



UNIVERSIDAD POPULAR AUTÓNOMA DEL
ESTADO DE PUEBLA

DECANATO DE INGENIERÍAS

**CRITERIOS DE SEGUIMIENTO EN EL SECTOR DE
EDUCACIÓN BÁSICA Y MEDIA SUPERIOR EN PUEBLA:
TASA DE TERMINACIÓN Y TASA DE COBERTURA
BASADOS EN UN ENFOQUE DE CIENCIA DE DATOS.**

Trabajo práctico para obtener el grado de
Maestra en Ciencia de Datos e Inteligencia de Negocios

Presenta

Mónica Angélica Portillo Pedraza

17970055

Director de Trabajo Práctico

Dr. Emmanuel Olivera Pérez

Co-directora de Trabajo Práctico

Dra. Argelia Berenice Urbina Najera

Verano 2022



UPAEP – Secretaría General

Dirección General de Apoyos Académicos

Dirección del Centro de Recursos para el Aprendizaje y la Investigación.

Biblioteca Central - **Karol Wojtyla**

Tesis Digitales Restricciones de uso:

DERECHOS RESERVADOS ©

PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de textos, imágenes, gráficas, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente de donde la obtuvo mencionando el autor o autores involucrados en el documento.

Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Puebla, Puebla a 03 de junio de 2022

Dra. Ma. del Rocío Guadalupe Morales Salgado

Directora de Programa Educativo

Facultad de Tecnologías de Información y Ciencia de Datos

UPAEP

PRESENTE

Asunto: Aprobación para digitalización de Trabajo práctico

Los que suscriben la presente, en calidad de comité de **Trabajo práctico** de la Lic. Mónica Angélica Portillo Pedraza, con número de matrícula **17970055** e ID **3342007**, le informamos que hemos revisado el documento de Trabajo práctico titulado **Criterios de seguimiento en el sector de Educación Básica y Media Superior en Puebla: Tasa de terminación y Tasa de cobertura basados en un enfoque de Ciencia de Datos**.

Como resultado de esta revisión consideramos que el documento de **Trabajo práctico** cumple con los requisitos necesarios para ser defendido con el propósito de obtener el grado de **Maestría en Ciencia de Datos e Inteligencia de Negocios**, por lo que aprobamos que se realice la digitalización correspondiente y por ende la programación del examen de grado.

La saludamos afectuosamente, anticipando nuestro agradecimiento por la atención a la presente

Cordialmente,



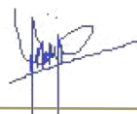
Dr. Emmanuel Olivera Pérez
Director de trabajo práctico



Dra. Argelia Berenice Urbina
Najera
Co-directora de trabajo práctico



Mtro. José Francisco Zeron
Jiménez
Asesor



Dra. Sandra R. Murillo Cano
Sinodal



Dra. Ma. del Rocío G. Morales Salgado
Sinodal



POSGRADOS
17 SUR 901
Barrio de Santiago
Puebla, Pue. México
C.P. 72410

Tel.: 01 (222) 229 9400
Fax: 01 (222) 2325251
01 800 224 2200
www.upaep.mx

Contenido

Introducción	3
Objetivos	8
Objetivo general	8
Objetivos específicos.....	8
Justificación	9
Alcances y limitaciones	9
Viabilidad	9
Resultados esperados.....	9
Procesos de desarrollo.....	10
Descripción de la base de datos	10
Método de recolección de datos.....	11
Métodos	12
Diagrama de datos.....	17
Resultados de los ejercicios.....	18
Conclusiones.....	41
Referencias	42

Introducción

El éxito de una organización depende principalmente de la calidad y eficacia de las decisiones que se pueden tomar por medio de un análisis de datos adecuado. Para la toma de decisiones se hace uso de sistemas, prácticas, aplicaciones y tecnologías que ayudan a analizar datos con el fin de comprender mejor el negocio. Esta práctica ha tomado fuerza con el desarrollo de sistemas de planificación de recursos empresariales, bases de datos, almacenes de datos y procesos de software. Business Analytics incluye herramientas de consulta, OLAP, informes y alertas que pueden responder consultas como qué pasa si la tendencia continúa, qué sucederá después, es decir, predicción. Se hace uso de los datos, análisis estadístico y cuantitativo, modelos predictivos y explicativos. (Khan, Nadeem, & Ali, 2019)

La analítica avanzada tiene 3 funciones: descriptiva, predictiva y de optimización. El análisis descriptivo se basa en mostrar información del estado y desempeño de una empresa haciendo uso de bases y almacenes de datos con el fin de obtener información de datos históricos que se presenta en forma de informes, cuadros de mando y agrupación. La visualización de datos es importante en este tipo de análisis. El análisis predictivo determina el probable futuro de un evento e identifica patrones y relaciones. Se aplican técnicas estadísticas, de minería de datos y en ocasiones de visualización. Se usan patrones en modelos que se pueden aplicar a nuevas transacciones para predecir comportamientos y resultados. Las técnicas pueden incluir; clasificación, clustering, asociación y optimización y analítica prescriptiva. (Ibídem)

Los macrodatos y el aprendizaje automático forman parte de la revolución en el análisis de datos y si se le agregan a la estadística, se obtiene la ciencia de datos y si a esto se le agrega economía, econometría y un enfoque en toma de decisiones, se obtiene ciencia de datos comerciales. El aprendizaje automático se centra en maximizar el rendimiento predictivo, evalúa qué tan bien un modelo entrenado en un conjunto de datos predecirá nuevos datos, el propósito es hacer predicción de un futuro que sigue los mismos patrones que en los datos pasados. (Taddy, 2019)

Robert Lucas de la Universidad de Chicago ganó el Premio Nobel por su trabajo en macroeconomía, se pronunció en contra de la práctica común de usar correlaciones pasadas de las variables macroeconómicas para guiar la formulación de políticas y argumentó que para entender adecuadamente cómo los cambios de política individuales afectan un sistema completo, se debe modelar la macroeconomía estructuralmente derivada de un fundamento de la teoría económica. Un responsable de políticas que pueda implementar aprendizaje automático para aquellas tareas de predicción que enfrenta, puede automatizar y acelerar los procesos de toma de decisión. (Taddy, 2019)

Una buena ciencia guarda sus esfuerzos económicos y estadísticos para difíciles cuestiones estructurales.

El análisis de políticas tradicional implica un proceso de extrapolación de datos recolectados de muestras representativas para después desarrollar políticas basadas en esa muestra. Por otra parte, la analítica de políticas aprovecha los nuevos datos y las nuevas tecnologías computacionales para proveer una visión más precisa de un sistema. El gobierno siempre ha recopilado, generado y compilado grandes cantidades de datos y con el desarrollo de las Tecnologías de la Información y Computación (TIC) digitales, estas actividades se han computarizado cada vez más. La analítica de políticas puede ayudar a respaldar decisiones, proporcionar nuevos enfoques analíticos para gestionar problemas públicos, identificación temprana de tendencias a través de análisis predictivo además de que los sistemas de soporte de decisiones analizan información en tiempo real con el fin de identificar patrones y correlaciones que pudieran estar ocultos. (Longo & McNutt, 2018)

Con el crecimiento de Internet, al análisis de políticas se enfrenta ahora a una gran cantidad de información llamada “macrodatos”, lo cual implica el aumento de volumen, velocidad y alcance de evidencia relevante para los formuladores de políticas. Esta información se genera a través de actividades que hacen las personas, de sus interacciones con servicios web y plataformas sociales y de su comportamiento de compra revelado a través de tarjetas electrónicas por mencionar algunas actividades. Los datos generados brindan un análisis de políticas basado en datos precisos, continuos y dinámicos. (Longo & McNutt, 2018)

Para este trabajo se toma en cuenta el componente de Educación de la base de datos Matriz de Indicadores de Resultados del programa PROSPERA en el estado de Puebla. Contiene información de niños, niñas y jóvenes que recibieron apoyos económicos para sus estudios y la compra de útiles escolares. Se otorgan becas educativas a los menores de 18 años que formen parte de familias afiliadas. El monto de apoyo incrementa conforme aumenta el grado escolar. Los beneficiarios deben acreditar su inscripción para cursar entre 1º de primaria y 3º de secundaria. Las becas de educación media superior son para las y los jóvenes inscritos de entre 14 y 21 años. (Secretaría de Bienestar, 2017)

El gobierno federal y el de cada entidad federativa con el fin de asegurar la educación como derecho y combatir la pobreza, acuden al financiamiento de la educación pública y se comprometen a consolidar las fuentes de financiamiento como a dedicar recursos reales crecientes. En específico siguen los criterios de crecimiento del gasto para el desarrollo social, el financiamiento federal a la educación pública debe aumentar anualmente, por lo menos en la misma proporción que el crecimiento del Producto Interno Bruto estimado en los Criterios Generales de Política Económica. El gobierno federal concentra la mayor parte de los ingresos

