 y 7 años).

En términos del mecanismo causal, este efecto podría pasar por un aumento en la infraestructura escolar o por una mejora en la calidad docente. Aunque no se cuenta con información precisa de lo anterior para el periodo de análisis, se utilizaron múltiples variables proxy: asistencia escolar, matrícula escolar, número de aulas, número de docentes y número de docentes con educación superior. Para ello, se utilizó la información del Formato 911 de la SEP exclusivamente para escuelas de educación indígena (excepto para asistencia escolar, donde se utilizó igualmente información de los Censos y Encuestas Intercensales del INEGI).

La información no resulta contundente, pero parece señalar un incremento 1.313 puntos porcentuales en la asistencia escolar por municipio e incrementos de 6.783 alumnas y alumnos en la matrícula escolar y 0.541 aulas promedio por escuela de educación indígena en municipios del programa.

Intervenciones relativamente simples como la del Programa de In-

clusión y Equidad Educativa pueden mejorar los aprendizajes del español durante el periodo crítico (previo a los 12 años). Sin embargo, si bien originalmente este programa se fragmentó en cuatro programas diferentes en 2020, el componente para educación indígena (ahora Programa de Atención a la Diversidad de la Educación Indígena) desapareció para el ejercicio 2021.

2. Marco Teórico

2.1. Alfabetización y diversidad lingüística

*“Los límites de mi lenguaje significan los límites de mi mundo.”
(Wittgenstein 1921).*

El analfabetismo es la carencia de habilidad para leer y escribir un breve recado (CONEVAL 2018). Cabe destacar que existe otra clase de analfabetismo denominado analfabetismo funcional, el cual se refiere a la falta de habilidad para comprender el contenido de los mensajes. Sin embargo, este estudio se enfocará, exclusivamente, al respecto del analfabetismo por el importante papel que juega para los procesos de aprendizaje en general y para el pleno desarrollo de la persona.

Lenneberg (1975), considerado el padre de la biología del lenguaje, plantea la hipótesis de que existe un periodo crítico para el aprendizaje del lenguaje durante la infancia cuando el cerebro tiene mayor sensibilidad para construir redes neuronales. Newport (1990) demuestra que, por ejemplo, aprender una segunda lengua durante los primeros 12 años genera que las niñas y los niños hablen ese idioma sin acento. Si el lenguaje se aprende después, muestra, en promedio, más deformaciones en la morfología y la sintaxis de las oraciones. Esto se podría deber a dos hipótesis:

“1. The crucial constraints needed for language acqui-

sition are those of a special language faculty, equipped in advance to expect only certain types of structures to occur in human language. Maturational change may occur in these constraints, leading the older language learner to less success in inducing the linguistic system to which she is exposed”.

“2. At least, some of the constraints crucial to success in language acquisition are nonlinguistic and the maturational changes which lead to more difficulty in language learning occur in these nonlinguistic constraints on perception and memory”. (Newport 1990, p. 27).

Es decir, es posible que ciertos tipos de estructuras del lenguaje requieran de determinadas facultades lingüísticas (o no lingüísticas) previas. Si estas no se desarrollaron previamente, aprender una segunda lengua o una lengua durante la etapa adulta sería más accidentado.

Algunos de los elementos esenciales del lenguaje son la sintaxis, semántica y morfología. La sintaxis es el conjunto de reglas del lenguaje que homologan estructuras sobre cómo coordinar y secuenciar palabras en una oración. El elemento que guía a la sintaxis es el verbo: “el verbo contiene información que no solo organiza la frase, sino que también ayuda a determinar su significado” (Cárdenas 2010). Por otra parte, la semántica se encuentra mu-

cho más vinculada con el estudio del significado de las palabras (Cárdenas 2010). Sin embargo, el significado siempre parte del significante, por lo que la semántica introduce el vínculo directo entre el lenguaje como algo mecánico o exógeno y el lenguaje como elemento de apropiación.

Finalmente, la morfología se refiere a la estructura de las palabras en lo individual. Es decir, a diferencia de la sintaxis que se refiere a la composición de oraciones, la morfología toma a cada palabra como unidad de análisis. ¿Por qué estos conceptos son importantes? Porque los lenguajes forman parte de familias: familias lingüísticas. Son grupos que mantienen una misma línea de origen; provienen del mismo punto y, por ende, mantienen cierta relación en sus estructuras ortográficas y gramaticales (Marlett 2019).

Cuando una niña crece aprendiendo español, le será relativamente más sencillo aprender otras lenguas romances en comparación con otra niña que creció hablando inglés. El principio aquí es el mismo, en un país con tanta riqueza cultural como lo es México, existe también una gran pluralidad de lenguas que pertenecen a múltiples familias lingüísticas. La Clasificación de Lenguas Indígenas del INEGI contiene 334 lenguas indígenas; de las cuales solo 95 se pueden agrupar entre 13 familias lingüísticas (INEGI 2010).

Solo como contraste para visualizar el grado de pluralidad lingüística que existe en México, se estima que, en África, existen cerca

de dos mil lenguas; en Asia del Sur, mil quinientas y, en Nueva Guinea, mil (Albó & at 2009). Sin embargo, en estos lugares, estas lenguas se suelen agrupar entre 10 y 27 familias lingüísticas (Albó & at 2009). Mientras que solo en México se agrupan un total de 13 familias y restan 239 lenguas a las que no se les puede vincular con ninguna de éstas (INEGI 2010).

Sumado a la dificultad lingüística para aprender español, un rezago durante su aprendizaje en la infancia le podría ocasionar efectos adversos en el largo plazo. Visualizando este contexto, una niña que crece aprendiendo una de estas lenguas indígenas tendrá mayor dificultad para aprender español, en comparación con otras niñas cuyo contexto es hispanoparlante. En otros términos, la niña indígena enfrentará mayor probabilidad de ser analfabeta que su contraparte.

Por supuesto, en esta comparación, subyacen otras diferencias importantes. Probablemente, la niña que creció aprendiendo una lengua indígena vive en una localidad rural con menor infraestructura y mayor rezago educativos; tiene mayor probabilidad de sufrir condiciones de pobreza y situaciones de desigualdad y discriminación, entre otros elementos sociodemográficos que la ponen en una situación de mucha desventaja respecto del resto de sectores de la población. Más adelante, se explicará la estrategia empírica para aislar el efecto del lenguaje sobre analfabetismo para realizar una comparación hipotética entre la misma niña indígena si hu-

biera nacido en un contexto hispanoparlante, pero bajo las mismas condiciones de pobreza, urbanización, desigualdad, etc.

2.2. Desarrollo de habilidades cognitivas y educación básica

“Ability gaps between individuals and across socioeconomic groups open up at early ages, for both cognitive and noncognitive skills.” (Cunha y Heckman 2007, p. 32).

La educación no es el único derecho que, históricamente, se le ha negado a la población indígena. Sin embargo, sí es uno de los más cruciales, debido a dos características esenciales del proceso educativo. Además del periodo crítico para el desarrollo de habilidades cognitivas (incluyendo el lenguaje), este proceso presenta dos características importantes: complementariedad dinámica y autoproduktividad (Cunha y Heckman 2007; Heckman y Mosso 2014).

Si tomamos el modelo de formación de capital humano de Cunha y Heckman (2007), tenemos que el stock de habilidades en el periodo $t+1$ (θ_{t+1}) depende, esencialmente, de tres factores: las características parentales (h), el stock de habilidades en t (θ_t) y la inversión en capital humano en t (I_t).

$$\theta_{t+1} = f_t(h, \theta_t, I_t)$$

La complementariedad dinámica se refiere a que mayores habilidades en un periodo incrementan la productividad de la inversión educativa en el periodo siguiente: esto es que $\frac{\partial^2 f_t}{\partial \theta_t I_t} > 0$. La auto-productividad significa que un mayor nivel de habilidades en un periodo incrementan el nivel de habilidades en el siguiente periodo: es decir, que $\frac{\partial f_t}{\partial \theta_t} > 0$ (Cunha y Heckman 2007; Heckman y Mosso 2014).

Esto implica que rezagos en la educación inicial y básica no solo tienen consecuencias inmediatas en el aprendizaje de las niñas y los niños, sino que, además, tiene un impacto negativo sobre sus habilidades y su capacidad de aprender con el paso de los años.

Existe una amplia literatura al respecto (Elango, et al. 2016), pero, por ejemplo, Chetty, et al. (2011) demuestran que determinados factores de la educación inicial tienen efectos importantes en el largo plazo a partir del *Project STAR*. Haber asistido a un grupo pequeño en Kinder incrementa la probabilidad de asistir a la universidad, adquirir una propiedad y tener mayores ahorros para el retiro. Mayores niveles de experiencia de sus maestras y haber pertenecido a un grupo con niñas y niños con mejores resultados promedio tienen un impacto positivo sobre sus ingresos al cumplir 27 años (Chetty, et al. 2011).

Asimismo, Carneiro, et al. (2019) encuentran que una intervención de bajo costo de sesiones parentales para madres y padres con bebés de 0 a 5 años pueden incrementar el vocabulario y ha-

bilidades socio-emocionales de las niñas y los niños 3 años después del experimento.

Al respecto de la educación básica, De Hoyos, Estrada y Vargas (2018) demuestran que los resultados en pruebas estandarizadas durante la educación primaria son buenos predictores no solo de aprendizajes en secundaria y preparatoria, sino de tener un empleo posteriormente.

2.3. Un programa para la educación indígena

“Las lenguas son importantes, pero mucho más importantes son sus hablantes. Las lenguas mueren porque sus hablantes son discriminados y violentados”, Elena Yasnaya Aguilar.

Si bien no existe un programa social en México que se concentre en la educación inicial indígena, sí existió uno para la educación básica: el Programa de Inclusión y Equidad Educativa coordinado por la SEP. Este programa surgió en 2014 como fusión de siete programas presupuestarios: cinco de educación básica, uno de media superior y uno de superior (MOCYR, 2018). Esto explica que, dentro de un solo programa, se concentran seis componentes: i) apoyo a escuelas de educación indígena, ii) apoyo a centros educativos para migrantes, iii) apoyo a servicios de educación especial, iv) apoyo a escuelas de educación telesecundaria, v) apoyo a Centros de Atención a Estudiantes con Discapacidad (CAED) de

educación media superior y vi) apoyo a Instituciones Públicas de Educación Superior (IPES) para la inclusión de estudiantes con discapacidad o en contextos de vulnerabilidad.

El presente estudio se enfoca únicamente en el componente de apoyo a escuelas de educación indígena, a cargo de la Dirección General de Educación Indígena de la Subsecretaría de Educación Básica. En el diagnóstico del programa, se identifica que el problema público que atiende es una “inadecuada oferta educativa a las diversas necesidades de las niñas, los niños y los jóvenes en contexto vulnerable” (SEP 2014; p. 21). Así, la población objetivo se definió como “todas las escuelas de Educación Básica que ofrecen los servicios de educación a indígenas, migrantes, educación especial y telesecundarias que decidan participar en el programa” (SEP 2014, p. 25).

Sin embargo, por otro lado, las Reglas de Operación (ROP) del programa definen que los beneficiarios son los gobiernos estatales y la Autoridad Educativa Federal en la Ciudad de México (AEFCM) que decidan participar voluntariamente y cuyas Autoridades Educativas Locales (AEL) atiendan escuelas de educación indígena, centros de educación migrante y servicios públicos de educación especial, para el caso del nivel básico (DOF, 2019).

Las ROP establecen que el objetivo específico del componente de educación indígena es “beneficiar a escuelas de educación indígena con acciones de fortalecimiento académico y contextualización”

(DOF 2019, p. 9). Asimismo, contemplan la entrega de tres tipos de bienes y servicios: apoyos técnicos, apoyos financieros y gastos de operación local.

El apoyo técnico se refiere a asesoría y acompañamiento para cumplir con la normatividad que regía al programa, así como para la ejecución de acciones para la mejora educativa. El gasto de operación local se refiere al 3% de los subsidios transferidos, que podrían utilizarse para cubrir gastos de operación estatales. Los apoyos financieros anuales se determinarían de acuerdo con “los objetivos del Programa para la Inclusión y la Equidad Educativa y los criterios de distribución establecidos por la Dirección General de Desarrollo Curricular (DGDC) y la DGEI” (DOF 2019, p. 14).

Básicamente, el programa entregaba subsidios a las escuelas primarias de educación indígena que lo solicitan (a través de las AEL), que podían tener como destino (DOF 2019):

1. Acciones para el fortalecimiento académico: diplomados, congresos, seminarios, viáticos y transporte, subsidios para educación continua de docentes indígenas, asesoría y asistencia técnica.
2. Acciones de contextualización: adaptación de contenidos y materiales, diseño de métodos y estrategias de enseñanza, ejecución de programas de lectura y escritura en lengua indígena y compra de materiales educativos.

Al respecto de las “acciones de contextualización”, Duarte, Moreno y Gargiulo (2011) afirman que la conexión a servicios de electricidad, telefonía, agua potable, desagüe y baños son determinantes para el logro de los estudiantes en pruebas de lenguaje y matemáticas. Además, Barret, et al. (2019) señalan que instalar protección frente a desastres naturales, acceso a servicios básicos (luz, agua, saneamiento y comunicaciones), espacios para recreación al aire libre y condiciones físicas buenas y seguras en las escuelas son elementos que tienen un impacto positivo sobre la asistencia escolar de alumnos y maestros, la salud de los alumnos en la escuela y la probabilidad de que los maestros se mantengan en dicha profesión.

“Investments in school infrastructure and the physical conditions for learning are not luxury but a need”.

(Barret, et al. 2019, p. 13).

No solo la infraestructura de las escuelas importan, sino también la adquisición de material educativo, como los libros de texto. Evans, et al. (2010) demuestran que las niñas y niños que crecen en hogares con muchos libros adquieren, en promedio, tres años más de escolaridad que niñas y niños en hogares sin libros.

Sin embargo, Glewwe, Kremer y Moulin (2009) argumentan que los estudiantes con peores antecedentes académicos no se benefician de la adquisición de libros de texto en las escuelas. De hecho,

el programa en que se basan y que entregó libros de texto aleatoriamente en Kenya no tuvo un impacto sobre los aprendizajes en el agregado.

Asimismo, sobre este tipo de acciones *input-based*, De Hoyos, Holland y Troiano (2015) demuestran que políticas basadas en incrementar el número de docentes por grupo y su salario, mejorar la infraestructura escolar e implementar el uso de computadoras no llevaron a mejores resultados en pruebas estandarizadas en el caso de Argentina.

“This is not to say that education inputs are not important in the learning process, but their impact is limited and seems to diminish over time because it is subject to the systems ability to use them in an effective way”.

(De Hoyos, Holland y Troiano 2015).

Es decir, las políticas basadas en la adquisición y mejora de infraestructura educativa dependen de la capacidad institucional y docente para transformarla en el aprendizaje de los estudiantes. Esto está fuertemente ligado con la relevancia que tendrían las “acciones para el fortalecimiento académico”.

Las “acciones para el fortalecimiento académico” del Programa para la Inclusión y la Equidad Educativa se podrían utilizar para subsidiar la formación continua de docentes. Mejorar la calidad docente puede tener un impacto importante sobre los aprendizajes

(Rivkin, Hanushek y Kain 2005; Bruns y Luque 2015), porque son el principal insumo del proceso educativo. Hanushek, Piopiunik y Wiederhold (2014) demuestran que las habilidades cognitivas de los docentes están altamente correlacionadas con los resultados de los estudiantes.