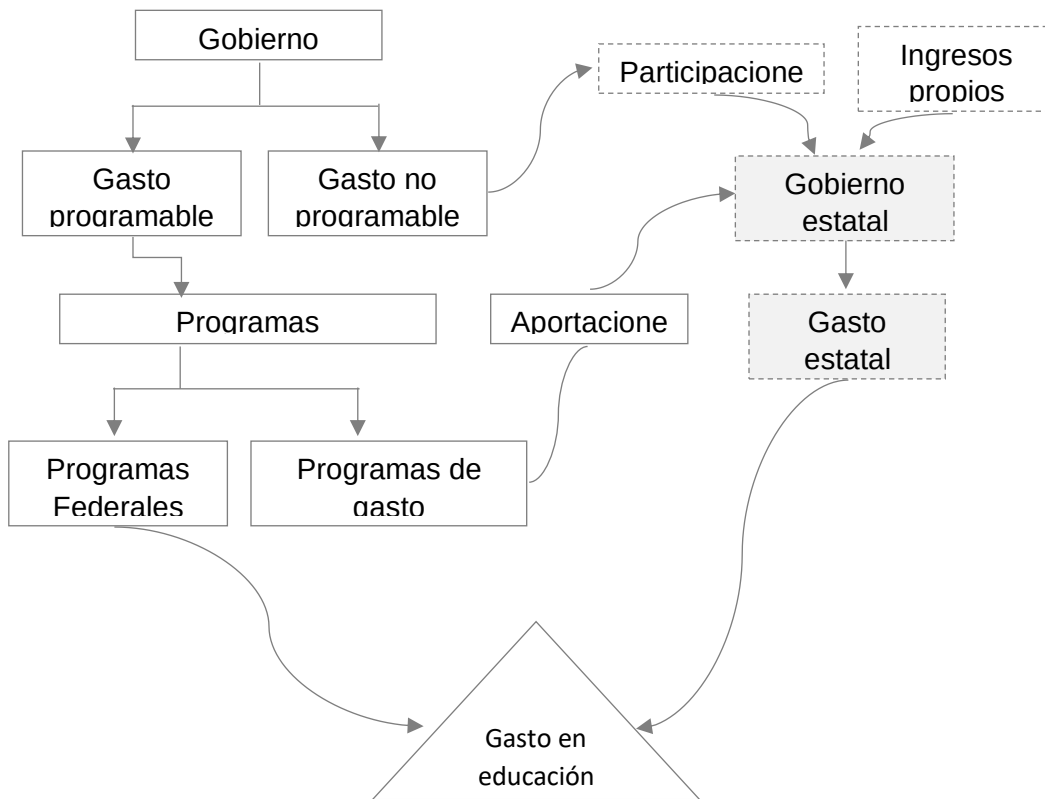
públicos para después redistribuir una parte entre los estados de la república. (Instituto Nacional para la Evaluación de la Educación, 2018)

A través de participaciones y programas presupuestarios llamados aportaciones, la federación retorna parte de los recursos a los gobiernos locales. Estos recursos son transferidos a las entidades federativas y no pueden ser desviados para fines distintos de los ya establecido por la normativa que los rige. Los criterios que deben seguir son equidad, eficiencia, eficacia, economía, honradez y transparencia. El gobierno de cada entidad y el federal deben rendir cuentas acerca del ejercicio de los recursos públicos que disponen en un año fiscal. (Instituto Nacional para la Evaluación de la Educación, 2018)

El gasto público en educación tiene dos principales fuentes de financiamiento: federal y estatal. La federación destina recursos financieros en primer lugar para el Gasto Programable del Sector Público Presupuestario que gestiona y provee bienes y servicios públicos a través de dos programas: Programas Federales y Programas de Gasto Federalizado (o aportaciones). En segundo lugar, transfiere a los estados recursos no condicionados para que los gobiernos locales, junto con sus ingresos locales, provean bienes y servicios educativos. Se sabe que el gobierno federal contribuye con la mayor parte del gasto en educación. (Instituto Nacional para la Evaluación de la Educación, 2018)

El Gasto Federal ejercido en Educación Obligatoria permite comparar los montos asignados a Educación Básica y Media Superior y distinguir los recursos federales, es decir los programas federales, destinados a proporcionar subsidios, proveer servicios públicos, evaluar, regular y mejorar procesos administrativos y los programas de gasto federalizado, que mantienen en operación los centros de Educación Básica, Media Superior y educación para adultos. (Instituto Nacional para la Evaluación de la Educación, 2019)

Se muestran las principales fuentes de financiamiento del gasto en educación en el siguiente diagrama:



El Gasto Federal Ejercido en Educación Obligatoria se calcula con la siguiente fórmula:

$$GFEO = G_{PF} + G_{PGF}$$

$$G_{PF} = \sum_{i=1}^m PF_i$$

$$G_{PGF} = \sum_{j=1}^n PGF_j$$

Donde:

$GFEO$ = Gasto Federal Ejercido en Educación Obligatoria.

G_{PF} = Gasto anual d-en educación obligatoria mediante programas federales.

G_{PGF} = Gasto anual en educación obligatoria mediante programas de gasto federalizado.

PF_i = Gasto anual del programa federal i .

PGF_j = Gasto anual del programa de gasto federalizado j .

m = Total de programas federales de educación obligatoria.

n = Total de programas de gasto federalizado de educación obligatoria.

Los Programas Federales representan acciones compensatorias que el gobierno federal implementa mediante convenios con los gobiernos estatales, como subsidios, gastos de operación, otros gastos corrientes, inversión y servicios personales. Por ejemplo, para Educación Básica, becas para alumnos, apoyo pedagógico para servicios educativos comunitarios, dotación de útiles escolares, capacitación para el aprovechamiento del apoyo a la gestión escolar, capacitación a supervisores de zona escolar, transferencia de recursos a las entidades federativas para incrementar la cobertura, la inclusión y la equidad, entre otros. (Instituto Nacional para la Evaluación de la Educación, 2018)

Los programas de Gasto Federalizado tienen su fundamento en la Ley de Coordinación Fiscal, donde se definen las reglas de distribución a las entidades. Se destinan en mayor medida a mantener la operación de los servicios educativos transferidos a los gobiernos locales mediante convenios de descentralización. En el segmento de Educación, los programas son el Fondo de Aportaciones Múltiples (FAM), el Fondo de Aportaciones para la Nómina Educativa y Gasto Operativo (FONE) y el Fondo de Aportaciones para la Educación Tecnológica y de Adultos (FAETA). Por medio de la firma de los convenios entre la federación y los gobiernos estatales en el marco del Acuerdo Nacional para la Modernización de la Educación Básica (ANMEB), se consolidaron los fondos con el fin de transferir a los gobiernos locales el control y gestión de las escuelas de EB y de los recursos financieros para su operación y aumentar la cobertura.

Los Programas de Gasto Federalizado sufren cambios menos frecuentemente a comparación de los Programas Federales, estos cambian las reglas y lineamientos de operación con base en los resultados de evaluaciones anuales. El incumplimiento de los gobiernos locales de la normativa de la contabilidad gubernamental limita el análisis del gasto educativo en forma desagregada.

La tasa neta de cobertura es un indicador que mide el acceso de los niños y jóvenes a la educación. Considera a los matriculados en edad idónea para cierto nivel educativo de acuerdo con la edad normativa (preescolar 3 a 4 años, primaria 6 a 11 años, secundaria 12 a 14 años y media superior 15 a 17 años). Este indicador es importante y da una idea de qué tan eficaz es el Sistema Educativo Nacional para matricular a la población que debería asistir al nivel educativo de acuerdo a su edad. Con esta información se puede estimar las brechas para alcanzar la matriculación universal en EB y trazar las metas en EMS. (Instituto Nacional para la Evaluación de la Educación, 2019)

La fórmula del cálculo es la siguiente:

$$\frac{tA_{en}^{niv}}{tP_{en}^{niv}} \times 100$$

Donde:

tA_{en}^{niv} = Alumnos matriculados en edades idóneas o típica en correspondiente al nivel o tipo educativo niv al inicio del ciclo escolar t .

tP_{en}^{niv} = Población a mitad de año t en edades idóneas o típica en para cursar el nivel o tipo educativo niv .

niv = Nivel o tipo educativo: preescolar, primaria, secundaria y media superior.

en = Edad idónea o típica correspondiente a cada nivel o tipo educativo niv . De 3 a 5 años en preescolar, 6 a 11 años en primaria, 12 a 14 años en secundaria y 15 a 17 años en media superior.

t = Ciclo escolar.

Para el cálculo se contabiliza a los alumnos que están en el nivel que les corresponde según su edad, los alumnos que no cumplen con la edad requerida quedan fuera de este cálculo. El indicador puede tener valores de cero a cien y en ocasiones el valor excede el 100% debido a que se utilizan distintas fuentes de información. (Instituto Nacional para la Evaluación de la Educación, 2019)

Objetivos

Objetivo general

Determinar los efectos y la magnitud sobre el Porcentaje de Terminación de Educación Básica y la Tasa Neta de Cobertura de Educación Media Superior, a partir de un conjunto de variables relacionadas con apoyos monetarios, gasto público y de cobertura integradas a la política de educación en el estado de Puebla.

Objetivos específicos

- Analizar cuantitativamente la dinámica del comportamiento del gasto público para mejorar la toma de decisiones en Educación Básica y Media Superior.
- Determinar mediciones del gasto público que permitan la evaluación oportuna del gasto en Educación Básica y Media Superior.
- Pronosticar el comportamiento de la Tasa de Terminación de EB y de la Tasa de Cobertura de EMS para los próximos 4 años.

Justificación

El seguimiento de programas es de suma importancia para identificar si han existido cambios a lo largo del tiempo y si estos han sido positivos. El presente trabajo busca aplicar modelos estadísticos basados en ciencia de datos para dar seguimiento del Porcentaje de Terminación de Educación Básica y de la Tasa Neta de Cobertura de Educación Media Superior con el fin de fortalecer la toma de decisiones en cuanto a la educación en el estado de Puebla.

Alcances y limitaciones

El análisis comprende información de la base de datos llamada Matriz de Indicadores de Resultados del programa PROSPERA que contiene información de los indicadores que miden los objetivos y los resultados esperados con PROSPERA por Entidad Federativa del ejercicio fiscal, de la cual se toma la información del estado de Puebla. Además de información del rubro de Educación obtenida del Instituto Nacional para la Evaluación de la Educación.

Los principales beneficiarios serán:

- Secretaría de Educación

Un limitante es el lapso del que se tiene información, el cual va del año 2011 al 2018, provocando desactualización. La información consultada se reporta anualmente, pero hasta el momento no se encuentra información actualizada en las páginas de acceso libre.

Viabilidad

El proyecto es viable ya que las bases de datos están disponibles en la página de Datos Abiertos del Gobierno de México y en la del INEE en el apartado de Banco de Indicadores y son de libre acceso.

Resultados esperados

Indicadores y mediciones que permitan dar seguimiento al Porcentaje de Terminación de EB y a la Tasa Neta de Cobertura de EMS con la ayuda de modelos matemáticos.

Procesos de desarrollo

Descripción de la base de datos

- Matriz de Indicadores de Resultados del programa PROSPERA

Contiene información de los indicadores que miden los objetivos y los resultados esperados con PROSPERA por Entidad Federativa del ejercicio fiscal, de la cual se toma la siguiente información del estado de Puebla:

Porcentaje de becarios de Educación Básica a los que se les emitieron apoyos monetarios para becas educativas
Porcentaje de becarias y becarios de Educación Básica a los que se emitió apoyos para útiles escolares
Porcentaje de terminación de educación básica de los jóvenes
Porcentaje de becarios de educación media superior a los que se les emitieron apoyos monetarios para becas educativas
Porcentaje de becarias y becarios de educación media superior a los que se emitió apoyos para útiles escolares

Tabla 1. Variables de la base Matriz de Indicadores de Resultados del programa PROSPERA

- Datos del Instituto Nacional para la Evaluación de la Educación

Contiene información nacional y del estado de Puebla sobre la educación.

Gasto de Programa Federal Total de Educación Básica
Programa de Gasto Federalizado Total de Educación Básica
Tasa Neta de Cobertura por entidad federativa de Educación Básica
Número de alumnos de Educación Básica
Número de docentes de Educación Básica
Número de planteles de Educación Básica
Gasto de Programa Federal Total de Educación Media Superior
Programa de Gasto Federalizado Total de Educación Media Superior
Tasa Neta de Cobertura por entidad federativa de Educación Media Superior
Número de alumnos de Educación Media Superior
Número de docentes de Educación Media Superior
Número de planteles de Educación Media Superior

Tabla 2. Variables del INEE

Método de recolección de datos

Los datos de Educación de los programas antes mencionados se obtienen a través de la página de Datos Abiertos del Gobierno de México, se descargan en formato RAR que contienen los archivos en formato .xls obtenidos de: <https://datos.gob.mx/busca/dataset/matriz-de-indicadores-de-resultados>.

Los datos del INEE se obtienen a través de su página oficial en línea, en el apartado de Banco de Indicadores Educativos, se descargan en formato .xls obtenidos de: <https://historico.mejoredu.gob.mx/evaluaciones/indicadores-educativos/>.

- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2011:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/04/2011_EF_Pue.pdf
- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2012:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/04/2012_EF_Pue.pdf
- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2013:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/04/2013_EF_Pue.pdf
- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2014:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/04/2014_EF_Pue.pdf
- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2015:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/04/2015_EF_Pue.pdf
- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2016:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/04/2016_EF_Pue.pdf
- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2017:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/04/2017_EF_Pue.pdf
- Número de alumnos, docentes, planteles, Tasa Neta de Cobertura de Educación Básica y Media Superior año 2018:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/09/2018_EF_Pue.pdf

- Gasto de Programa Federal Total y Programa de Gasto Federalizado Total de Educación Básica y Media Superior:
https://historico.mejoredu.gob.mx/wp-content/uploads/2019/09/2018_AR03_b.2-1.pdf

Nota: Para el uso de datos de Número de alumnos, docentes y planteles y Tasa Neta de Cobertura en Educación Básica se promediaron los datos de Primaria y Secundaria.

Métodos

- Mínimos Cuadrados Ordinarios

El modelo de regresión simple se usa para estudiar la relación entre dos variables, se interesa en explicar y en términos de x . Se estudia cómo varía y cuando se cambia x . Se confrontan 3 problemas, el primero es que no hay una relación exacta entre dos variables entonces cómo se permite que otros factores afecten y . El segundo es cuál es la relación funcional entre y y x y el tercero es cómo saber si se captura una relación ceteris paribus entre y y x si ese es el objetivo. (Wooldridge, 2013)

Ecuación de modelo de regresión lineal simple:

$$y = \beta_0 + \beta_1 x + u$$

y es la variable dependiente, la variable explicada y x es la variable independiente o variable explicativa. También se le conoce como modelo de regresión lineal de dos variables o modelo de regresión lineal bivalente ya que se refiere a dos variables, que son x y y .

La variable u es llamada término de error en la relación y representa factores distintos de x que afectan y . Si estos factores son fijos, es decir que cambian en un valor de 0, entonces se dice que x tiene un efecto lineal en y .

$$\Delta y = \beta_1 \Delta x \text{ si } \Delta u = 0$$

Entonces el cambio en y es β_1 multiplicado por el cambio en x lo que quiere decir que β_1 es el parámetro de pendiente en la relación entre y y x , donde se mantienen fijos los factores en u . β_0 es el parámetro de intercepción y también es llamado término constante.

Se puede asumir que el valor promedio de u es cero, es decir $E(u) = 0$.

El coeficiente de correlación es una medida natural de la asociación entre dos variables aleatorias. Si u y x no están correlacionados, no están linealmente

relacionados. Se asume que el valor promedio de u no depende del valor de x , y se representa de la siguiente manera:

$$E(u|x) = E(u)$$

Con esto se asume que $E(u|x) = 0$, es decir una media condicional de cero.