Sin embargo, es importante delimitar que las ROP del programa no permiten delimitar exactamente a qué tipo de seminarios, conferencias y viáticos se refiere y, por lo tanto, no es posible garantizar que dichas actividades tengan un impacto efectivo sobre la calidad docente.

En conclusión, las transferencias del programa bien podrían tener un impacto positivo sobre los aprendizajes de los alumnos de escuelas indígenas. Sin embargo, es también probable que estos subsidios no sean suficientes; sobre todo en contextos donde hubiera docentes con menor nivel pedagógico o acervo de capital humano y las “acciones para el fortalecimiento académico” no hubieran sido suficientes para mejorar la formación continua de docente.

Bajo el contexto en el que hablar una lengua indígena podría dificultar el proceso de alfabetización y tomando en cuenta que el proceso educativo tiene un periodo crítico durante la infancia y que presenta las características denominadas como complementariedad dinámica y autoproduktividad, se vuelve relevante analizar el impacto del programa para mejorar los aprendizajes de la lengua española entre la población indígena, especialmente en el nivel

básico.

3. Presentación de Variables y Análisis Exploratorio

3.1. Analfabetismo, población indígena y características sociodemográficas

Para realizar el análisis empírico de la primer hipótesis, se utilizaron cifras del Censo de Población 2010 y la Encuesta Intercensal 2015 que realizan el Instituto Nacional de Estadística y Geografía (INEGI) (se excluye de esta parte del análisis el Censo de Población 2020, ya que, para el momento de escribir este documento, no se han publicado las estimaciones 2020 de pobreza del CONEVAL). La información se encontraba a nivel individuo; sin embargo, se colapsó en promedios y porcentajes a nivel municipal, de tal forma que todo el proyecto analiza información sobre un total de 2,445 municipios.

De estas bases, se extrajeron las variables del porcentaje de analfabetismo, de población que habla lengua indígena, de mujeres, de asistencia escolar y de localidades rurales y los promedios de edad y escolaridad. Para construir el porcentaje de mujeres, se construyó una dummy de mujer a partir de la variable sexo y, para el porcentaje de localidades rurales, se construyó una dummy de localidad rural si el tamaño de la localidad era menor a 2,500 personas.

Para incorporar información de otros controles, se utilizó la base de la Medición Multidimensional de la Pobreza del Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL). Así, se obtuvieron los porcentajes de pobreza y de pobreza extrema, de población con ingreso inferior a la línea de bienestar y bienestar mínimo, de población con carencia por acceso a los derechos sociales (educación, salud, vivienda, servicios básicos en la vivienda y alimentación), las cifras de población y el coeficiente de Gini a nivel municipal.

A continuación, se muestran algunas estadísticas descriptivas de estas variables en 2010 y 2015:

Table 1: Estadísticas descriptivas en 2010

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Mujeres	2,445	51.41	1.95	42.92	50.21	52.62	62.19
Edad	2,445	32.57	3.38	23.87	30.35	34.44	47.87
Alfabetismo	2,445	84.51	8.76	35.35	80.00	91.33	98.52
Hablan lengua indígena	2,445	19.36	31.33	0.00	0.26	24.71	99.91
Asistencia escolar	2,445	30.03	3.96	15.00	27.61	32.33	46.04
Nivel académico	2,445	2.39	0.42	1.08	2.11	2.63	4.67
Rural	2,445	59.44	35.97	0.00	27.05	100.00	100.00
Pobreza	2,445	67.01	20.15	3.18	53.05	83.46	99.40
Pobreza extrema	2,445	25.53	20.64	0.10	8.90	37.40	90.50
Población	2,445	46,786.00	135,122.60	90	4,181	33,577	1,807,581
Rezago educativo	2,445	31.44	10.67	2.60	23.60	38.90	65.40
Carencia en salud	2,445	34.66	18.47	1.20	21.50	43.80	98.20
Carencia en seguridad social	2,445	75.55	15.63	3.30	67.10	87.50	98.70
Carencia en vivienda	2,445	25.70	16.63	0.00	12.00	36.40	85.50
Car. en servicios de vivienda	2,445	47.15	29.95	0.00	19.40	72.40	100.00
Carencia alimentaria	2,445	30.38	18.26	0.00	16.80	40.00	99.30
Ingreso inferior a LB	2,445	70.15	18.19	4.20	57.40	84.90	99.40
Coficiente de Gini	2,445	0.43	0.05	0.29	0.40	0.46	0.66
Programa	2,445	0.33	0.47	0	0	1	1

Table 2: Estadísticas descriptivas en 2015

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Mujeres	2,445	51.52	1.69	44.21	50.52	52.49	59.78
Edad	2,445	33.87	3.41	24.21	31.74	35.74	50.99
Alfabetismo	2,445	87.17	7.42	46.65	83.52	92.78	98.54
Hablan lengua indígena	2,445	19.24	30.94	0.00	0.31	24.42	99.97
Asistencia escolar	2,445	28.66	3.88	13.75	26.40	30.98	44.54
Nivel académico	2,445	2.55	0.42	1.37	2.27	2.80	4.85
Rural	2,445	59.10	36.02	0.00	26.62	100.00	100.00
Pobreza	2,445	65.52	21.51	2.73	50.13	83.06	99.94
Pobreza extrema	2,445	19.71	17.96	0.00	5.60	28.50	97.50
Población	2,445	49,627.23	144,460.50	90	4,362	34,090	1,903,552
Rezago educativo	2,445	27.86	10.07	2.50	20.50	34.70	60.60
Carencia en salud	2,445	14.17	7.00	0.90	9.10	18.10	77.40
Carencia en seguridad social	2,445	73.99	15.73	5.90	66.30	85.60	97.00
Carencia en vivienda	2,445	19.83	13.87	1.00	8.90	28.10	82.70
Car. en servicios de vivienda	2,445	45.97	30.55	0	17.1	73.9	100
Carencia alimentaria	2,445	24.42	11.78	0	16.2	30.6	86
Ingreso inferior a LB	2,445	69.27	19.08	3.50	56.60	84.20	99.90
Coefficiente de Gini	2,445	0.39	0.04	0.30	0.37	0.41	0.64
Programa	2,445	0.33	0.47	0	0	1	1

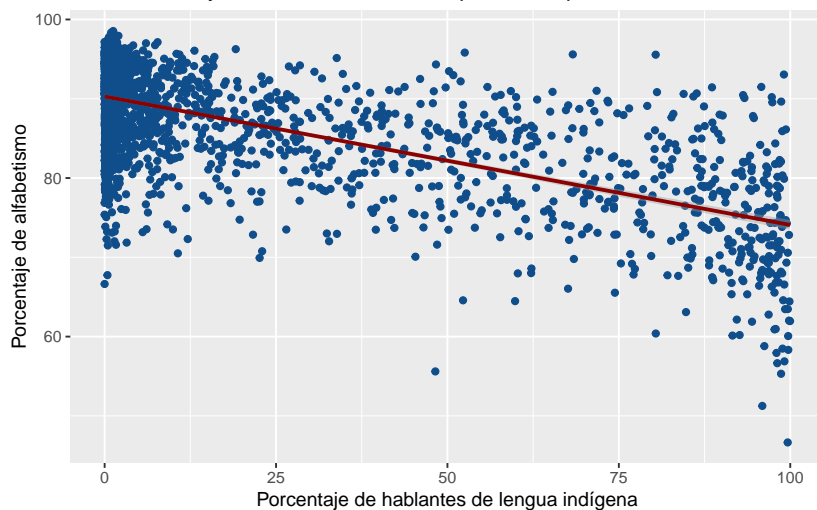
En 2015, los municipios de México tienen una tasa de alfabetismo promedio del 87.17% y cuentan con un porcentaje promedio de 19.24% de población que habla lengua indígena. El municipio con mayor porcentaje de analfabetismo es Cochoapa el Grande, ubicado en el estado de Guerrero. Encabezaba la lista en 2010 y se mantuvo en 2015.

Personalmente, tuve la oportunidad de conocerlo hace unos 6 años y me conmueve encontrarlo nuevamente en estas cifras. Es un ejemplo de cómo el crecimiento del país se ha olvidado de estos municipios perdidos en la sierra; sin infraestructura para entrar y salir (y mucho menos para llevar y sacar productos y materias primas), sin escuelas ni hospitales, sin servicios públicos básicos y sin acceso al desarrollo.

En el caso del porcentaje de población que habla lengua indígena, los tres municipios que encabezan la lista se encuentran en el estado de Chiapas: Santiago el Pinar, San Juan Cancuc, Chalchihuitán. Prácticamente, toda su población habla lengua indígena.

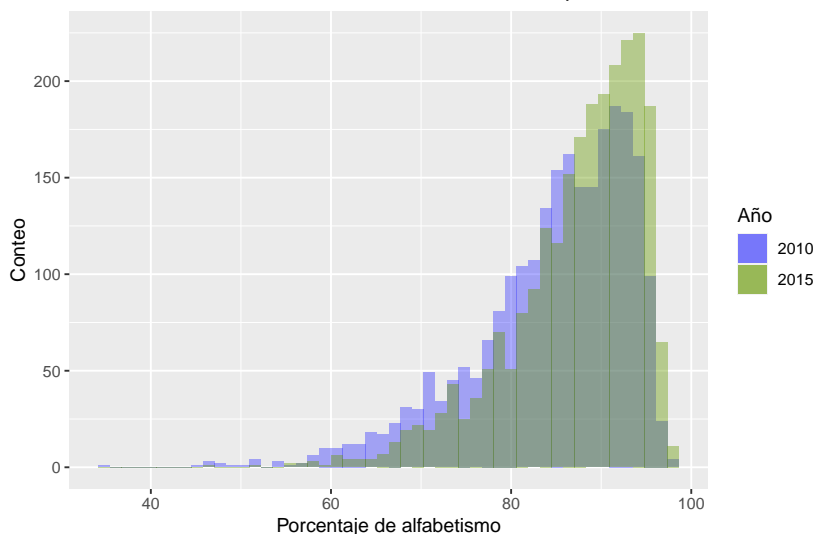
La relación entre el alfabetismo y el indigenismo es bastante clara. Cuanto mayor es el porcentaje de población que habla lengua indígena, menor es el nivel de alfabetismo en el municipio. Si bien esto no es evidencia para hacer inferencia causal, sí es una primera intuición sobre el vínculo inverso que podría existir entre ambas variables.

Gráfica 1. Relación entre el porcentaje de hablantes de lengua indígena y el nivel de alfabetismo por municipio en 2015



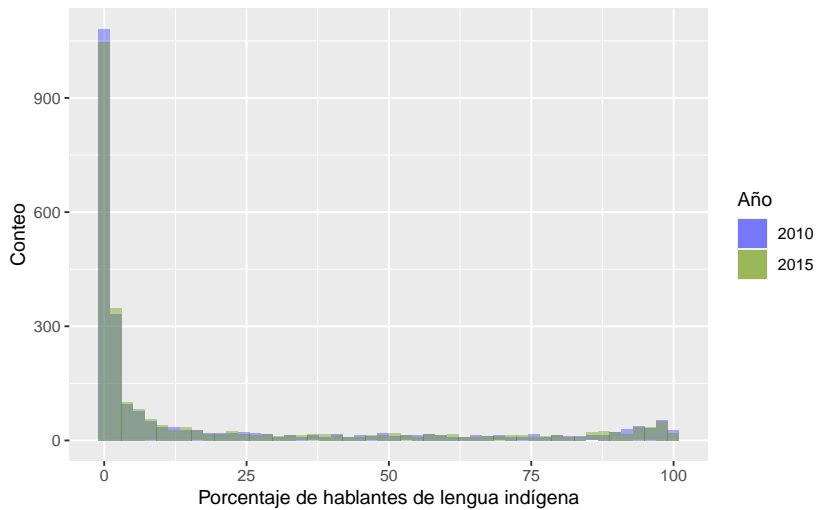
A continuación, se observa la distribución del alfabetismo entre los municipios de México: parece ser que la distribución se ha movido hacia una mejor posición entre 2010 y 2015. Sin embargo, aún se observan muchos municipios rezagados en el lado izquierdo de la distribución.

Gráfica 2. Densidad del alfabetismo en los municipios de México



En lo que respecta a la distribución de la población que habla lengua indígena, no parece haber cambiado mucho. Esto podría ser problemático a la hora de estimar el modelo empírico, debido a que la estrategia de identificación se basa en la estimación de diferencias temporales. Si la variable no cambia a lo largo del tiempo, no se podría medir su efecto utilizando un modelo de efectos fijos o efectos aleatorios.

Gráfica 3. Densidad de hablantes de lengua indígena en los municipios de México



Para verificar si las variables cambiaron en el tiempo de manera significativa; a continuación, se muestra una tabla de balance con la prueba de hipótesis de medias: la hipótesis nula se refiere a la igualdad de medias.

Table 3: Diferencia de medias entre 2010 y 2015

	Variables	Medias_2010	Medias_2015	Diferencia	P_Value
alfabet	Alfabetismo	84.5065	87.1711	2.6646	0.0000
hlengua	Hablan lengua indígena	19.3595	19.2410	-0.1185	0.8941
mujer	Mujeres	51.4127	51.5186	0.1059	0.0428
edad	Edad	32.5720	33.8675	1.2955	0.0000
asisten	Asistencia escolar	30.0262	28.6599	-1.3663	0.0000
nivacad	Nivel académico	2.3924	2.5517	0.1593	0.0000
rural	Rural	59.4417	59.0960	-0.3458	0.7370
pobreza	Pobreza	67.0130	65.5243	-1.4887	0.0125
pobreza_ext	Pobreza extrema	25.5272	19.7073	-5.8199	0.0000
pob	Población	46786.0041	49627.2335	2841.2294	0.4776
rezed	Rezago educativo	31.4447	27.8621	-3.5826	0.0000
salud	Carencia en salud	34.6581	14.1706	-20.4876	0.0000
seg_soc	Carencia en seguridad social	75.5520	73.9899	-1.5620	0.0005
vivienda	Carencia en vivienda	25.7040	19.8263	-5.8777	0.0000
serv_viv	Car. en servicios en vivienda	47.1535	45.9688	-1.1848	0.1709
alim	Carencia alimentaria	30.3815	24.4176	-5.9638	0.0000
inf_lb	Ingreso inferior a LB	70.1499	69.2662	-0.8837	0.0974
gini	Desigualdad	0.4327	0.3916	-0.0411	0.0000

Estadísticamente, las variables que muestran un cambio significativo entre 2010 y 2015 fueron las medias de edad, nivel académico y desigualdad y los porcentajes de alfabetismo, mujeres, asistencia escolar, pobreza, pobreza extrema, rezago educativo y todas las carencias sociales excepto servicios en la vivienda.

Nuevamente, el cambio de 19.36% a 19.24% en el porcentaje de la población que habla lengua indígena parece no ser significativo. Si bien esto podremos comprobarlo también más adelante con las regresiones, esta prueba de balance deja una advertencia sobre la precisión y la veracidad de las estimaciones en el próximo capítulo.

3.2. Cobertura del Programa de Inclusión y Equidad Educativa

Para realizar el análisis empírico de la segunda hipótesis, se utilizó la misma base de datos anterior, incorporando el Censo de Población 2005 y el Censo de Población 2020, así como información sobre los municipios que recibieron recursos por parte del componente de educación indígena del Programa de Inclusión y Equidad Educativa de 2014 a 2019. Desgraciadamente, esta información se limita a señalar qué municipios percibieron recursos por parte del programa durante cada ejercicio fiscal, no se cuenta con información de la magnitud de estas transferencias a cada municipio y tampoco se cuenta con información desglosada por escuelas.