La fórmula para estimar los parámetros β_0 y β_1 es la siguiente:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Donde $[(x_i, y_i): i = 1, \dots, n]$, n es el tamaño de la población. Aquí, u_i es el término de error para la observación i porque contiene todos los factores que afectan a y_i distintos de x_i .

$$E(u) = 0$$

y

$$Cov(x, u) = E(xu) = 0$$

$$n^{-1} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) =$$

y

$$n^{-1} \sum_{i=1}^n x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) =$$

Usando las propiedades básicas de la suma, se puede escribir como:

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$

Donde $\bar{y} = n^{-1} \sum_{i=1}^n y_i$ es el promedio muestral de y_i e igual para x_i . Esta ecuación permite escribir $\hat{\beta}_0$ en términos de $\hat{\beta}_1$, $\bar{y}$ y $\bar{x}$:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Una vez que se tiene la estimación de la pendiente $\hat{\beta}_1$, se obtiene la estimación de la intersección $\hat{\beta}_0$ dado $\bar{y}$ y $\bar{x}$:

$$\sum_{i=1}^8 x_i [y_i - (\bar{y} - \hat{\beta}_1 \bar{x}) - \hat{\beta}_1 x_i] =$$

Que tras la reordenación da:

$$\sum_{i=1}^8 x_i(y_i - \bar{y}) - \hat{\beta}_1 \sum_{i=1}^8 x_i(x_i - \bar{x})$$

De las propiedades básicas de la suma

$$\sum_{i=1}^8 x_i(x_i - \bar{x}) = \sum_{i=1}^8 (x_i - \bar{x})^2$$

Y

$$\sum_{i=1}^8 x_i(y_i - \bar{y}) = \sum_{i=1}^8 (x_i - \bar{x})(y_i - \bar{y})$$

Por lo tanto

$$\sum_{i=1}^8 (x_i - \bar{x})^2 > 0$$

La pendiente estimada es:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^8 (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^8 (x_i - \bar{x})^2}$$

Para cualquier $\hat{\beta}_0$ y $\hat{\beta}_1$ se define un valor ajustado para y cuando $x = x_i$ como

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Este es el valor que se predice para y cuando $x = x_i$ para la intersección y pendiente dadas. Hay un valor ajustado para cada observación. El residual de la observación i es la diferencia entre el y_i real y su valor ajustado:

$$\hat{u}_i = y_i - \bar{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$$

Ahora se eligen $\hat{\beta}_0$ y $\hat{\beta}_1$ para hacer la suma de residuos al cuadrado

$$\sum_{i=1}^8 \hat{u}_i^2 = \sum_{i=1}^8 (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Una vez que se han determinado las estimaciones de la intersección y la pendiente de MCO, se forma la línea de regresión de MCO:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Esta ecuación se puede llamar función de regresión muestral porque es la versión estimada de la función de regresión poblacional $E(x|y) = \beta_0 + \beta_1 x$.

La estimación de la pendiente se puede escribir como:

$$\hat{\beta}_1 = \Delta \hat{y} / \Delta x$$

Dice la cantidad en la que $\hat{y}$ cambia cuando x aumenta en una unidad. Equivalentemente:

$$\Delta \hat{y} = \hat{\beta}_1 \Delta x$$

- Análisis Factorial con factores

El objetivo del análisis factorial es describir las relaciones de covarianza entre factores, es decir muchas variables en términos de pocas cantidades aleatorias subyacentes, pero no observables. Se basa en que las variables se pueden agrupar por correlaciones. Es decir, en un grupo en especial puede suceder que todas las variables estén altamente correlacionadas entre sí pero tienen correlaciones relativamente pequeñas respecto a las variables de un grupo diferente. Aquí se dice que un grupo representa un factor responsable de las correlaciones observadas. Este análisis se puede considerar una extensión del análisis de componentes principales. (Johnson & Wichern, 2014)

El vector aleatorio observable X , con p componentes tiene media y matriz de covarianza. El modelo factorial dice que X es linealmente dependiente de unas pocas variables aleatorias no observables que llevan por nombre factores comunes y p fuentes adicionales de variación llamadas errores o factores específicos. El modelo es el siguiente:

$$X_1 - \mu_1 = l_{11}F_1 + l_{12}F_2 + \cdots + l_{1m}F_m + \varepsilon_1$$

$$X_2 - \mu_2 = l_{21}F_1 + l_{22}F_2 + \cdots + l_{2m}F_m + \varepsilon_2$$

$$X_p - \mu_p = l_{p1}F_1 + l_{p2}F_2 + \cdots + l_{pm}F_m + \varepsilon_p$$

O en notación matricial:

$$\underset{(p \times 1)}{X} - \underset{(p \times 1)}{\mu} = \underset{(p \times m)}{L} \underset{(m \times 1)}{F} + \underset{(p \times 1)}{\varepsilon}$$

μ_i = media de la variable i

ε_i = i -ésimo factor específico

F_j = j -ésimo factor común

El coeficiente l_{ij} se llama “carga” de la i -ésima variable sobre el j -ésimo factor así que se le llama matriz de cargas de los factores a la matriz $\mathbf{L}$. Las desviaciones de p $X_1 - \mu_1, X_2 - \mu_2, \dots, X_p - \mu_p$ son expresadas en términos de variables aleatorias de $p + m$ que no son observables las cuales son $F_1, F_2, \dots, F_m, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$. Con algunos supuestos de los vectores aleatorios de $\mathbf{F}$, el anterior modelo implica ciertas relaciones de covarianza, entonces se asume que:

$$E(\mathbf{F}) = \mathbf{0}_{(m \times 1)}, \quad \text{Cov}(\mathbf{F}) = E[\mathbf{F}\mathbf{F}'] = \mathbf{I}_{(m \times m)}$$

$$E(\boldsymbol{\varepsilon}) = \mathbf{0}_{(p \times 1)}, \quad \text{Cov}(\boldsymbol{\varepsilon}) = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \boldsymbol{\Psi}_{(p \times p)} = \begin{bmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi_p \end{bmatrix}$$

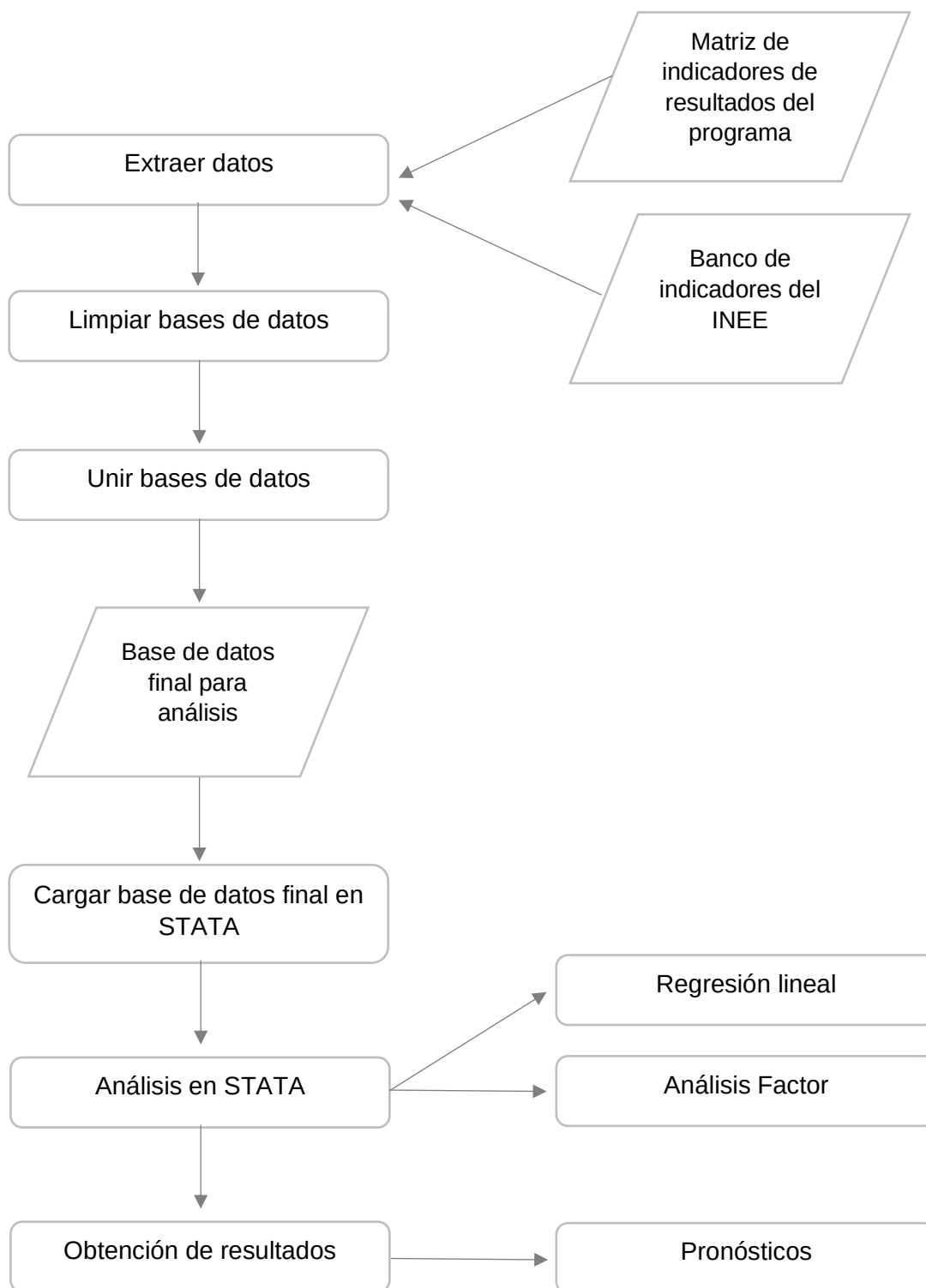
$\mathbf{F}$ y $\boldsymbol{\varepsilon}$ son independientes así que:

$$\text{Cov}(\boldsymbol{\varepsilon}, \mathbf{F}) = E(\boldsymbol{\varepsilon}\mathbf{F}') = \mathbf{0}_{(p \times m)}$$

Método de análisis de componentes principales (y factores)

$$\boldsymbol{\Sigma}_{(p \times p)} = \mathbf{L}_{(p \times p)} \mathbf{L}'_{(p \times p)} + \mathbf{0}_{(p \times p)} = \mathbf{L}\mathbf{L}'$$

Diagrama de datos



Resultados de los ejercicios

Estimación de mínimos cuadrados ordinarios

Educación Básica

Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa

Estimación 1

Programa de Gasto
Federalizado Total

Coeficiente	0.0001486***
Error estándar	0.0000304
Constante	0.2293719
R-cuadrado	0.7991

Estimación 2

Número de docentes

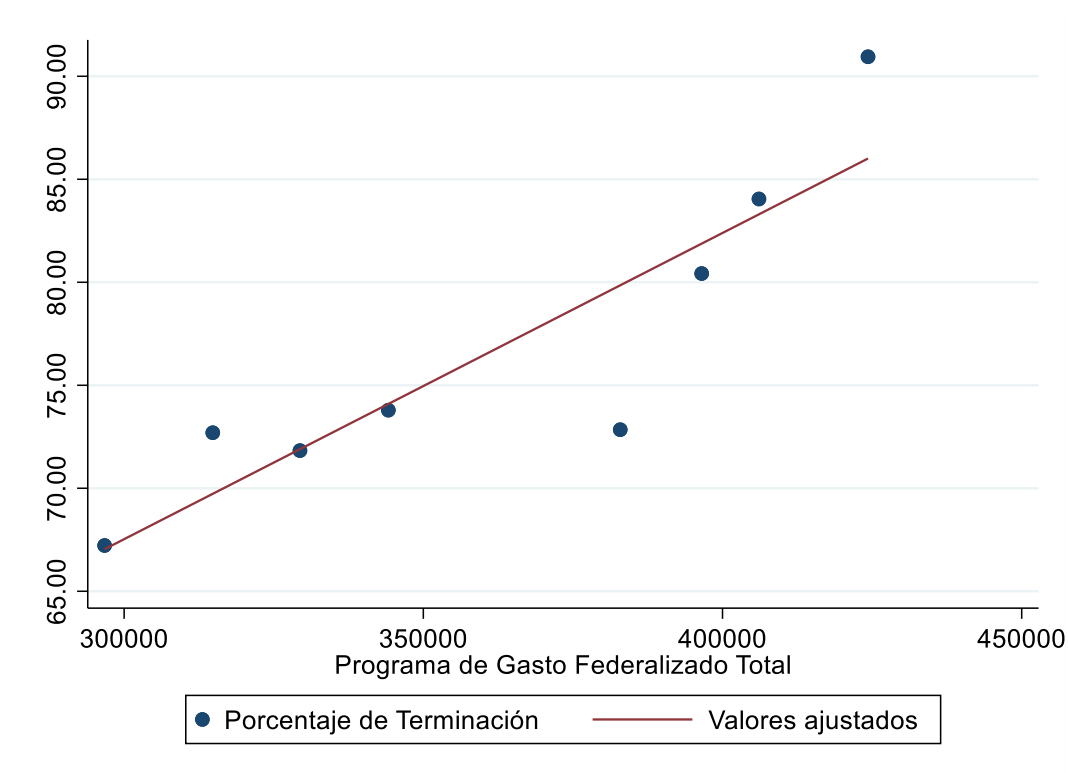
Coeficiente	0.0095083****
Error estándar	0.0010025
Constante	-378.5662
R-cuadrado	0.9375

Estimación 3

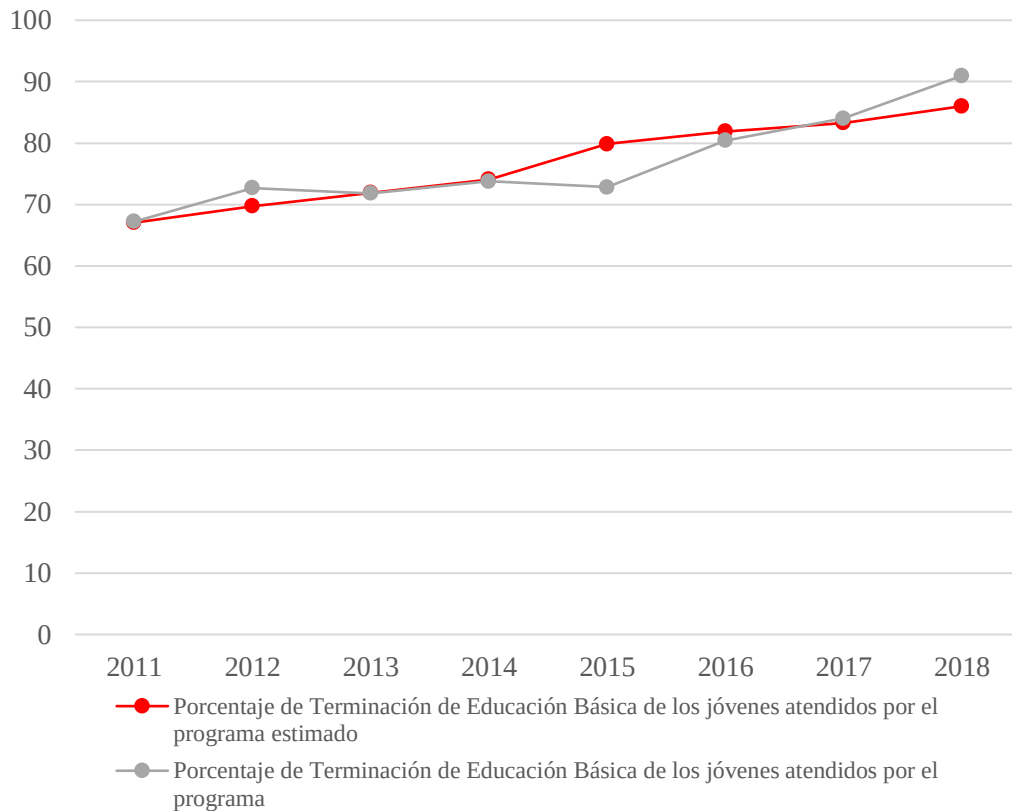
Número de planteles

Coeficiente	0.0687934***
Error estándar	0.0126406
Constante	-394.0367
R-cuadrado	0.8315

Estimación 1



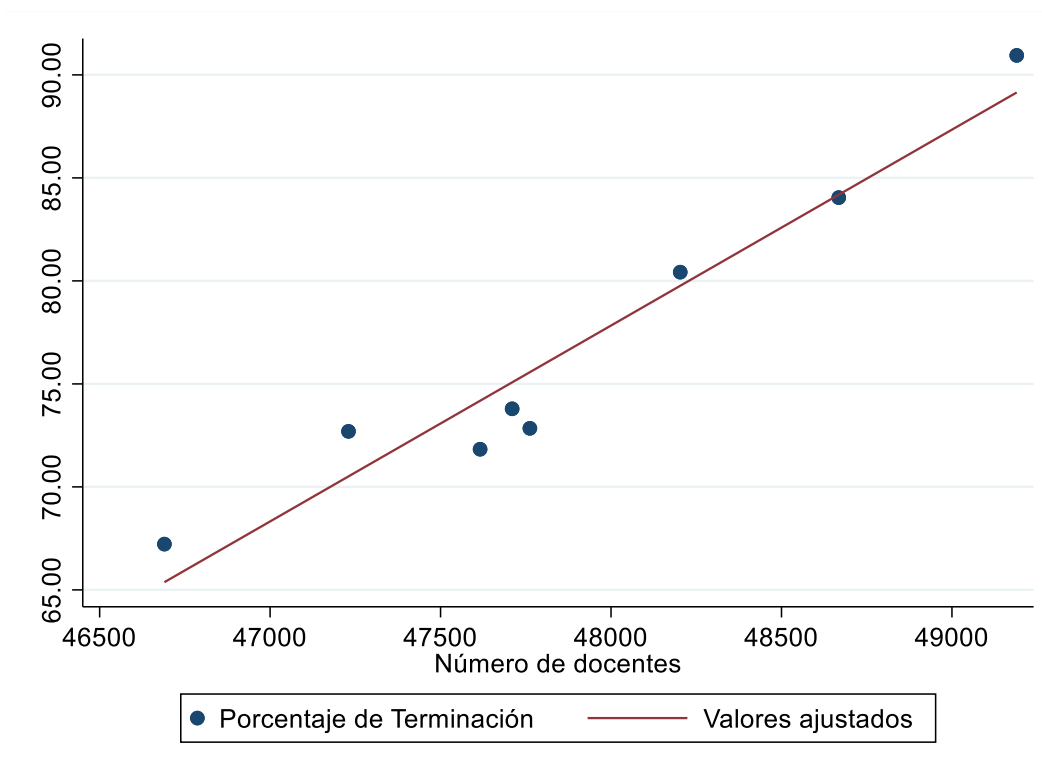
Año	Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa estimado	Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa	Programa de gasto Federalizado Total
2011	67.04523	67.22	296744.54
2012	69.73165	72.7	314817.91
2013	71.89763	71.83	329389.9
2014	74.09258	73.79	344156.82
2015	79.85211	72.84	382905.09
2016	81.87531	80.42	396516.44
2017	83.29834	84.04	406090.16
2018	86.00719	90.95	424314.43