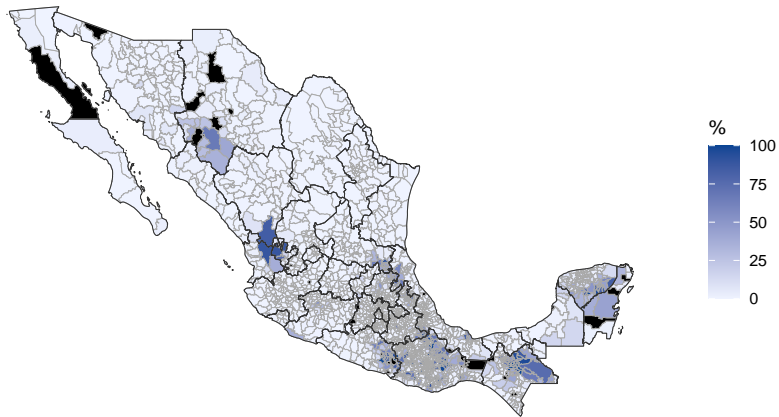
Esta información se obtuvo a través de la solicitud número 0001100442220 y el recurso de revisión RRA11287/20, mediante la Plataforma Nacional de Transparencia del Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI). Cabe destacar que la SEP señaló que los beneficiarios directos del programa eran las Entidades Federativas y que, por lo tanto, no cuenta con información presupuestaria desglosada por escuela y municipio. Sin embargo, para atender a la solicitud de información, se recolectó la información emitida por las Entidades Federativas al cierre de la cuenta pública de cada ejercicio fiscal para conocer qué municipios fueron beneficiados en cada caso.

Es posible que existan imprecisiones o errores de registro en la base de datos, toda vez que la captura de la información presentaba una importante heterogeneidad: por ejemplo, en algunos casos no se reportaba el municipio, sino la localidad o el nombre de la escuela. Para ello, se tuvo que limpiar la base de datos haciendo correcciones manuales para identificar el código del municipio correspondiente.

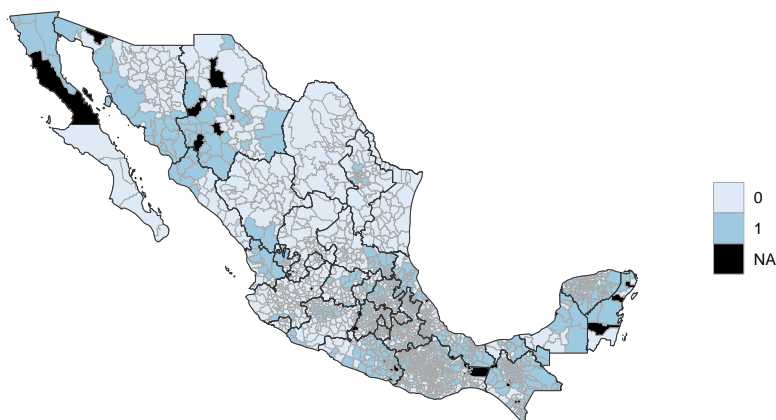
Como se puede observar a continuación, los estados del sureste, así como algunos municipios de Puebla, Oaxaca y otros estados por donde pasa la Sierra Madre Occidental concentran un alto porcentaje de población hablante de lenguas indígenas. Por otro lado, a pesar de algunas divergencias, la cobertura municipal del

Programa de Inclusión y Equidad Educativa pinta un mapa similar, lo cual no es una sorpresa, ya que los subsidios están destinados justamente para apoyar a las escuelas de educación indígena. Entre 2014 y 2019, el 33.36% de los municipios percibieron subsidios por parte del programa (817 municipios).

Mapa 1. Porcentaje de niñas y niños de 6 a 11 años que habla lengua indígena por municipio en 2010



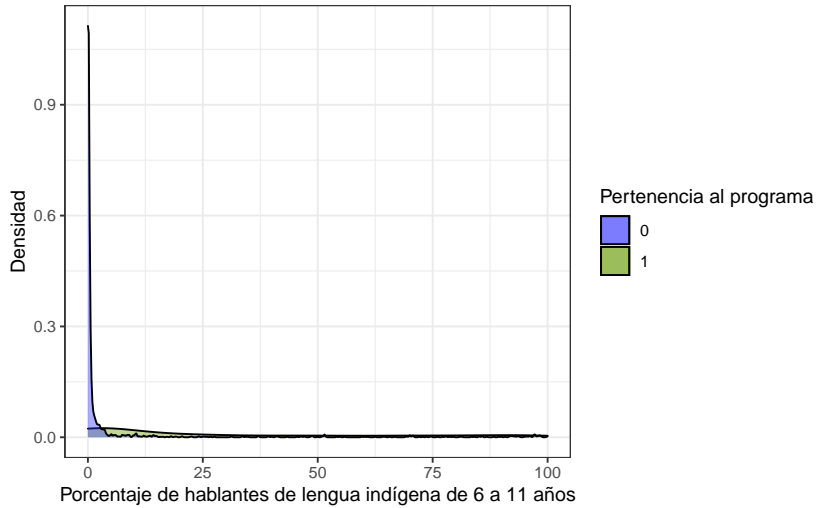
Mapa 2. Cobertura del Programa de Inclusión y Equidad Educativa entre 2014 y 2019



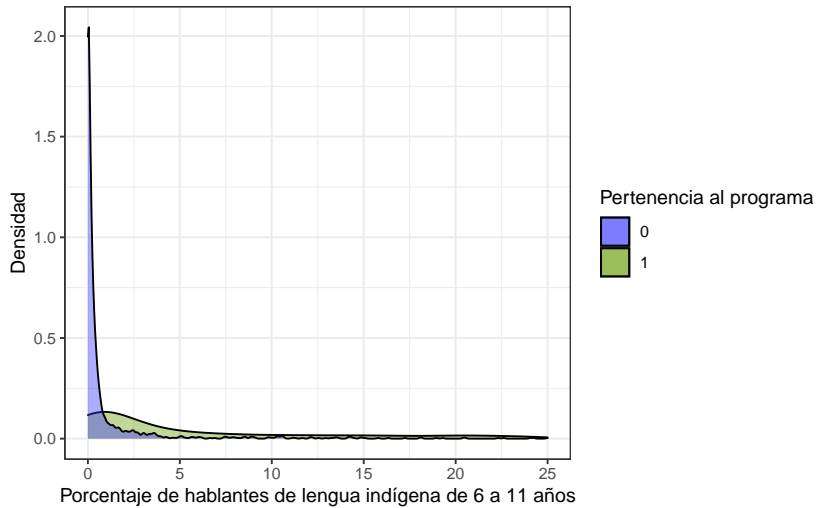
Ahora, si observamos la distribución del porcentaje de población que habla lengua indígena entre los municipios con y sin subsidios del programa, queda más claro aún que la distribución del programa no es aleatoria, sino que se encuentra focalizada en las localidades con mayor porcentaje de población indígena.

Es cierto que existe una mucho mayor concentración de municipios con poco porcentaje de población indígena (alrededor del 0%) pertenecientes al grupo de control. Sin embargo, es importante y, hasta cierto punto sorpresivo, que la distribución alrededor del extremo derecho concentra distribuciones similares de municipios de control y tratamiento. Es decir, existe un número importante de municipios con alto porcentaje de población indígena (entre 75% y 100%) que no formó parte del programa.

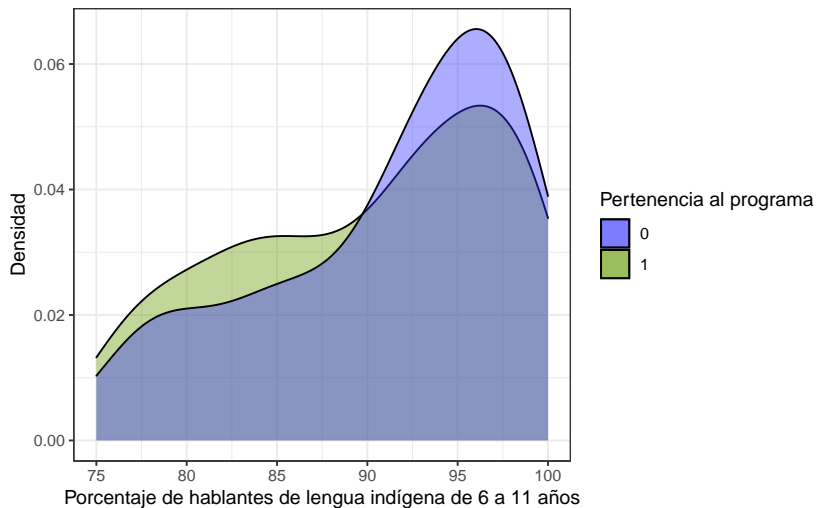
Gráfica 4. Hablantes de lengua indígena de 6 a 11 años entre municipios de control y tratamiento



Gráfica 5. Hablantes de lengua indígena de 6 a 11 años entre municipios de control y tratamiento (entre 0% y 25%)



Gráfica 6. Hablantes de lengua indígena de 6 a 11 años entre municipios de control y tratamiento (entre 75% y 100%)



Será importante visualizar las diferencias entre los municipios que formaron parte del programa y los que no previo a su ejecución. En los municipios del programa, el 29.97% de la población de 6 a 11 años hablaban una lengua indígena y contaban con un porcentaje de alfabetismo de 72.12% en promedio, el grupo de control solo tiene el 5.99% de población de 6 a 11 años indígena y 79.27% de alfabetismo en 2010.

En promedio, los municipios del programa son prácticamente diferentes en todo a los municipios del grupo de control, excepto en el porcentaje de mujeres, población en localidades rurales, carencia en acceso a servicios de salud y desigualdad. Su población tiene menor edad y nivel académico y mayor pobreza, pobreza extrema, asistencia escolar, población, rezago educativo, carencias sociales

e ingresos inferiores a la línea de bienestar mínimo.

Table 4: Diferencia de medias entre municipios de control y del programa en 2010 (población de 6 a 11 años)

	Variablen	Control	Programa	Diferencia	P_Value
alfabet	Alfabetismo	79.2683	72.1232	-7.1452	0.0000
hlengua	Hablan lengua indígena	5.9927	29.9667	23.9740	0.0000
mujer	Mujeres	49.2452	49.4423	0.1971	0.1412
asisten	Asistencia escolar	97.1099	96.1972	-0.9127	0.0000
nivacad	Nivel académico	1.7153	1.6903	-0.0250	0.0000
rural	Rural	58.5558	62.5362	3.9804	0.0077
pobreza	Pobreza	63.1767	74.2939	11.1172	0.0000
pobreza_ext	Pobreza extrema	20.9212	34.4274	13.5061	0.0000
pob	Población	41542.0149	58327.8886	16785.8737	0.0095
rezed	Rezago educativo	30.0183	34.1502	4.1319	0.0000
salud	Carencia en salud	34.2172	35.4705	1.2534	0.1191
seg_soc	Carencia en seguridad social	73.6726	79.2861	5.6135	0.0000
vivienda	Carencia en vivienda	21.3646	34.1672	12.8026	0.0000
serv_viv	Car. en servicios en vivienda	40.2788	60.5370	20.2582	0.0000
alim	Carencia alimentaria	26.3558	38.3059	11.9501	0.0000
inf_lb	Ingreso inferior a LB	66.7940	76.5068	9.7128	0.0000
gini	Desigualdad	0.4334	0.4312	-0.0021	0.2804

Debido a todo lo anterior, queda claro que una simple diferencia de medias entre los municipios de control y los del programa arrojaría una estimación sesgada. Este sesgo puede (y seguramente es el caso) venir de dos fuentes: diferencias en observables y diferencias en no observables.

Para controlar por las diferencias en observables, se utilizarán controles sociodemográficos entre 2010 y 2015 (el conjunto de variables que componen la tabla anterior), pero la más importante se refiere al porcentaje de la población que habla lengua indígena, ya que determina la distribución del programa.

Para controlar por las diferencias en no observables y como respuesta al problema de autoselección, se utilizará el método de Diferencias en Diferencias para estimar el impacto del programa, respecto del escenario hipotético en el que los municipios del programa no hubieran percibido los subsidios. De esta forma, se busca reforzar el supuesto de independencia condicional.

4. Estrategia Empírica

4.1. Efectos Fijos y Aleatorios

“It’s turtles all the way down”.

Como ya se mencionó anteriormente, no basta realizar una comparación entre el alfabetismo (o analfabetismo) entre la población que habla lengua indígena y la que no, ya que existen muchos sesgos de por medio vinculados con variables ya antes referidas: pobreza, desigualdad, ingresos, educación, etc. Por ello, no solo se buscará controlar estas variables introduciéndolas en el modelo, sino que se realizará una estimación de efectos fijos para controlar por cualquier otra variable que pudiera causar sesgo, pero que no cambie a lo largo del tiempo en cada municipio. El planteamiento es el siguiente:

$$Y_{i,t} = \beta_0 + \beta_1 X_{i,t} + \beta_2 Z_{i,t} + \beta_3 A_i + U_{i,t}$$

Donde $Y_{i,t}$ es el porcentaje de alfabetismo en el municipio i y año t , $X_{i,t}$ es el porcentaje de población que habla lengua indígena en el municipio i y año t , $Z_{i,t}$ es un vector de controles en el municipio i y año t , A_i es una variable no observable para el municipio i que no cambia en el tiempo y $U_{i,t}$ es el error de estimación en el municipio i y año t .

Si a este planteamiento le restamos la media de cada variable en el tiempo, se produce la siguiente ecuación:

$$(Y_{i,t} - \bar{Y}_i) = \beta_1(X_{i,t} - \bar{X}_i) + \beta_2(Z_i - \bar{Z}_i) + \beta_3(A_i - \bar{A}_i) + (U_{i,t} - \bar{U}_i)$$

Esta es la transformación de efectos fijos o transformación intra-grupal. Debido a que la variable A_i no cambia en el tiempo, se elimina de la especificación. Es decir, aunque no sea observable, su varianza está capturada en el modelo de efectos fijos para cada municipio y, por lo tanto, no causa sesgo en la estimación (Wooldridge 2015).

Otra forma de hacer esto es utilizando un modelo de efectos aleatorios: en este caso, se utiliza un elemento $\theta = 1 - (\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_a^2})^{\frac{1}{2}}$ para realizar la transformación del planteamiento de manera ponderada (Wooldridge 2015).

A partir de ello, nos quedamos con la siguiente estructura: $1 - \theta = (\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_a^2})^{\frac{1}{2}}$. Esto es equivalente a la proporción de la varianza que capta el modelo a través del error de estimación (σ_u^2) y del total de los errores “aleatorios” para cada municipio ($T\sigma_a^2$). Por lo tanto, $T\sigma_a^2$ es el componente que aporta homocedasticidad en los errores.

Si θ tiende a 1, significa que la mayoría de esta varianza se agrupa en σ_u^2 y casi nada en $T\sigma_a^2$, por lo que resulta más preciso utilizar la versión heterocedástica del modelo de efectos fijos. Pero, si θ tiende a 0, significa que $T\sigma_a^2$ explica casi toda la varianza del modelo: esto es, los errores aleatorios. En este caso, el modelo degenera

en Mínimos Cuadrados Ordinarios, toda vez que las variables omitidas que se controlaban con efectos fijos no explicaban nada del modelo.