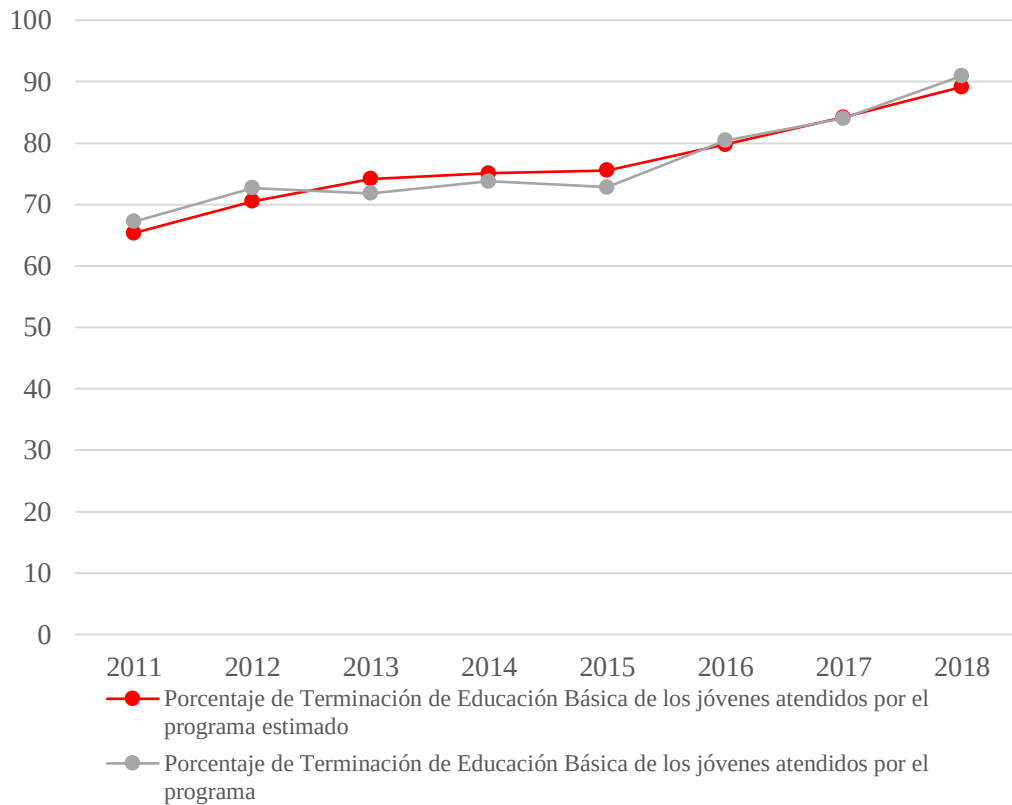
Para la estimación 1 se observa que el coeficiente es estadísticamente significativo al 1%. La magnitud del efecto parcial es de 0.0000304. Existe una relación positiva entre el Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa y el Programa de Gasto Federalizado Total. El valor de R-cuadrado que es de .7991, esto significa que el gasto explica hasta un 80% el porcentaje de terminación.

Se observa que los valores estimados del Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa no son muy diferentes de los valores reales específicamente en los años 2011, 2013, 2014 y 2017.

Estimación 2



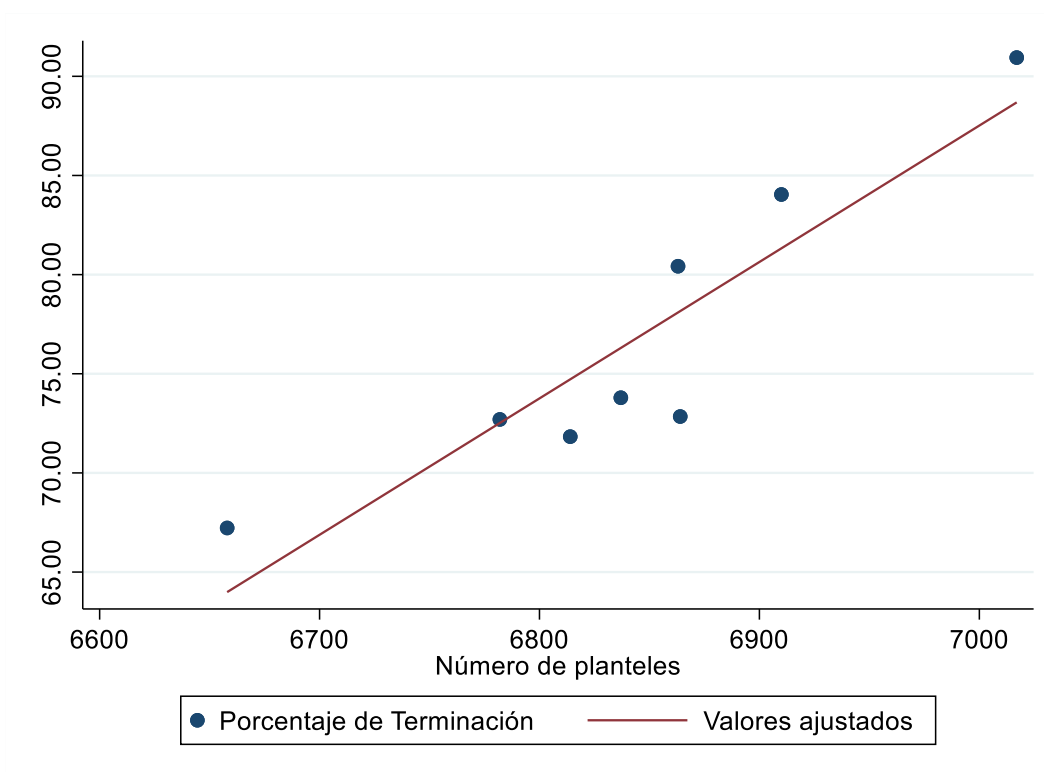
Año	Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa estimado	Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa	Número de docentes
2011	65.37568	67.22	46690
2012	70.51015	72.7	47230
2013	74.18035	71.83	47616
2014	75.07413	73.79	47710
2015	75.56856	72.84	47762
2016	79.76171	80.42	48203
2017	84.18307	84.04	48668
2018	89.14639	90.95	49190



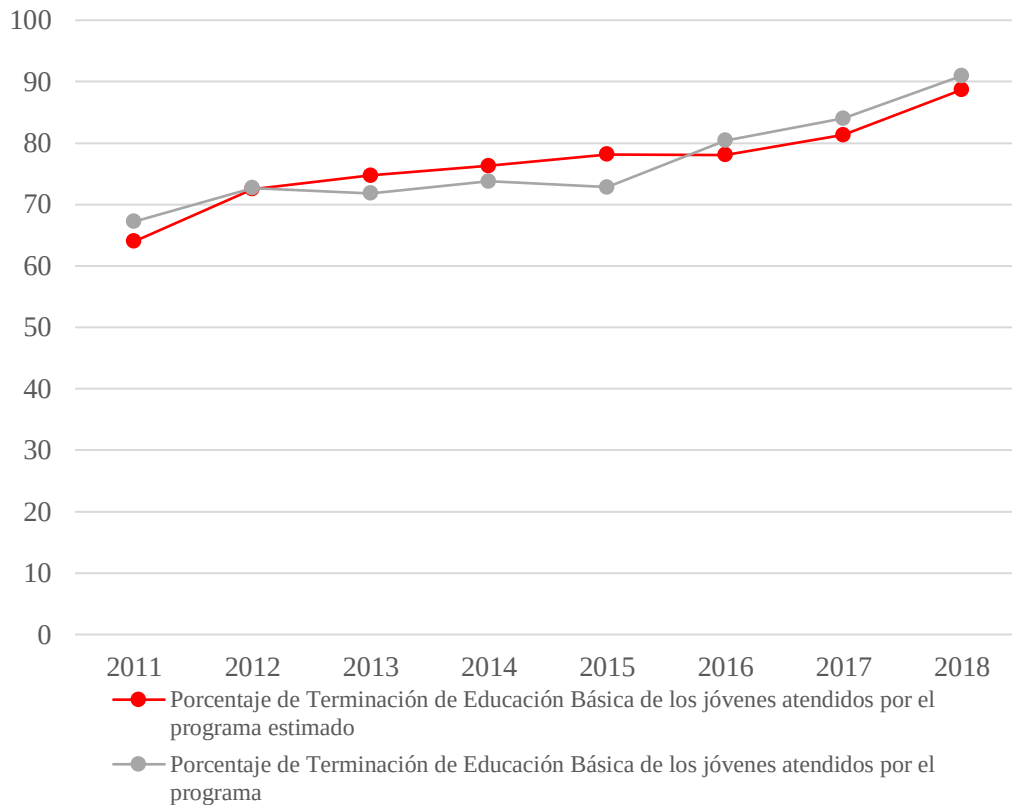
Para la estimación 2 se observa que el coeficiente es estadísticamente significativo al .1%. La magnitud del efecto parcial es de 0.0010025. Existe una relación negativa entre el Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa y el número de docentes. El valor de R-cuadrado que es de .9375, esto significa que el número de docentes explica hasta un 94% el porcentaje de terminación.

Se observa que los valores estimados del Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa no son muy diferentes de los valores reales y se acercan más en comparación de la estimación 1 para los años 2012, 2015, 2016, 2017 y 2018.

Estimación 3



Año	Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa estimado	Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa	Número de planteles
2011	63.98963	67.22	6658
2012	72.52	72.7	6782
2013	74.7214	71.83	6814
2014	76.30364	73.79	6837
2015	78.16106	72.84	6864
2016	78.09227	80.42	6863
2017	81.32556	84.04	6910
2018	88.68645	90.95	7017



Para la estimación 3 se observa que el coeficiente es estadísticamente significativo al 1%. La magnitud del efecto parcial es de 0.0126406. Existe una relación negativa entre el Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa y el número de planteles. El valor de R-cuadrado que es de .7991, esto significa que el número de planteles explica hasta un 80% del porcentaje de terminación.

Se observa que los valores estimados del Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa no son muy diferentes de los valores reales y se acercan más en comparación de la estimación 1 para los años 2012, 2015 y 2018. En comparación de la estimación 2, se acerca más para el año 2012.

Educación Media Superior

Tasa Neta de Cobertura

Estimación 1

Programa Federal Total

Coefficiente 0.0003401**

Error estándar 0.0001357

Constante 33.48996

R-cuadrado 0.5112

Estimación 2

Programa de Gasto
Federalizado Total

Coefficiente 0.0066564**

Error estándar 0.0024114

Constante 32.95691

R-cuadrado 0.5595

Estimación 3

Número de alumnos

Coefficiente 0.0002138**

Error estándar 0.000075