Entonces, la principal diferencia del modelo de efectos aleatorios respecto del de efectos fijos es que se sostiene sobre un gran supuesto: que $Cov(X_{i,t}, A_i) = 0$. Es decir, que el factor A_i , a diferencia del planteamiento de efectos fijos, no estaría causando sesgo por omisión. Sin embargo, al menos en el contexto de esta evaluación, este supuesto parece poco realista.

Por ejemplo, volviendo al caso de Cochoapa el Grande, ¿por qué este era el municipio con mayor analfabetismo en 2005 (74%) y quince años después, en 2020, continúa ocupando la primer posición nacional (56%)? Claramente, hay algo ahí que persiste a lo largo de los años y de los sexenios. Entrar y salir de Cochoapa es muy difícil: si vas en carro (un privilegio de por sí) hay que dejarlo en la ciudad de Tlapa de Comonfort porque el camino se termina ahí. Posteriormente, debes de tomar un taxi o una camioneta de montaña para cruzar el camino hacia Cochoapa (o puedes intentarlo con tu auto, si es que llegas).

Una vez ahí, la infraestructura de servicios básicos es muy precaria: hogares sin drenaje y sin luz, sin alumbrado público e incommunicados, porque no hay señal de celular y el único teléfono fijo del pueblo se encuentra en el palacio municipal que de noche se convierte en albergue para muchas personas que no tienen un techo

para dormir. Impactan ahí la geografía, instituciones políticas, estructura demográfica, comercio, vías de transporte y comunicación, entre una larga lista de factores que escapan de las grandes mediciones nacionales de pobreza y de crecimiento.

Son pequeños detalles que persisten y que impactan de manera importante sobre el desarrollo y que, por lo tanto, en este contexto pareciera relevante hacer un intento por disminuir su sesgo. Por ello, en adelante, se utilizará como principal referencia la especificación de efectos fijos para hacer inferencia causal y, en su caso, se presentarán modelos OLS Pooled y de efectos aleatorios solo para verificar la robustez en signo, magnitud y significancia de los resultados.

4.2. Diferencias en Diferencias

“The first thing to realize about parallel universes... is that they are not parallel.” (Angrist y Pischke 2009, p. 221)

El modelo de Diferencias en Diferencias es una extensión de los modelos para datos tipo panel, por lo que puede incorporar efectos fijos o aleatorios. Cuando la distribución del tratamiento no es aleatoria, pero existe un grupo de control y un grupo de tratamiento y, además, se cuenta con información de tipo panel para ambos grupos antes y después de la aplicación del tratamiento, es posible recurrir a una especificación de dobles

diferencias. Esto es, incorporar la interacción entre la dummy del tratamiento y la dummy del año posterior a su aplicación.

“The source of OVB [Omitted Variable Bias] when evaluating these policies must therefore be unobserved variables at the state and year level. In some cases, group-level omitted variables can be captured by group-level fixed effects, an approach that leads to the differences-in-differences identification strategy”.

(Angrist y Pischke 2009, p. 227).

Dado que la aplicación del programa no es aleatoria, una simple diferencia de medias entre grupos estaría sesgada como ya se discutió anteriormente. Incorporar la interacción de dobles diferencias permite estimar la diferencia entre la media del *outcome* de interés, en este caso alfabetismo, entre el grupo del programa y un contrafactual que no es precisamente el grupo de control, sino el escenario hipotético sobre qué hubiera pasado con el grupo del programa de haber seguido el mismo comportamiento que el grupo de control.

El planteamiento es el siguiente:

$$Y_{i,t} = \beta_0 + \beta_1 P_i + \beta_2 Post_t + \beta_{DD}(P_i * Post_t) + U_{i,t}$$

Donde $Y_{i,t}$ es el porcentaje de alfabetismo en el municipio i y año t , P_i es una dummy que indica si el municipio i formó parte del

programa, $Post_t$ es una dummy que indica si el año t es posterior a la aplicación del programa (en este caso, posterior a 2013), $P_i * Post_t$ es la interacción entre las dos variables anteriores y, por lo tanto, es una dummy que indica si el municipio i formó parte del programa y, además, si el año t es posterior a su aplicación y $U_{i,t}$ es el error de estimación en el municipio i y año t .

De esta forma, el elemento de la interacción o doble diferencia ($P_i * Post_t$) captura el efecto del programa, siendo que:

$$\beta_{DD} = [E(Y_{i,t}|P_i = 1, Post_t = 1) - E(Y_{i,t}|P_i = 1, Post_t = 0)] - [E(Y_{i,t}|P_i = 0, Post_t = 1) - E(Y_{i,t}|P_i = 0, Post_t = 0)]$$

Además, es posible incorporar efectos fijos a esta especificación para controlar por variables que no cambian en el tiempo aunque, idealmente, el modelo de Diferencias en Diferencias simple tendría que controlar la varianza necesaria de tal forma que no exista sesgo por variable omitidas.

Dicho lo anterior, es importante señalar que el principal supuesto del modelo de diferencias en diferencias es denominado: tendencias paralelas. Esto es que el comportamiento del grupo de tratamiento o del programa seguiría la misma tendencia que el grupo de control, en ausencia del programa. Por lo tanto, podrían existir fuentes de sesgo (observables y no observables) que hagan que esto no suceda.

Si bien no es posible comprobar o afirmar que esto sucede (i.e. es un supuesto), sí es posible verificar qué tan creíble es utilizando efectos placebo con periodos previos. En este caso, se utilizará la información de 2000 a 2010 para identificar si los grupos del programa y de control muestran comportamientos similares antes de la ejecución del programa.

Para ello, se incorporarán dobles diferencias para cada año: las interacciones del programa con los años 2000 y 2005 serán los efectos placebo y las interacciones del programa con 2015 y 2020 estimarán su impacto. Idealmente, los coeficientes de los efectos placebo tendrían que ser pequeños y no significativos para que el supuesto de tendencias paralelas sea más creíble.

5. Resultados

5.1. El efecto de hablar una lengua indígena sobre el analfabetismo

A continuación, se muestra la tabla de regresión con la estimación para cada especificación. El primer modelo se refiere a una regresión *pooled* con errores heterocedásticos. Los modelos dos y tres corresponden a efectos fijos y efectos aleatorios, respectivamente, incorporando errores clusterizados por municipio, ya que es factible pensar que no exista independencia entre determinado municipio en 2010 y ese mismo municipio en 2015.

El coeficiente no cambia mucho entre modelos, lo cual es una buena noticia para la consistencia del resultado. En todos los casos, el coeficiente es negativo y significativo al 1%; la estimación puntual se encuentra entre -0.038 y -0.052 .

Debido a la significancia en los coeficientes de los modelos estimados, es factible pensar que sí existe un nivel de varianza en la diferencia temporal del porcentaje de población que habla lengua indígena, tal que permite estimar su efecto causal sobre el analfabetismo, controlando por variables incluidas como controles y por aquellas omitidas que podrían no estar cambiando en el tiempo.

Caeteris paribus, un incremento de un punto en el porcentaje de la población que habla lengua indígena causa un decremento prome-

dio de 0.052 puntos en el porcentaje de alfabetismo de los municipios de México, independientemente del nivel de pobreza, desigualdad, género, urbanización, población, rezago educativo y carencias sociales, nivel académico, asistencia escolar y bienestar económico.

Adicionalmente, se observa que existe una relación negativa entre los promedios de edad y desigualdad y los porcentajes de mujeres, asistencia escolar, pobreza extrema, población, rezago educativo, todas las carencias sociales (excepto seguridad social) y población con ingreso inferior a la línea de bienestar con el nivel de alfabetismo. Mientras que el nivel académico, localidades rurales, pobreza y carencia por acceso a seguridad social se asocian positivamente con el alfabetismo en México, sin que estos resultados tengan una interpretación causal.

Para observar mejor este resultado, a continuación, se muestra el efecto parcial de hablar una lengua indígena sobre el alfabetismo a partir de simular la estimación del modelo de efectos fijos por el método de *partial-out*.

Gráfica 7. Efecto parcial de hablar una lengua indígena sobre el alfabetismo

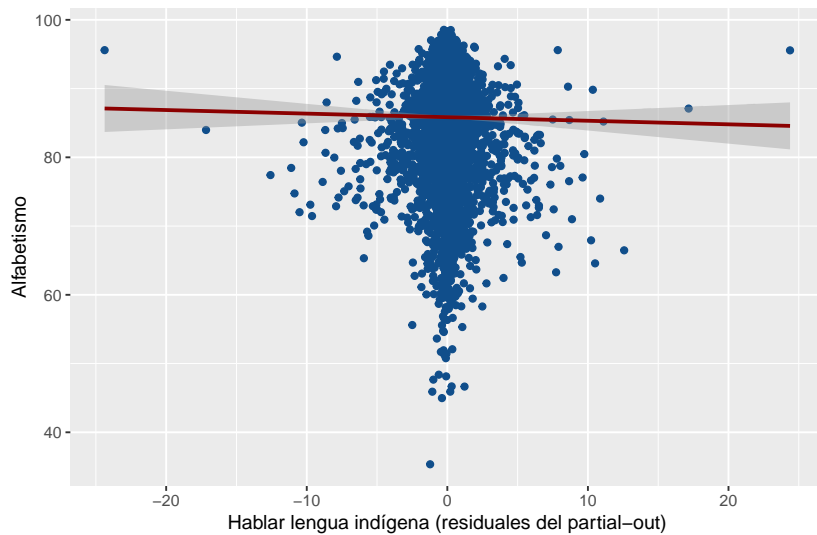


Table 5: Modelos de regresión lineal

	<i>Dependent variable:</i>		
	Pooled OLS	Alfabetismo Efectos Fijos	Efectos Aleatorios
	(1)	(2)	(3)
Hablan lengua indígena	−0.038*** (0.002)	−0.052*** (0.018)	−0.051*** (0.003)
Mujeres	−0.308*** (0.029)	−0.060 (0.052)	−0.223*** (0.032)
Edad	−0.191*** (0.029)	−0.205*** (0.070)	−0.122*** (0.032)
Asistencia escolar	−0.329*** (0.025)	−0.066** (0.030)	−0.231*** (0.025)
Nivel académico	7.426*** (0.433)	10.841*** (0.794)	9.252*** (0.540)
Rural	0.009*** (0.002)	0.026* (0.013)	0.004* (0.002)
Pobreza	0.077*** (0.024)	0.125*** (0.022)	0.077*** (0.020)
Pobreza extrema	−0.111*** (0.008)	−0.025** (0.011)	−0.072*** (0.008)
Log(Población)	−0.699*** (0.041)	−0.311 (0.314)	−0.754*** (0.054)
Rezago educativo	−0.286*** (0.014)	−0.104*** (0.019)	−0.212*** (0.015)
Carencia en salud	0.016*** (0.004)	−0.022*** (0.004)	−0.004 (0.003)
Carencia en seguridad social	0.019*** (0.006)	0.020*** (0.007)	0.023*** (0.005)
Carencia en vivienda	−0.044*** (0.007)	−0.017** (0.008)	−0.039*** (0.006)
Car. en servicios de vivienda	−0.004 (0.003)	0.013*** (0.004)	−0.00001 (0.003)
Carencia alimentaria	−0.024*** (0.004)	−0.023*** (0.003)	−0.021*** (0.003)
Ingreso inferior a LB	−0.016 (0.021)	−0.100*** (0.020)	−0.045** (0.018)
Desigualdad	1.509 (1.046)	2.248** (1.045)	1.026 (0.964)
Constant	112.575*** (2.272)		98.760*** (2.661)
Errores	Heterocedásticos	Cluster	Cluster
Efectos	No	FE	RE
Observations	4,890	4,890	4,890
R ²	0.878	0.688	0.825
Adjusted R ²	0.878	0.372	0.824
Residual Std. Error	2.874 (df = 4872)		

Note:

*p<0.1; **p<0.05; ***p<0.01

5.2. El impacto del Programa de Inclusión y Equidad Educativa

5.2.1. Diferencias en Diferencias

En la siguiente tabla de regresión, se muestran los resultados de múltiples especificaciones de Diferencias en Diferencias para el grupo de niñas y niños entre 6 y 11 años: el periodo anterior al programa corresponde a 2010; para 2015, el programa llevaba, aproximadamente, dos años de operación y, para 2020, se agregarían cuatro años más.

El primer modelo se refiere a una especificación de mínimos cuadrados ordinarios (OLS), el segundo incorpora efectos fijos por municipio y tiempo, el tercer modelo controla por el porcentaje de población que habla lengua indígena, el cuarto controla por el resto de variables socio-demográficas (solo se cuentan con cifras de pobreza multidimensional para 2010 y 2015) y el quinto es una especificación de efectos aleatorios. En todos los casos, se utilizan errores clusterizados por municipio.

Para verificar la credibilidad del supuesto de tendencias paralelas, se incorporan estimaciones de Diferencias en Diferencias para 2000 y 2005, cuando aún no entraba en operación el Programa de Inclusión y Equidad Educativa. Por lo tanto, esta intervención placebo del programa no tendría que mostrar impacto alguno para poder suponer que los municipios del programa se pudieron haber

comportado igual que los de control de 2010 a 2020.

Todas las especificaciones arrojan resultados similares y todos significativos al 1%. En promedio, la presencia del Programa de Inclusión y Equidad Educativa incrementó el alfabetismo de cada municipio en 3.424 puntos porcentuales de 2010 a 2020: separando este efecto, el incremento fue mayor entre 2014 y 2015 (2.898 puntos porcentuales) que el del periodo 2016 a 2019 (0.526 puntos porcentuales). Esto no implica que el efecto del programa se redujo: en la gráfica 8, se observa que estos 0.526 puntos son, más bien, un incremento en el efecto acumulado.

Al respecto del supuesto de tendencias paralelas, no se detectan diferencias significativas para los coeficientes del efecto placebo. Además, el coeficiente de diferencias en diferencias parece no cambiar mucho en signo, magnitud y significancia al incorporar controles.

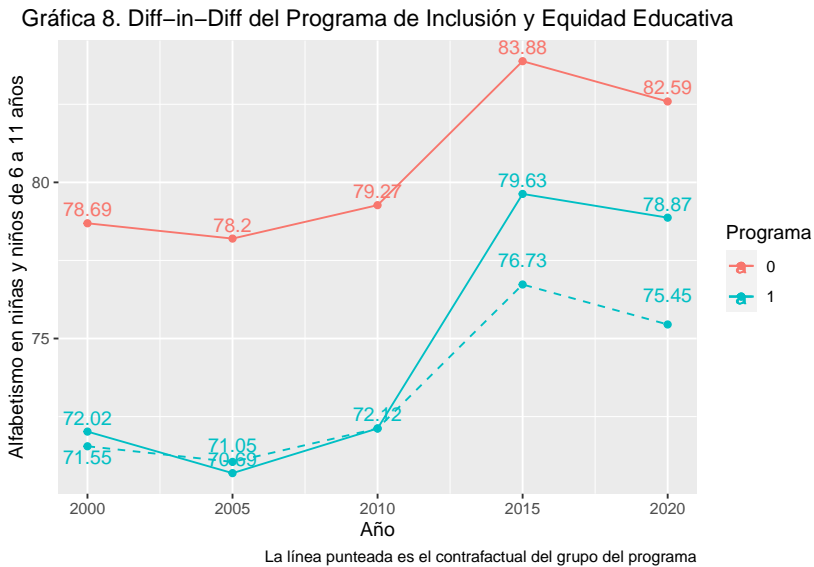
Table 6: Modelos de Diferencias en Diferencias

	<i>Dependent variable:</i>				
	Pooled OLS	Efectos Fijos	Alfabetismo Efectos Fijos	Efectos Fijos	Efectos Aleatorios
	(1)	(2)	(3)	(4)	(5)
Diff-in-Diff 2000	0.475 (0.380)	0.475 (0.380)	0.719* (0.382)		0.475 (0.380)
Diff-in-Diff 2005	-0.364 (0.391)	-0.364 (0.391)	-0.263 (0.392)		-0.364 (0.391)
Diff-in-Diff 2015	2.898*** (0.292)	2.898*** (0.292)	2.786*** (0.291)	1.685*** (0.291)	2.898*** (0.292)
Diff-in-Diff 2020	3.424*** (0.346)	3.424*** (0.346)	2.980*** (0.342)		3.424*** (0.346)
Hablan lengua indígena			-0.113*** (0.016)	-0.008 (0.016)	
Constant	78.693*** (0.245)				78.693*** (0.245)
Edades	6 a 11 años	6 a 11 años	6 a 11 años	6 a 11 años	6 a 11 años
Errores	Cluster	Cluster	Cluster	Cluster	Cluster
Efectos	No	FE	FE	FE	RE
Otros Controles	No	No	No	Sí	No
Observations	12,095	12,095	12,095	4,838	12,095
R ²	0.158	0.219	0.228	0.572	0.201
Adjusted R ²	0.158	0.023	0.034	0.137	0.201
Residual Std. Error	9.192 (df = 12085)				

Note:

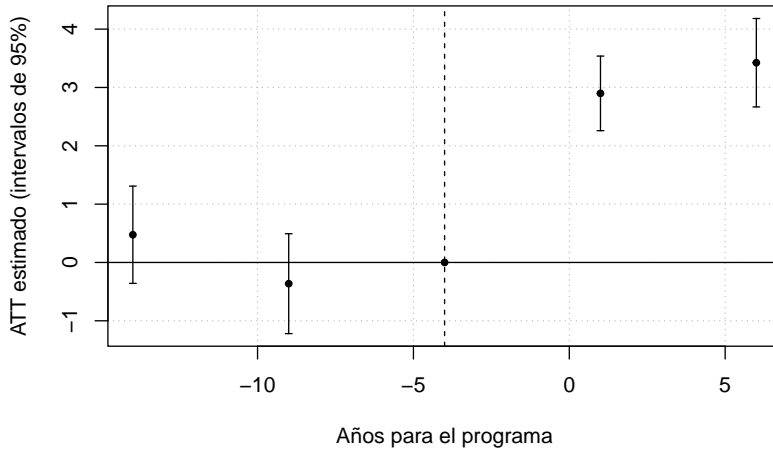
*p<0.1; **p<0.05; ***p<0.01

Como referencia visual de este efecto, a continuación, se muestra el comportamiento de los grupos entre 2000 y 2020, estimando el efecto del programa a partir de la curva del supuesto de tendencias paralelas: es decir, qué hubiera pasado si los municipios del programa no hubieran percibido estos subsidios (siguiendo la tendencia del grupo de control). La diferencia entre la curva punteada que representa el contrafactual y la curva azul que representa al grupo de tratamiento es el efecto del programa.



Asimismo, se presenta un gráfico de *event study* para visualizar el efecto del programa antes (placebo) y después de su aplicación en 2014 para cada año, incorporando intervalos de confianza al 95% para el ATT estimado de cada año.

Gráfica 9. Event Study del Programa de Inclusión y Equidad Educativa



5.2.2. Impacto por edades

También, es posible separar este efecto por edades, para identificar si el programa tuvo un impacto heterogéneo por generación. Estos resultados se muestran en la siguiente tabla utilizando efectos fijos por municipio y tiempo para estimar las Diferencias en Diferencias separando por edades entre 6 y 11 años.

Según estas estimaciones, el impacto del programa fue mayor para las niñas y niños que tenían 7 años: 5.741 puntos porcentuales entre 2010 y 2020. Los resultados para los grupos de 6 y 8 años son similares, con efectos de 4.220 y 4.494 puntos, respectivamente. Sin embargo, para las niñas y niños de 9 a 11 años, el impacto del

programa parece ser que se mantuvo o no incrementó respecto del efecto de 2014-2015.

Cabe destacar que el supuesto de tendencias paralelas podría ser inválido para el caso de 6 y 7 años, toda vez que muestran coeficientes estadísticamente significativos al 5% y 1% para los placebos de 2000 o 2005. Para los grupos de 8 a 11 años, por el contrario, no parece haber diferencias significativas antes de la aplicación del programa.

Table 7: Modelos de Diferencias en Diferencias por edades

	<i>Dependent variable:</i>					
	Alfabetismo					
	Efectos Fijos	Efectos Fijos	Efectos Fijos	Efectos Fijos	Efectos Fijos	Efectos Fijos
	(1)	(2)	(3)	(4)	(5)	(6)
Diff-in-Diff 2000	1.578** (0.668)	1.548** (0.658)	0.433 (0.500)	0.042 (0.369)	-0.239 (0.261)	-0.178 (0.187)
Diff-in-Diff 2005	2.162*** (0.708)	-1.843*** (0.687)	-0.737 (0.535)	-0.439 (0.370)	-0.031 (0.258)	-0.085 (0.202)
Diff-in-Diff 2015	2.210*** (0.489)	4.623*** (0.506)	3.829*** (0.401)	2.580*** (0.278)	1.669*** (0.188)	1.021*** (0.130)
Diff-in-Diff 2020	4.220*** (0.576)	5.741*** (0.589)	4.494*** (0.442)	2.660*** (0.296)	1.600*** (0.207)	1.083*** (0.139)
Edades	6 años	7 años	8 años	9 años	10 años	11 años
Errores	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Efectos	FE	FE	FE	FE	FE	FE
Observations	10,805	10,805	10,805	10,805	10,805	10,805
R ²	0.441	0.180	0.120	0.099	0.106	0.078
Adjusted R ²	0.300	-0.026	-0.101	-0.127	-0.119	-0.154
F Statistic (df = 8; 8636)	850.414***	236.199***	147.506***	118.869***	127.339***	91.072***

Note:

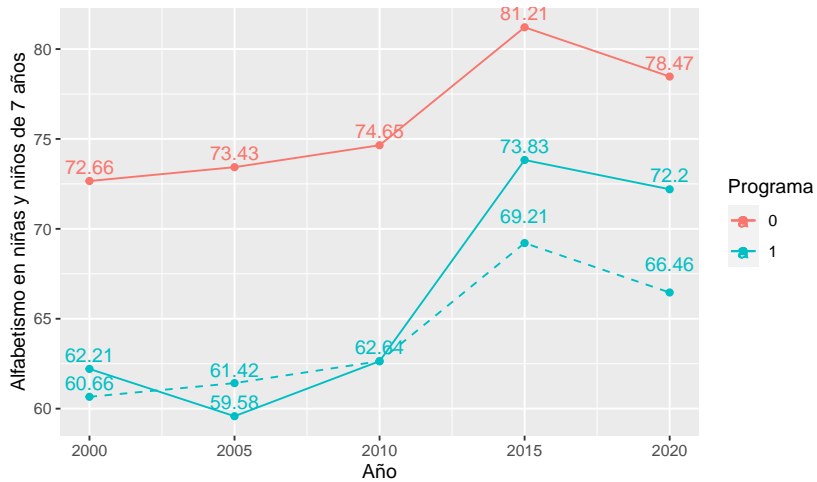
*p<0.1; **p<0.05; ***p<0.01

Gráfica 10. Diff-in-Diff del programa por edades:
población de 6 años



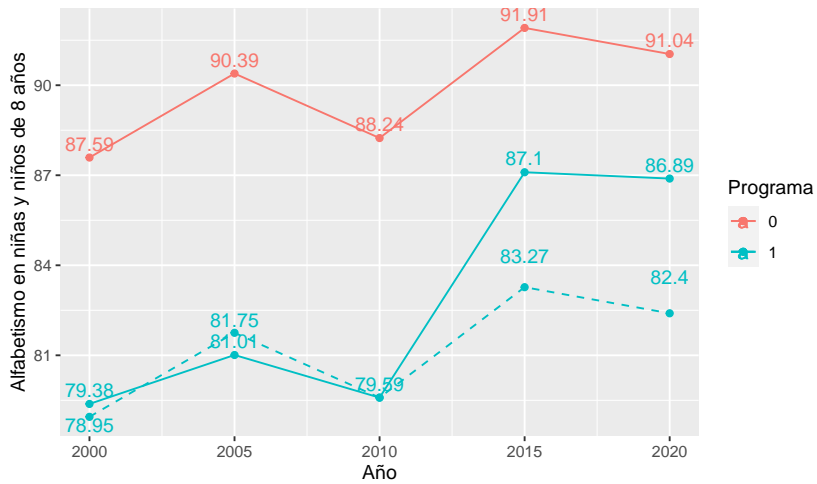
La línea punteada es el contrafactual del grupo del programa

Gráfica 11. Diff-in-Diff del programa por edades:
población de 7 años



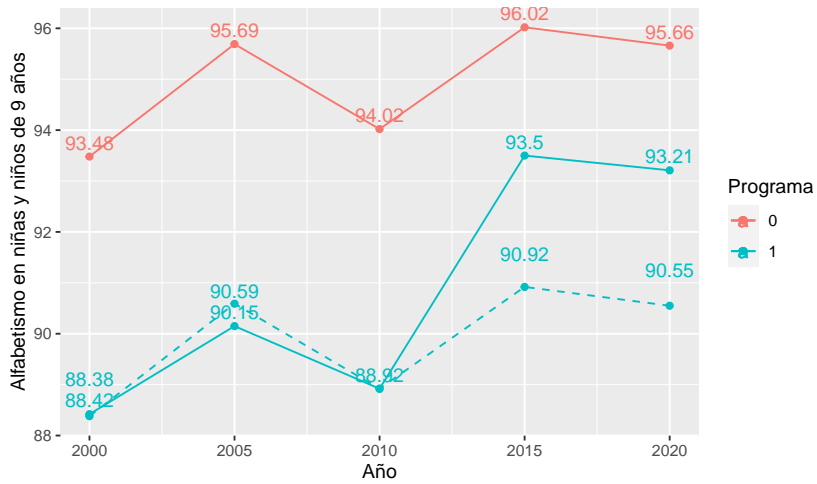
La línea punteada es el contrafactual del grupo del programa

Gráfica 12. Diff-in-Diff del programa por edades:
población de 8 años



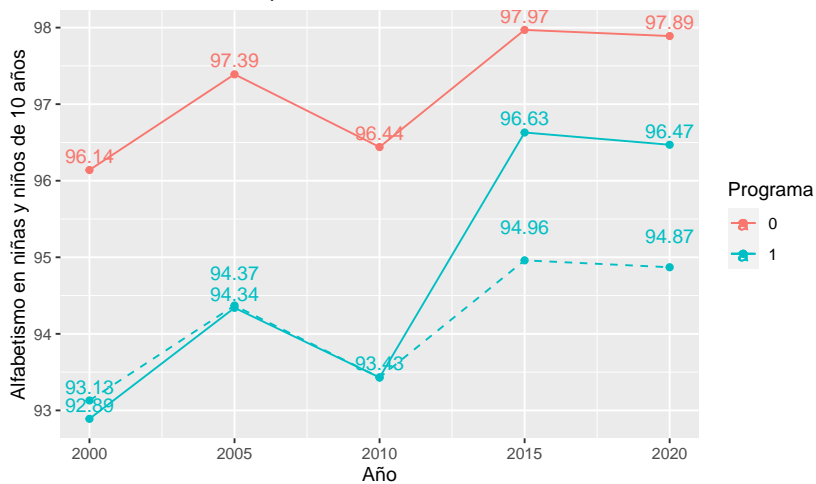
La línea punteada es el contrafactual del grupo del programa

Gráfica 13. Diff-in-Diff del programa por edades:
población de 9 años



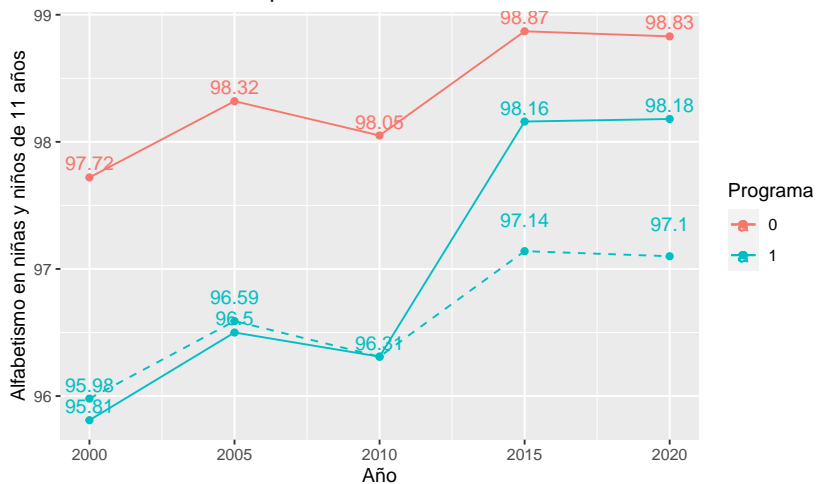
La línea punteada es el contrafactual del grupo del programa

Gráfica 14. Diff-in-Diff del programa por edades:
población de 10 años



La línea punteada es el contrafactual del grupo del programa

Gráfica 15. Diff-in-Diff del programa por edades:
población de 11 años

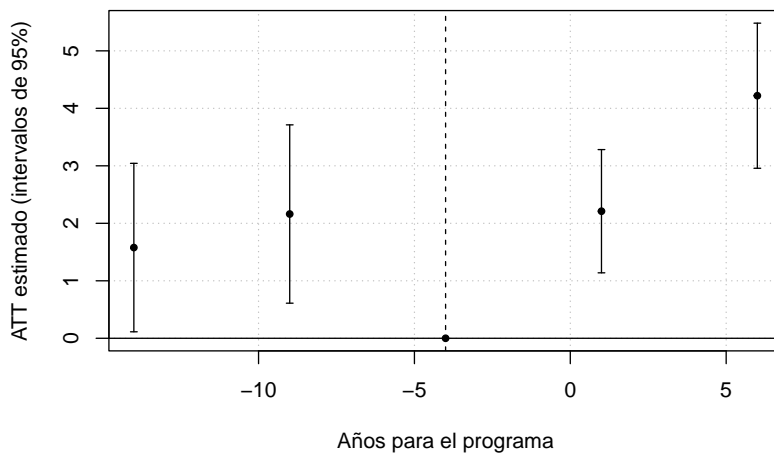


La línea punteada es el contrafactual del grupo del programa

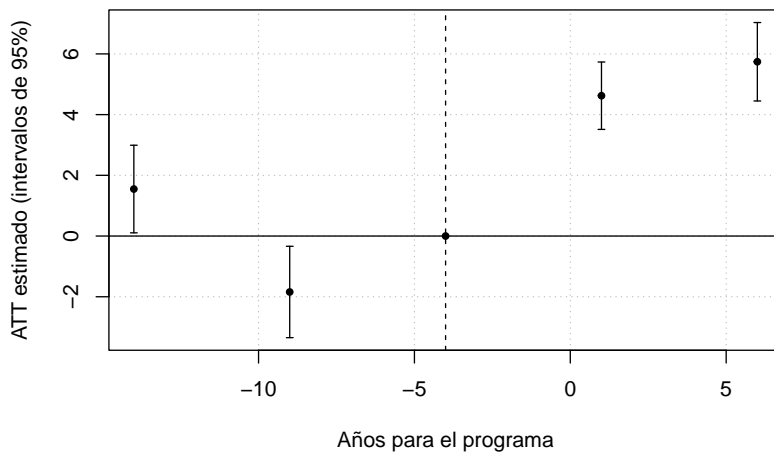
Nuevamente, se presentan los gráficos de *event study* para cada grupo de edad, con le objetivo de visualizar los coeficientes antes

y después de la aplicación del programa en 2014, incluyendo los intervalos de confianza al 95%.

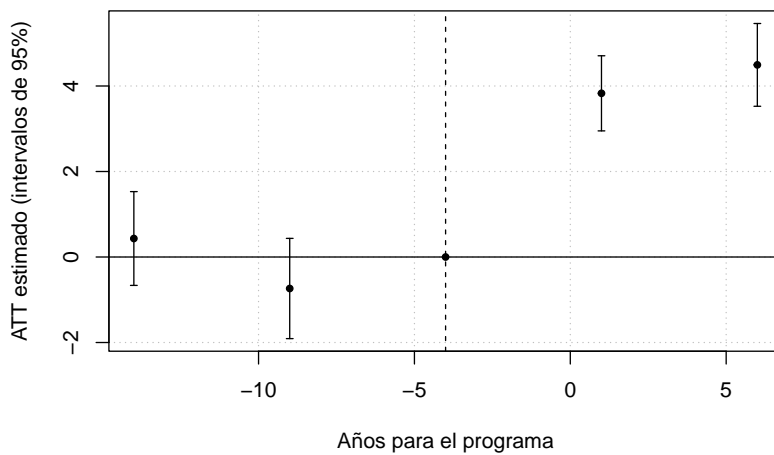
Gráfica 16. Event Study por edades: población de 6 años



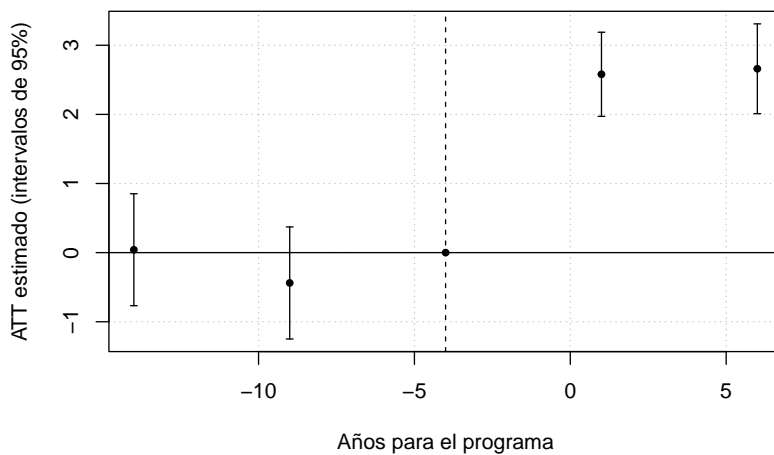
Gráfica 17. Event Study por edades: población de 7 años



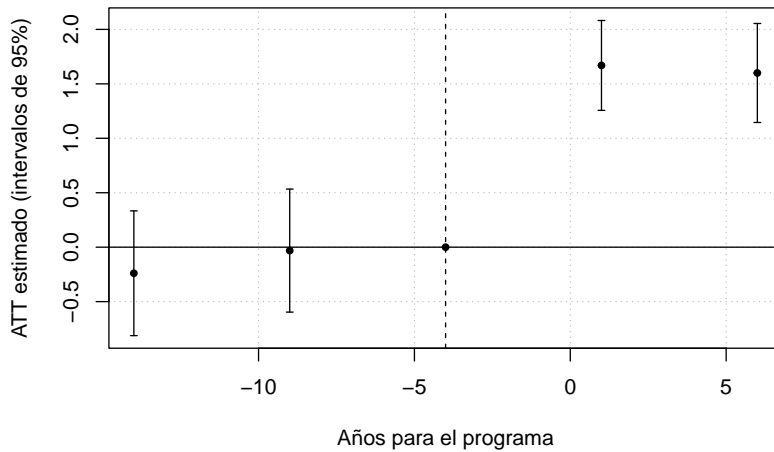
Gráfica 18. Event Study por edades: población de 8 años



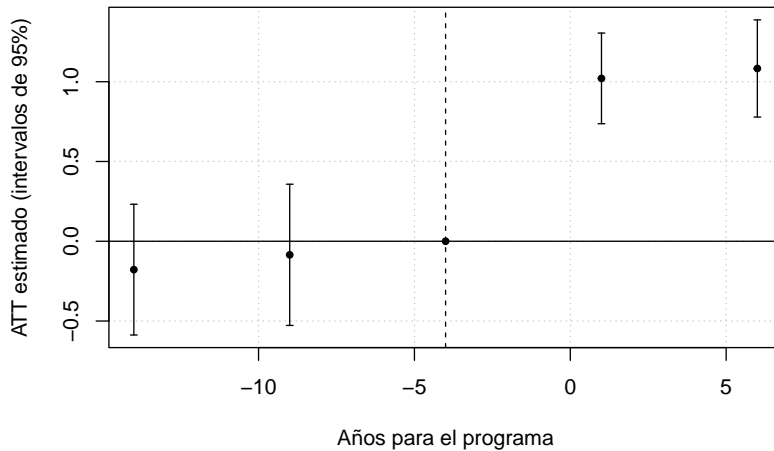
Gráfica 19. Event Study por edades: población de 9 años



Gráfica 20. Event Study por edades: población de 10 años



Gráfica 21. Event Study por edades: población de 11 años



Adicionalmente, se intento darle seguimiento a una misma cohorte de estudiantes a lo largo del tiempo: por ejemplo, cuando tenían

6 años en 2010, luego 11 años en 2015 y 16 años en 2020. Sin embargo, no fue posible estimar el efecto del programa por este medio, ya que, para periodos posteriores a la conclusión de la primaria, se alcanzan valores cercanos al 100% de alfabetización tanto en el grupo de control, como en el grupo del programa.

Dado que la variable dependiente está topada al 100%, no es posible construir correctamente el contrafactual de dobles diferencias: por construcción, la línea del contrafactual siempre tendría menor curvatura que la del grupo de control, arrojando siempre un coeficiente positivo, alto y con significancia.

A partir de lo anterior, podría surgir la siguiente duda: ¿entonces, para qué sirve el programa, si de todas formas las niñas y niños van a estar alfabetizados cuando cumplan más de 11 años? La respuesta se remite al marco teórico: porque no solo importa que las personas aprendan (en algún punto de su vida), sino que aprendan de forma temprana, porque existe el periodo crítico (antes de los 12 años) cuando adquieren mejor estos conocimientos y porque estas habilidades cognitivas ayudan al desarrollo de habilidades posteriores con el tiempo.

5.3. Mecanismo Causal y Efectos Indirectos

Idealmente, la mejor alternativa para probar el mecanismo causal entre los subsidios del programa y el alfabetismo sería utilizando cifras sobre i) infraestructura escolar y ii) calidad docente. Sin embargo, la información sobre infraestructura escolar es muy limitada: existe un *Censo de escuelas, maestros y alumnos de educación básica y especial*, pero solo se realizó en 2013 (SEP & INEGI 2013). Por otro lado, existe información sobre evaluación del Servicio Profesional Docente, pero es representativa por entidades federativas y no se encuentra a nivel municipal (INEE 2019).

Otra alternativa para medir, indirectamente, la calidad docente, sería utilizando los resultados de pruebas estandarizadas en alumnas y alumnos de primaria. Sin embargo, la prueba ENLACE se realizó hasta 2013 para, posteriormente, ser sustituida por PLANEA para el ciclo escolar 2014-2015, mientras que el programa nació en 2014.

Debido a lo anterior, se utilizaron múltiples proxys, principalmente, en materia de infraestructura escolar para información municipal de tipo panel. El porcentaje de asistencia escolar por municipio (de niñas y niños entre 6 y 11 años) se construyó utilizando los Conteos, Censos y Encuesta Intercensal del INEGI.

Por otro lado, se utilizaron las cifras del Formato 911 de la SEP, exclusivamente entre escuelas de educación indígena, para medir

la matrícula escolar, número de docentes, número de docentes con educación superior y número de aulas por escuela en municipios de control y tratamiento por año de 2000 a 2019.

Adicionalmente, se planteó la posibilidad de observar efectos indirectos del programa. Es posible que un mayor nivel alfabetismo entre estudiantes de escuelas del programa se tradujera en incrementos en el alfabetismo de sus madres y abuelas. Para ello, se utilizó el porcentaje de alfabetismo de mujeres entre 21 y 40 años, 40 y 60 años y 60 años o más en domicilios con niñas y niños de 6 a 11 años con información de los Conteos, Censos y Encuesta Intercensal del INEGI.

5.3.1. Mecanismo causal: Asistencia escolar

A continuación, se presentan los resultados del modelo de Diferencias en Diferencias sobre asistencia escolar entre las niñas y niños de 6 a 11 años, es decir, durante la primaria. La primera columna se refiere a una especificación OLS *pooled*, la segunda, a efectos fijos, la cuarta, efectos fijos controlando por población que habla lengua indígena, la quinta, por efectos fijos controlando por el resto de variables sociodemográficas y la sexta, efectos aleatorios. En todos los casos, se utilizan errores cluster por municipio.

Además, se incorpora la tabla con estimaciones por edad (6, 7, 8, 9, 10 y 11 años) para identificar efectos heterogéneos.

Según la estimación, el programa incrementó la asistencia escolar en 1.313 puntos porcentuales en los municipios cuyas escuelas de educación indígena percibieron subsidios. Aunque el coeficiente del placebo de 2000 es positivo y significativo al 5%.

La evidencia más robusta se refiere a los grupos de 9 y 11 años con efectos de 1.069 y 0.852 puntos porcentuales, respectivamente. En ambos casos, la evidencia parece señalar que se cumple el supuesto de tendencias paralelas.

Gráfica 22. Mecanismo causal: Diff-in-Diff sobre asistencia escolar



La línea punteada es el contrafactual del grupo del programa

Table 8: Modelos de Diferencias en Diferencias

	<i>Dependent variable:</i>				
	Pooled OLS	Asistencia escolar			Efectos Aleatorios
	(1)	Efectos Fijos	Efectos Fijos	Efectos Fijos	(5)
	(1)	(2)	(3)	(4)	(5)
Diff-in-Diff 2000	0.257** (0.118)	0.257** (0.118)	0.356*** (0.120)		0.257** (0.118)
Diff-in-Diff 2005	-0.258* (0.132)	-0.258* (0.132)	-0.218 (0.132)		-0.258* (0.132)
Diff-in-Diff 2015	0.535*** (0.086)	0.535*** (0.086)	0.490*** (0.086)	0.318*** (0.086)	0.535*** (0.086)
Diff-in-Diff 2020	1.313*** (0.143)	1.313*** (0.143)	1.133*** (0.143)		1.313*** (0.143)
Hablan lengua indígena			-0.046*** (0.006)	-0.021*** (0.006)	
Constant	97.490*** (0.048)				97.490*** (0.048)
Edades	6 a 11 años	6 a 11 años	6 a 11 años	6 a 11 años	6 a 11 años
Errores	Cluster	Cluster	Cluster	Cluster	Cluster
Efectos	No	FE	FE	FE	RE
Otros Controles	No	No	No	Sí	No
Observations	12,095	12,095	12,095	4,838	12,095
R ²	0.112	0.181	0.192	0.400	0.153
Adjusted R ²	0.112	-0.025	-0.010	-0.209	0.152
Residual Std. Error	2.723 (df = 12085)				

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 9: Modelos de Diferencias en Diferencias por edades

	<i>Dependent variable:</i>					
	Asistencia escolar					
	Efectos Fijos	Efectos Fijos	Efectos Fijos	Efectos Fijos	Efectos Fijos	Efectos Fijos
	(1)	(2)	(3)	(4)	(5)	(6)
Diff-in-Diff 2000	0.814*** (0.281)	0.496*** (0.187)	-0.035 (0.148)	-0.108 (0.149)	-0.144 (0.157)	-0.098 (0.192)
Diff-in-Diff 2005	-0.584* (0.319)	-0.169 (0.226)	-0.558*** (0.206)	-0.103 (0.219)	-0.420** (0.192)	-0.365 (0.228)
Diff-in-Diff 2015	1.232*** (0.187)	0.686*** (0.136)	0.301*** (0.116)	0.269** (0.108)	0.038 (0.110)	0.099 (0.122)
Diff-in-Diff 2020	2.339*** (0.252)	1.780*** (0.190)	1.177*** (0.177)	1.069*** (0.163)	0.837*** (0.172)	0.852*** (0.178)
Edades	6 años	7 años	8 años	9 años	10 años	11 años
Errores	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Efectos	FE	FE	FE	FE	FE	FE
Observations	10,805	10,805	10,805	10,805	10,805	10,805
R ²	0.086	0.109	0.133	0.116	0.119	0.100
Adjusted R ²	-0.143	-0.115	-0.085	-0.106	-0.102	-0.126
F Statistic (df = 8; 8636)	101.791***	131.509***	164.964***	141.216***	145.628***	120.222***

Note:

*p<0.1; **p<0.05; ***p<0.01

5.3.2. Mecanismo causal: Matrícula escolar

Al respecto de la matrícula escolar, se cuenta con información anual de 2000 a 2019 por escuela de educación indígena entre los municipios de control y del programa. La siguiente tabla muestra la estimación del modelo de Diferencias en Diferencias donde la primera columna es un OLS *pooled* para toda la línea de tiempo de 2000 a 2019, la segunda incluye efectos fijos y la tercera, efectos aleatorios. En todos los casos, se utilizan errores cluster por municipio.

Se observa un salto acumulado de 6.783 con significancia de apenas 10% y, al observar la línea de tiempo hacia atrás, no se encuentran diferencias significativas en los placebo. De hecho, en la gráfica se observa con mayor claridad que la curva del contrafactual parecería ser bastante “paralela” a la curva del grupo del programa hasta el salto de 2.502 estudiantes por escuela en 2014, cuando inició la operación del programa.

Gráfica 23. Mecanismo causal: Diff-in-Diff sobre matrícula escolar



Table 10: Modelos de Diferencias en Diferencias

	<i>Dependent variable:</i>		
	Matrícula escolar		
	Pooled OLS	Efectos Fijos	Efectos Aleatorios
	(1)	(2)	(3)
Diff-in-Diff 2000	-5.056 (4.028)	-5.056 (4.028)	-5.056 (4.028)
Diff-in-Diff 2001	-4.169 (3.762)	-4.169 (3.762)	-4.169 (3.762)
Diff-in-Diff 2002	-4.735 (3.637)	-4.735 (3.637)	-4.735 (3.637)
Diff-in-Diff 2003	-5.711 (3.491)	-5.711 (3.491)	-5.711 (3.491)
Diff-in-Diff 2004	-4.347 (3.135)	-4.347 (3.135)	-4.347 (3.135)
Diff-in-Diff 2005	-2.034 (2.856)	-2.034 (2.856)	-2.034 (2.856)
Diff-in-Diff 2006	-2.167 (2.824)	-2.167 (2.824)	-2.167 (2.824)
Diff-in-Diff 2007	0.479 (2.132)	0.479 (2.132)	0.479 (2.132)
Diff-in-Diff 2008	1.654 (1.945)	1.654 (1.945)	1.654 (1.945)
Diff-in-Diff 2009	2.088 (1.592)	2.088 (1.592)	2.088 (1.592)
Diff-in-Diff 2010	1.888 (1.296)	1.888 (1.296)	1.888 (1.296)
Diff-in-Diff 2011	0.922 (1.126)	0.922 (1.126)	0.922 (1.126)
Diff-in-Diff 2012	0.697 (0.753)	0.697 (0.753)	0.697 (0.753)
Diff-in-Diff 2014	2.502* (1.391)	2.502* (1.391)	2.502* (1.391)
Diff-in-Diff 2015	2.632* (1.543)	2.632* (1.543)	2.632* (1.543)
Diff-in-Diff 2016	3.374** (1.442)	3.374** (1.442)	3.374** (1.442)
Diff-in-Diff 2017	3.751** (1.644)	3.751** (1.644)	3.751** (1.644)
Diff-in-Diff 2018	5.460 (4.294)	5.460 (4.294)	5.460 (4.294)
Diff-in-Diff 2019	6.783* (3.906)	6.783* (3.906)	6.783* (3.906)
Errores	Cluster	Cluster	Cluster
Observations	15,080	15,080	15,080
R ²	0.015	0.090	0.086
F Statistic	5.759*** (df = 39; 15040)	36.977*** (df = 38; 14288)	1,408.004***

Note:

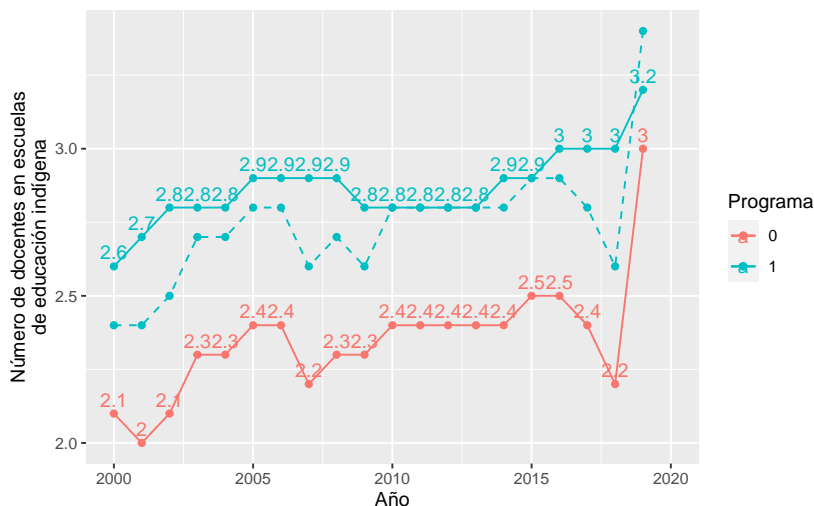
*p<0.1; **p<0.05; ***p<0.01

5.3.3. Mecanismo causal: Número de docentes

Al respecto del número de docentes en escuelas de educación indígena, la siguiente tabla muestra los resultados del modelo. De nuevo, la primer columna es un OLS *pooled* para toda la línea de tiempo, la segunda incluye efectos fijos y la tercera, efectos aleatorios. En todos los casos, se utilizan errores cluster por municipio.

En este caso, pareciera haber una caída importante en el grupo de control entre 2017 y 2018, mientras que el grupo del programa se mantuvo constante en el número de docentes. Derivado de esto, se observa un coeficiente de 0.397 significativo al 5% en 2018. Sin embargo, para 2019, pareciera haber una recuperación por parte del grupo de control igualando, nuevamente, a las escuelas del programa.

Gráfica 24. Mecanismo causal: Diff-in-Diff sobre el número de docentes



La línea punteada es el contrafactual del grupo del programa

Table 11: Modelos de Diferencias en Diferencias

	<i>Dependent variable:</i>		
	Pooled OLS	Número de docentes Efectos Fijos	Efectos Aleatorios
	(1)	(2)	(3)
Diff-in-Diff 2000	0.176 (0.184)	0.176 (0.184)	0.176 (0.184)
Diff-in-Diff 2001	0.249 (0.180)	0.249 (0.180)	0.249 (0.180)
Diff-in-Diff 2002	0.250 (0.174)	0.250 (0.174)	0.250 (0.174)
Diff-in-Diff 2003	0.098 (0.168)	0.098 (0.168)	0.098 (0.168)
Diff-in-Diff 2004	0.091 (0.157)	0.091 (0.157)	0.091 (0.157)
Diff-in-Diff 2005	0.096 (0.152)	0.096 (0.152)	0.096 (0.152)
Diff-in-Diff 2006	0.104 (0.150)	0.104 (0.150)	0.104 (0.150)
Diff-in-Diff 2007	0.254** (0.129)	0.254** (0.129)	0.254** (0.129)
Diff-in-Diff 2008	0.148 (0.115)	0.148 (0.115)	0.148 (0.115)
Diff-in-Diff 2009	0.163 (0.112)	0.163 (0.112)	0.163 (0.112)
Diff-in-Diff 2010	0.039 (0.089)	0.039 (0.089)	0.039 (0.089)
Diff-in-Diff 2011	0.020 (0.084)	0.020 (0.084)	0.020 (0.084)
Diff-in-Diff 2012	0.059 (0.077)	0.059 (0.077)	0.059 (0.077)
Diff-in-Diff 2014	0.058 (0.092)	0.058 (0.092)	0.058 (0.092)
Diff-in-Diff 2015	0.029 (0.097)	0.029 (0.097)	0.029 (0.097)
Diff-in-Diff 2016	0.085 (0.096)	0.085 (0.096)	0.085 (0.096)
Diff-in-Diff 2017	0.179* (0.108)	0.179* (0.108)	0.179* (0.108)
Diff-in-Diff 2018	0.397** (0.192)	0.397** (0.192)	0.397** (0.192)
Diff-in-Diff 2019	-0.190 (0.208)	-0.190 (0.208)	-0.190 (0.208)
Errores	Cluster	Cluster	Cluster
Observations	15,080	15,080	15,080
R ²	0.010	0.020	0.020
F Statistic	3.989*** (df = 39; 15040)	7.837*** (df = 38; 14288)	304.183***

Note:

* p<0.1; ** p<0.05; *** p<0.01

5.3.4. Mecanismo causal: Docentes con educación superior

Sobre los docentes con educación superior y siguiendo la misma estructura de la tabla de regresión de los casos anteriores, no se observa un impacto vinculado con la presencia del programa.

Solo se observa un coeficiente de 0.347 docentes con educación superior significativo al 5% en 2018; sin embargo, para 2019, ambos grupos se vuelven a igualar. El comportamiento es bastante similar al del caso anterior del número de docentes en general.

Gráfica 25. Mecanismo causal: Diff-in-Diff sobre el número de docentes con educación superior

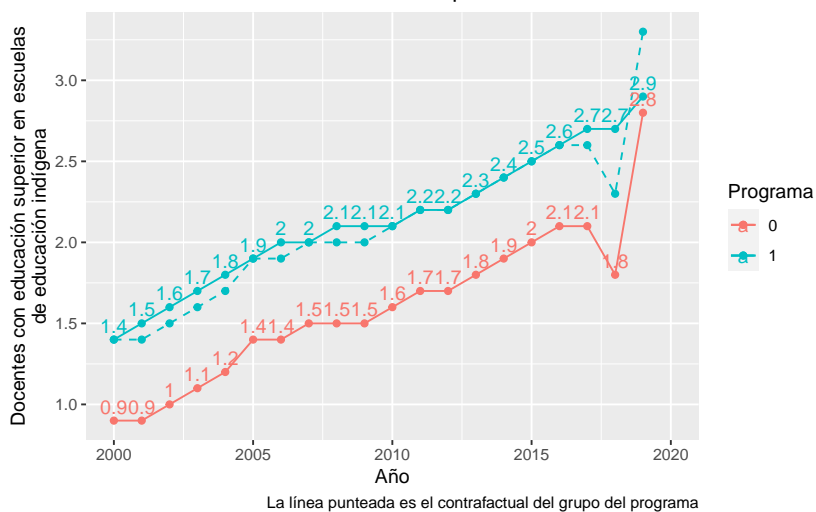


Table 12: Modelos de Diferencias en Diferencias

	<i>Dependent variable:</i>		
	Docentes con educación superior		
	Pooled OLS	Efectos Fijos	Efectos Aleatorios
	(1)	(2)	(3)
Diff-in-Diff 2000	0.020 (0.153)	0.020 (0.153)	0.020 (0.153)
Diff-in-Diff 2001	0.095 (0.145)	0.095 (0.145)	0.095 (0.145)
Diff-in-Diff 2002	0.088 (0.130)	0.088 (0.130)	0.088 (0.130)
Diff-in-Diff 2003	0.081 (0.127)	0.081 (0.127)	0.081 (0.127)
Diff-in-Diff 2004	0.117 (0.129)	0.117 (0.129)	0.117 (0.129)
Diff-in-Diff 2005	0.008 (0.114)	0.008 (0.114)	0.008 (0.114)
Diff-in-Diff 2006	0.062 (0.111)	0.062 (0.111)	0.062 (0.111)
Diff-in-Diff 2007	0.046 (0.113)	0.046 (0.113)	0.046 (0.113)
Diff-in-Diff 2008	0.012 (0.101)	0.012 (0.101)	0.012 (0.101)
Diff-in-Diff 2009	0.054 (0.094)	0.054 (0.094)	0.054 (0.094)
Diff-in-Diff 2010	0.025 (0.077)	0.025 (0.077)	0.025 (0.077)
Diff-in-Diff 2011	-0.001 (0.063)	-0.001 (0.063)	-0.001 (0.063)
Diff-in-Diff 2012	0.029 (0.059)	0.029 (0.059)	0.029 (0.059)
Diff-in-Diff 2014	0.065 (0.065)	0.065 (0.065)	0.065 (0.065)
Diff-in-Diff 2015	-0.009 (0.073)	-0.009 (0.073)	-0.009 (0.073)
Diff-in-Diff 2016	-0.041 (0.085)	-0.041 (0.085)	-0.041 (0.085)
Diff-in-Diff 2017	0.056 (0.089)	0.056 (0.089)	0.056 (0.089)
Diff-in-Diff 2018	0.347** (0.159)	0.347** (0.159)	0.347** (0.159)
Diff-in-Diff 2019	-0.389** (0.187)	-0.389** (0.187)	-0.389** (0.187)
Errores	Cluster	Cluster	Cluster
Observations	15,080	15,080	15,080
R ²	0.053	0.185	0.178
F Statistic	21.418*** (df = 39; 15040)	85.486*** (df = 38; 14288)	3,259.918***

Note:

*p<0.1; **p<0.05; ***p<0.01

5.3.5. Mecanismo causal: Número de aulas

Utilizando ahora el número de aulas como proxy de infraestructura escolar, la siguiente tabla muestra los resultados del modelo de regresión.

En promedio, la presencia del programa está asociada con un incremento de 0.541 aulas por escuela de educación indígena. La diferencia se empieza a hacer notoria desde 2015, haciéndose cada vez mayor con el paso del tiempo. Previo a la aplicación del programa, no se observan diferencias significativas.

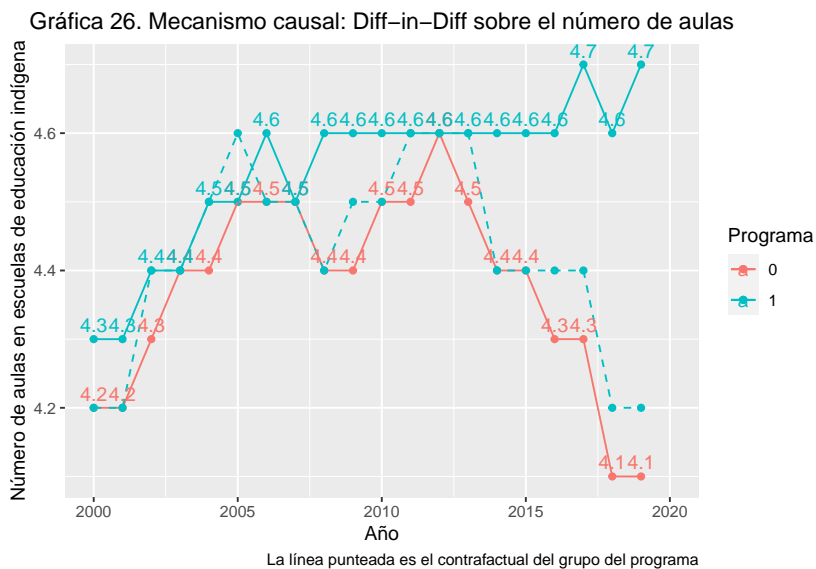


Table 13: Modelos de Diferencias en Diferencias

	<i>Dependent variable:</i>		
	Pooled OLS	Número de aulas	
		Efectos Fijos	Efectos Aleatorios
	(1)	(2)	(3)
Diff-in-Diff 2000	0.079 (0.142)	0.079 (0.142)	0.079 (0.142)
Diff-in-Diff 2001	0.112 (0.142)	0.112 (0.142)	0.112 (0.142)
Diff-in-Diff 2002	0.028 (0.139)	0.028 (0.139)	0.028 (0.139)
Diff-in-Diff 2003	0.030 (0.134)	0.030 (0.134)	0.030 (0.134)
Diff-in-Diff 2004	0.0004 (0.130)	0.0004 (0.130)	0.0004 (0.130)
Diff-in-Diff 2005	-0.053 (0.124)	-0.053 (0.124)	-0.053 (0.124)
Diff-in-Diff 2006	0.038 (0.126)	0.038 (0.126)	0.038 (0.126)
Diff-in-Diff 2007	0.018 (0.113)	0.018 (0.113)	0.018 (0.113)
Diff-in-Diff 2008	0.141 (0.108)	0.141 (0.108)	0.141 (0.108)
Diff-in-Diff 2009	0.136 (0.095)	0.136 (0.095)	0.136 (0.095)
Diff-in-Diff 2010	0.092 (0.085)	0.092 (0.085)	0.092 (0.085)
Diff-in-Diff 2011	0.040 (0.072)	0.040 (0.072)	0.040 (0.072)
Diff-in-Diff 2012	-0.028 (0.071)	-0.028 (0.071)	-0.028 (0.071)
Diff-in-Diff 2014	0.143 (0.099)	0.143 (0.099)	0.143 (0.099)
Diff-in-Diff 2015	0.175* (0.104)	0.175* (0.104)	0.175* (0.104)
Diff-in-Diff 2016	0.249** (0.097)	0.249** (0.097)	0.249** (0.097)
Diff-in-Diff 2017	0.362*** (0.103)	0.362*** (0.103)	0.362*** (0.103)
Diff-in-Diff 2018	0.433** (0.216)	0.433** (0.216)	0.433** (0.216)
Diff-in-Diff 2019	0.541** (0.218)	0.541** (0.218)	0.541** (0.218)
Errores	Cluster	Cluster	Cluster
Observations	15,080	15,080	15,080
R ²	0.003	0.015	0.014
F Statistic	1.007 (df = 39; 15040)	5.777*** (df = 38; 14288)	220.125***

Note:

*p<0.1; **p<0.05; ***p<0.01

5.3.6. Efectos indirectos: Alfabetización de madres y abuelas

Finalmente, se hizo un esfuerzo por visualizar posibles efectos indirectos vinculados con la aplicación del programa. En específico, se analizó el alfabetismo de madres y abuelas medido como el porcentaje de alfabetismo de mujeres en hogares con niñas y niños de 6 a 11 años entre municipios de control y del programa, utilizando las cifras del Conteo, Censos y Encuesta Intercensal del INEGI.

La siguiente tabla muestra los resultados del modelo de Diferencias en Diferencias por grupos de edad de las mujeres: 21 a 40 años, 40 a 60 años y 60 o más años. En todos los casos, se utilizaron efectos fijos y errores cluster por municipio. Sin embargo, no se observa ningún tipo de efecto vinculado con la aplicación del programa para ningún grupo de edad.

Table 14: Modelos de Diferencias en Diferencias

	<i>Dependent variable:</i>		
	Efectos Fijos	Alfabetismo Efectos Fijos	Efectos Fijos
	(1)	(2)	(3)
Diff-in-Diff 2000	0.552*** (0.139)	0.241* (0.127)	0.147 (0.119)
Diff-in-Diff 2005	0.559*** (0.178)	0.167 (0.193)	-0.008 (0.157)
Diff-in-Diff 2015	0.076 (0.117)	0.154 (0.140)	0.055 (0.140)
Diff-in-Diff 2020	0.062 (0.112)	0.189 (0.123)	0.090 (0.121)
Edades	21 a 40 años	40 a 60 años	60 o más años
Errores	Cluster	Cluster	Cluster
Efectos	FE	FE	FE
Observations	12,262	12,256	12,221
R ²	0.075	0.069	0.020
Adjusted R ²	-0.159	-0.167	-0.229
F Statistic	98.758*** (df = 8; 9787)	90.190*** (df = 8; 9781)	25.025*** (df = 8; 9746)
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01	

6. Conclusiones

“Living on 99 cents a day means you have limited acces to information -newspapers, television and books all cost money- and so you often just don’t know certain facts that the rest of the world take as given, like, for example, that vaccines can stop your child from getting measles. It means living in a world whose institutions are not built for someone like you.” (Banerjee y Dufló 2011, p. ix).

El proceso de alfabetización juega un papel fundamental para el pleno desarrollo de las personas. Uno de los sectores que ven más vulnerado su acceso a la educación es la población indígena. Este estudio se plantea una primer hipótesis de que, en parte, hablar una lengua indígena tiene efectos negativos sobre la alfabetización, ya que el español y las lenguas indígenas pertenecen a familias lingüísticas diferentes, lo que dificulta los procesos de aprendizaje de la morfología y sintaxis del español y se traduce en mayores niveles de analfabetismo. La segunda hipótesis que se analiza es que el componente de educación indígena del Programa de Inclusión y Equidad Educativa de 2014 a 2019 tuvo un impacto positivo sobre el proceso de alfabetización de la población indígena.

Utilizando información a nivel municipal provenientes del Censos de Población 2010 y la Encuesta Intercensal 2015, se estimaron modelos de regresión para datos panel: efectos fijos y efec-

tos aleatorios. Sobre la primer hipótesis, se concluye que, *ceteris paribus*, un incremento de un punto en el porcentaje de la población que habla lengua indígena causa un decremento promedio de 0.052 puntos en el alfabetismo de los municipios de México, independientemente del nivel de pobreza, desigualdad, género, urbanización, población, rezago educativo, escolaridad, asistencia escolar y bienestar económico.

Al respecto de la segunda hipótesis, se utilizó información a nivel municipal sobre las escuelas de educación indígena que percibieron recursos por parte del Programa de Inclusión y Equidad Educativa entre 2014 y 2019 y se anexan las cifras del Censo de Población 2020, así como del Censo de Población 2000 y el Conteo de Población 2005 para probar el supuesto de tendencias paralelas.

Utilizando múltiples especificaciones para estimar modelos de Diferencias en Diferencias, se concluye que, *ceteris paribus*, la presencia del Programa de Inclusión y Equidad Educativa generó un incremento de 3.424 puntos porcentuales en la media del alfabetismo de niñas y niños entre 6 y 11 años de los municipios con escuelas de educación indígena en México entre 2010 y 2020. El mayor impacto se parece concentrar entre las niñas y niños de 6 a 8 años (entre 4.220 y 5.741 puntos porcentuales), pero la evidencia más robusta se refiere al efecto para la población de 8 a 11 años: 4.494, 2.660, 1.600 y 1.083 puntos porcentuales,

respectivamente.

Para cada especificación de modelos, se incorporaron efectos placebo entre 2000 y 2010 con el objetivo de aportar evidencia sobre el supuesto de tendencias paralelas. En general, el supuesto pareciera creíble a la luz de que ninguno de los coeficientes del placebo mostraron significancia estadística (a excepción de las regresiones para 6 y 7 años).

En términos del mecanismo causal, este efecto podría pasar por un aumento en la infraestructura escolar o por una mejora en la calidad docente; sin embargo, no se cuenta con información municipal de tipo panel para probar el efecto del programa sobre dichos ámbitos con precisión. Sin embargo, se utilizaron distintas variables proxy para probar el mecanismo causal.

Se encontró que la aplicación del programa está vinculada con un incremento 1.313 puntos porcentuales en la asistencia escolar por municipio e incrementos de 6.783 alumnas y alumnos en la matrícula escolar y 0.541 aulas por escuela de educación indígena en municipios del programa.

Lo que esta evidencia señala es que el sistema educativo mexicano debe adaptarse para permitir una mayor inclusión de la población indígena en los procesos de alfabetización, sin que dejen de lado su herencia cultural. De hecho, la evidencia tiende a señalar que la enseñanza básica se tendría que basar en la lengua materna

(Bühmann y Trudell 2007; Ganuza y Hedman 2017; Nkonde, et al. 2018), mientras que el español se aprenda solo como segunda lengua durante los primeros años. Para ello, es esencial fortalecer a las escuelas de educación indígena: por ejemplo, a través de intervenciones como el Programa de Inclusión y Equidad Educativa. Sin embargo, si bien este programa originalmente se fragmentó en cuatro programas diferentes en 2020, el componente para educación indígena (ahora Programa de Atención a la Diversidad de la Educación Indígena) desapareció para el ejercicio 2021. La educación es una condición necesaria para el desarrollo, pero no se puede garantizar sin presupuesto (sobre todo, donde más se necesita).

Los procesos educativos deben ser plurales y adaptables al contexto de cada región; esto no solo representa una condición necesaria para el desarrollo de sectores de la población históricamente discriminados, sino que es una condición mínima de justicia social para garantizar la igualdad en el acceso a los derechos sociales. Intervenciones relativamente simples como la del Programa de Inclusión y Equidad Educativa pueden mejorar los aprendizajes del español durante el periodo crítico (previo a los 12 años).

7. Bibliografía

Albó, X., & at, e. (2009). Atlas sociolingüístico de pueblos indígenas en América Latina. Obtenido de https://www.unicef.org/honduras/tomo_1_atlas.pdf

Angrist, Joshua y Pischke, Jörn-Steffen. (2009). *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.

Banerjee, Abhijit y Duflo, Esther. (2011). *Poor Economics: A radical rethinking of the way to fight global poverty*. New York, NY: PublicAffairs.

Barret, P., et al. (2019). *The Impact of School Infrastructure on Learning: A Synthesis of the Evidence*. Washington, DC: The World Bank.

Behrman, J., et al. (2015). *Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools*. *Journal of Political Economy*, Vol. 123, No. 2, pp. 325-364.

Bruns, B. y Luque, J. (2015). *Profesores Excelentes: Cómo mejorar el aprendizaje en América Latina y el Caribe*. Washington, DC: The World Bank.

Bühmann, Dörthe y Trudell, Barbara. (2007). *Mother tongue matters: local language as a key to effective learning*. Francia: UNESCO.

Cárdenas, V. (2010). La relación entre semántica y sintaxis desde la perspectiva de la producción de lenguaje escrito. Obtenido de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-12002010000100008

Carneiro, P. et al. (2019). *Parental Beliefs, Investments and Child Development: Evidence from a Large-Scale Experiment*. IZA Discussion Paper 12506.

Chetty, R., et al. (2011). *How does your kindergarten classroom affect your earnings? Evidence from Project Star*. NBER Working Paper 16381.

Chomsky, N. (2017). *50 years later: a conversation about the biological study of language with Noam Chomsky. An interview with Patrick Trettenbrein*. Biolinguistics. Vol. 11.

CONEVAL. (2015). Medición de la Pobreza: La Cohesión Social. Obtenido de https://www.coneval.org.mx/Medicion/Paginas/Cohesion_Social.aspx

CONEVAL. (2018). *Estudio Diagnóstico del Derecho a la Educación 2018*. México: CONEVAL.

CONEVAL. (2021). *Inventario CONEVAL de Programas y Acciones Federales de Desarrollo Social*. Consultado el 09/08/2021 en <https://www.coneval.org.mx/evaluacion/ipfe/Paginas/default.aspx>

Cunha, F., & Heckman, J. (2007). *The technology of skill forma-*

tion. *American Economic Review*, 97(2), 31-47.

De Hoyos, R. Estrada, R. y Vargas, M. (2018). *Predicting individual wellbeing through test scores: evidence from a national assessment in Mexico*. Policy Research Working Paper 8459, The World Bank.

De Hoyos, R., Holland, P. y Troiano, S. (2015). *Understanding the trends in learning outcomes in Argentina, 2000 to 2012*. Policy Research Working Paper 7518, The World Bank.

DOF. (28 de diciembre de 2013). Acuerdo número 711 por el que se emiten las Reglas de Operación del Programa para la Inclusión y la Equidad Educativa.

DOF. (28 de febrero de 2019). Acuerdo número 04/02/19 por el que se emiten las Reglas de Operación del Programa para la Inclusión y la Equidad Educativa para el ejercicio fiscal 2019.

Duarte, Moreno y Gargiulo. (2011). *Infraestructura Escolar y Aprendizajes en la Educación Básica Latinoamericana: Un análisis a partir del SERCE*. Banco Interamericano de Desarrollo.

Elango, S., et al. (2016). *Early Childhood Education*. Economics of Means-Tested Transfer Programs in the United States. Vol. 2, Moffitt.

Evans, M., et al. (2010). *Family scholarly culture and educational success: books and schooling in 27 nations*. Research in Social Stratification and Mobility.

- Ganuza, Natalia y Hedman, Christina. (2017). *The impact of mother tongue instruction on the development of biliteracy: evidence from Somali-Swedish bilinguals*. Applied Linguistics, 1-25. Oxford University Press.
- Gargarella, Roberto. (1999). *Las teorías de la justicia después de Rawls*. España: Paidós.
- Heckman, J. et al. (1998). *Characterizing selection bias using experimental data*. Econometrica, Vol. 66, No. 5, pp. 1017-1098.
- Heckman, J. y Mosso, S. (2014). *The economics of human development and social mobility*. Annu. Rev. Econ., 6(1), 689-733.
- Hanushek, E., Piopiunik, M. y Wiederhold, S. (2014). *The value of smartet teachers: international evidence on teacher cognitive skills and student performance*. NBER Working Paper 20727.
- Iacus, S., King, G. y Porro, G. 2011. *Causal Inference without Balance Checking: Coarsened Exact Matching*. Oxford University Press.
- INEE. (2019). *Bases de datos INEE*. Disponible en <https://historico.mejoredu.gob.mx/bases-de-datos-inee-2019#spd>
- INEGI. (2015). *Medición de la Pobreza: Pobreza a nivel municipio 2010 y 2015*. Obtenido de <https://www.coneval.org.mx/Medicion/Paginas/Pobreza-municipal.aspx>
- INEGI. (2010). *Censo de Población y Vivienda 2010*. Obtenido

de <https://www.inegi.org.mx/programas/ccpv/2010/>

INEGI. (2010). *Clasificación de lenguas indígenas 2010*. México: INEGI.

INEGI. (s.f.). Encuesta Intercensal 2015. Obtenido de <https://www.inegi.org.mx/programas/intercensal/2015/>

King et al, 2011. *Comparative Effectiveness of Matching Methods for Causal Inference*. Working Paper.

Lenneberg, E. (1975). *Fundamentos biológicos del lenguaje*. Madrid: Alianza editorial.

Marlett, S. (2019). Familia y tronco: definiciones. Obtenido de <http://www.mexico.sil.org/es/familia-tronco-definiciones>

Miscevic-Kadijevic, Gordana. (2015). *TIMSS 2011: Relationship between self-confidence and cognitive achievement for Serbia and Slovenia*. Revista Electrónica de Investigación Educativa, 17(3), 109-115.

MOCYR. (2018). *Evaluación de Consistencia y Resultados 2017-2018: Programa para la Inclusión y la Equidad Educativa*. México: SEP.

Muralidharan, K. y Sundararaman, V. (2011). *Performance Pay: Experimental Evidence from India*. Journal of Political Economy, Vol. 119, No. 1, pp. 39-77.

Narro, J., & Moctezuma, D. (2012). Analfabetismo en México:

una deuda social. Obtenido de https://www.inegi.org.mx/rde/RDE_07/Doctos/RDE_07_Art1.pdf

Newport, E. (1990). *Maturational constraints on language learning*. Cognitive Science. 14, p. 11-28.

Nkonde, et al. (2018). *Evaluating the impact of teaching and learning of mathematics and science using local language (language of play) in primary schools in Muchinga Province, Zambia, a case of Chinsali District*. American Journal of Educational Research, 6(8), pp. 1153-1163.

Rivkin, S., Hanushek, E. y Kain, J. (2005). *Teachers, Schools and Academic Achievement*. Econometrica, Vol. 73, No. 2, pp. 417-458.

SEP & INEGI. (2013). *Censo de escuelas, maestros y alumnos de educación básica y especial 2013*. Disponible en <https://www.inegi.org.mx/sistemas/mapa/atlas/>

SEP. (2014). *Diagnóstico del Programa: Programa S244 Inclusión y Equidad Educativa*. México: SEP.

Smits, Huisman y Kruijff. (2008). *Home language and education in the developing world*. Netherlands: UNESCO.

Todd, Petra y Wolpin, Kenneth. (2003). *On the specification and estimation of the production function for cognitive achievement*. The Economic Journal, Vol. 113., No. 485. Pp. F3-F33.

Wittgenstein, Ludwig. (1921). Tractatus logico-philosophicus. México: ITAM.

Wooldridge, Jeffrey. (2015). Introducción a la Econometría. Quinta Edición. México, DF: Cengage Learning Editores.

World Bank. (2021). Loud and Clear: Effective Language of Instruction Policies for Learning. World Bank: Washington, DC.

8. Replicabilidad

Con el objetivo de ser transparente y para contribuir a futuros estudios que estén interesados en utilizar los datos con los que se construyó la parte empírica de esta tesis, queda a disposición pública el repositorio de GitHub que contiene el R Markdown con el código que genera este documento y todas las bases de datos que se utilizaron: https://github.com/alegrim13/ITAM_Tesis