Constante 4.09232

R-cuadrado 0.5754

Estimación 4

Número de docentes

Coefficiente 0.0079041**

Error estándar 0.0028257

Constante -69.60009

R-cuadrado 0.5660

Estimación 5

Número de planteles

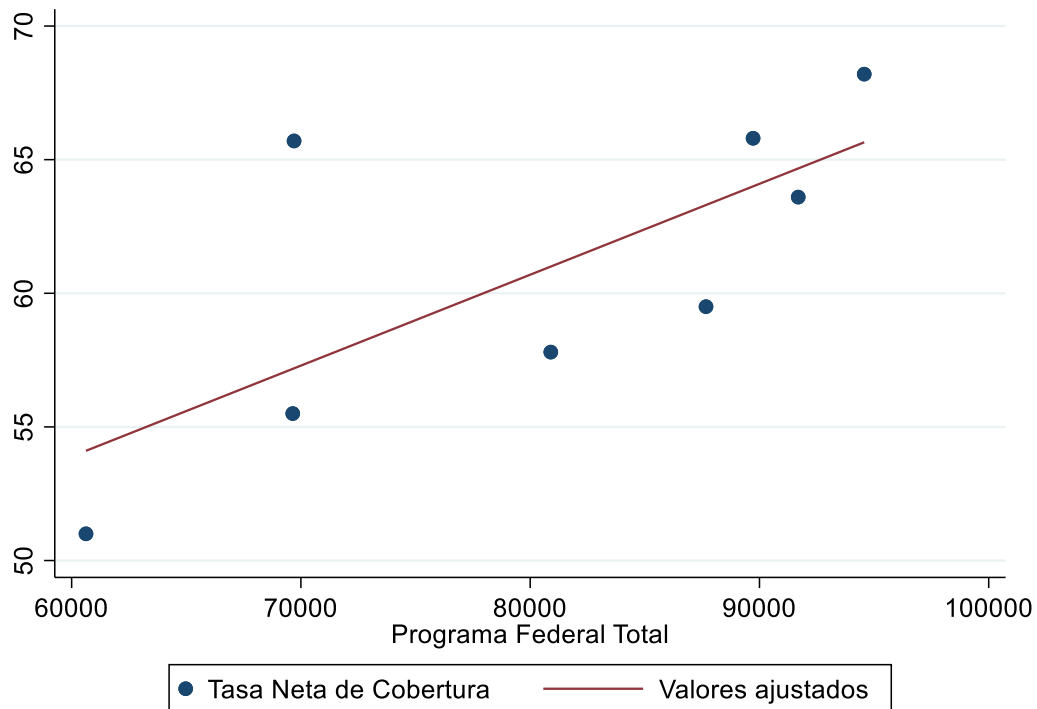
Coefficiente 0.0372764*

Error estándar 0.017291

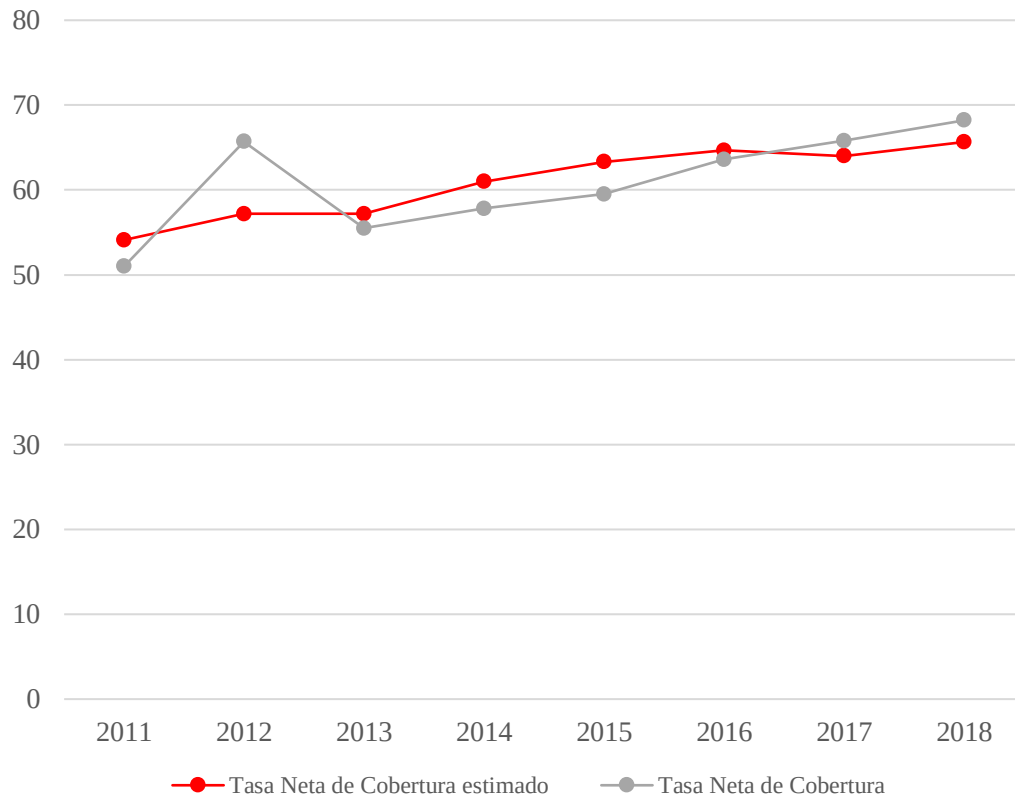
Constante 0.4718063

R-cuadrado 0.4365

Estimación 1



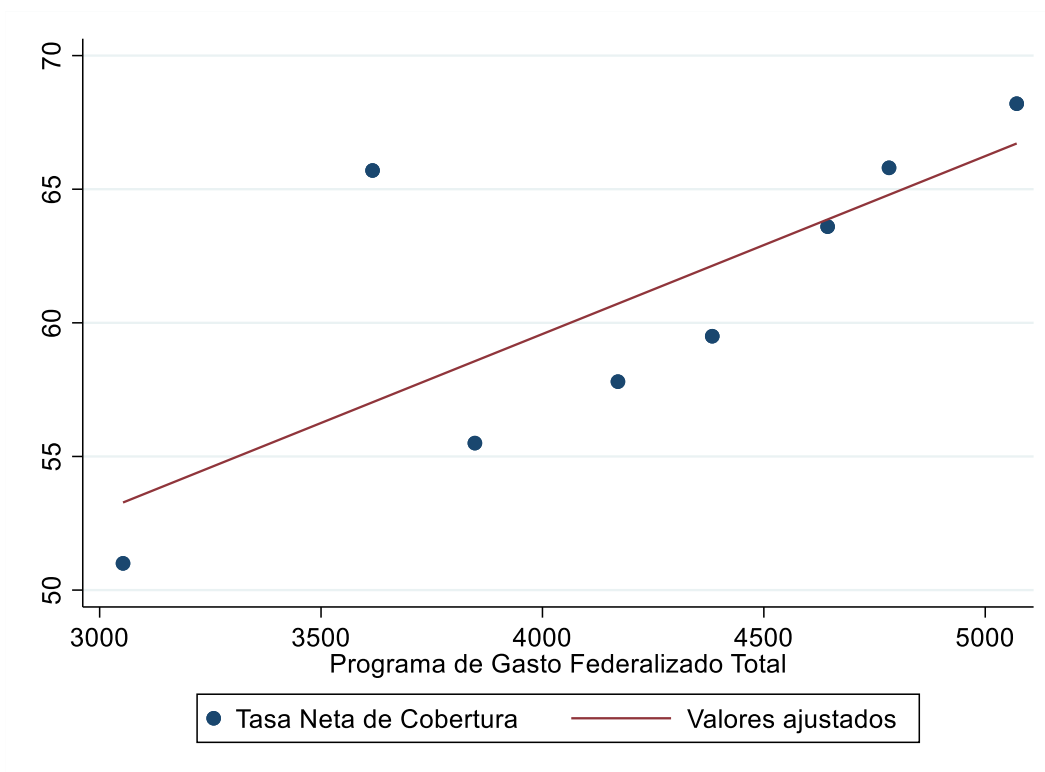
AÑO	Tasa Neta de Cobertura estimado	Tasa Neta de Cobertura	Gasto de Programa Federal Total
2011	54.10703	51	60627.07
2012	57.1932	65.7	69702.34
2013	57.17381	55.5	69645.32
2014	61.00104	57.8	80899.79
2015	63.30336	59.5	87670.06
2016	64.67057	63.6	91690.51
2017	64.001	65.8	89721.55
2018	65.65	68.2	94570.63



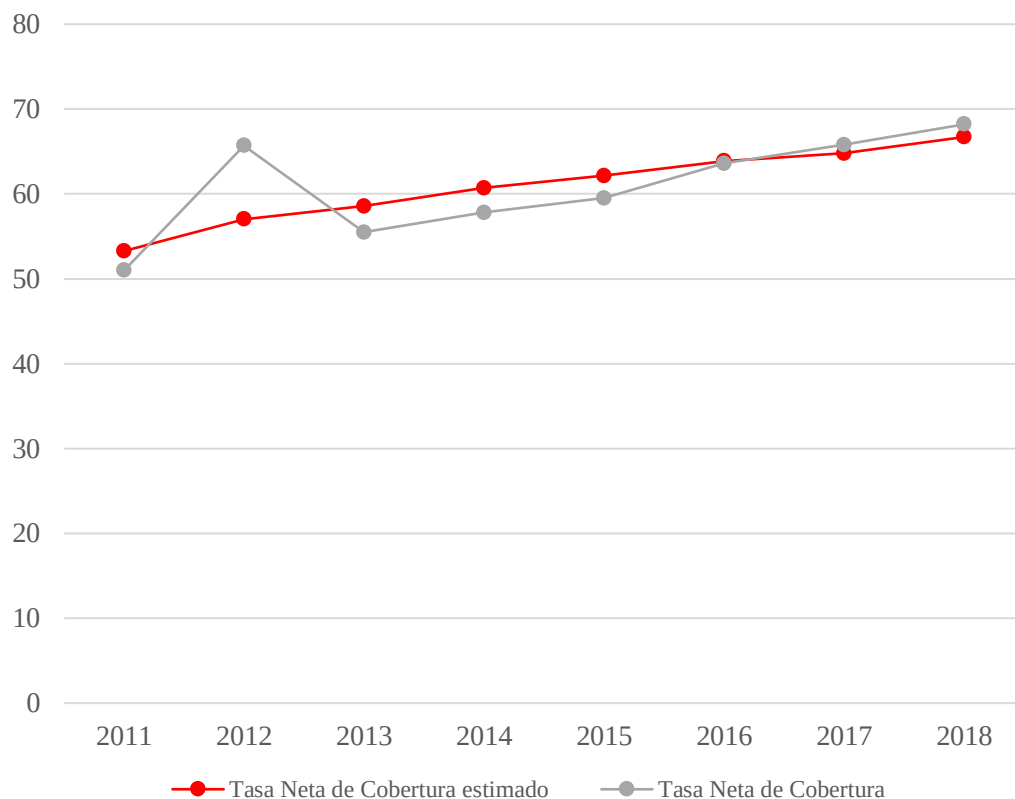
Para la estimación 1 se observa que el coeficiente es estadísticamente significativo al 5%. La magnitud del efecto parcial es de 0.0001357. Existe una relación positiva entre la Tasa Neta de Cobertura y el Programa Federal Total. El valor de R-cuadrado que es de .5112 esto significa que el Programa Federal explica hasta un 52% la tasa neta de cobertura.

Se observa que los valores estimados de la Tasa Neta de Cobertura no son muy diferentes de los valores reales en específico para los años 2013, 2016 y 2017.

Estimación 2



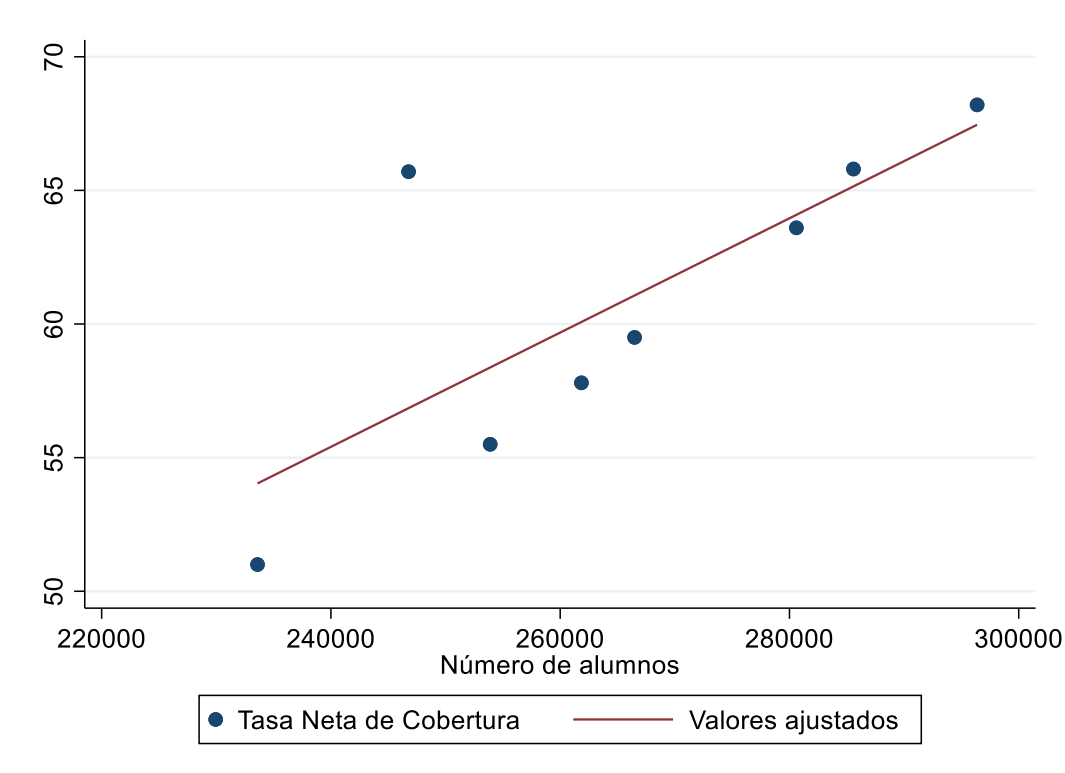
Año	Tasa Neta de Cobertura estimado	Tasa Neta de Cobertura	Programa de Gasto Federalizado Total
2011	53.27802	51	3052.88
2012	57.02888	65.7	3616.38
2013	58.56677	55.5	3847.42
2014	60.71711	57.8	4170.47
2015	62.13505	59.5	4383.49
2016	63.86957	63.6	4644.07
2017	64.79274	65.8	4782.76
2018	66.71186	68.2	5071.0733



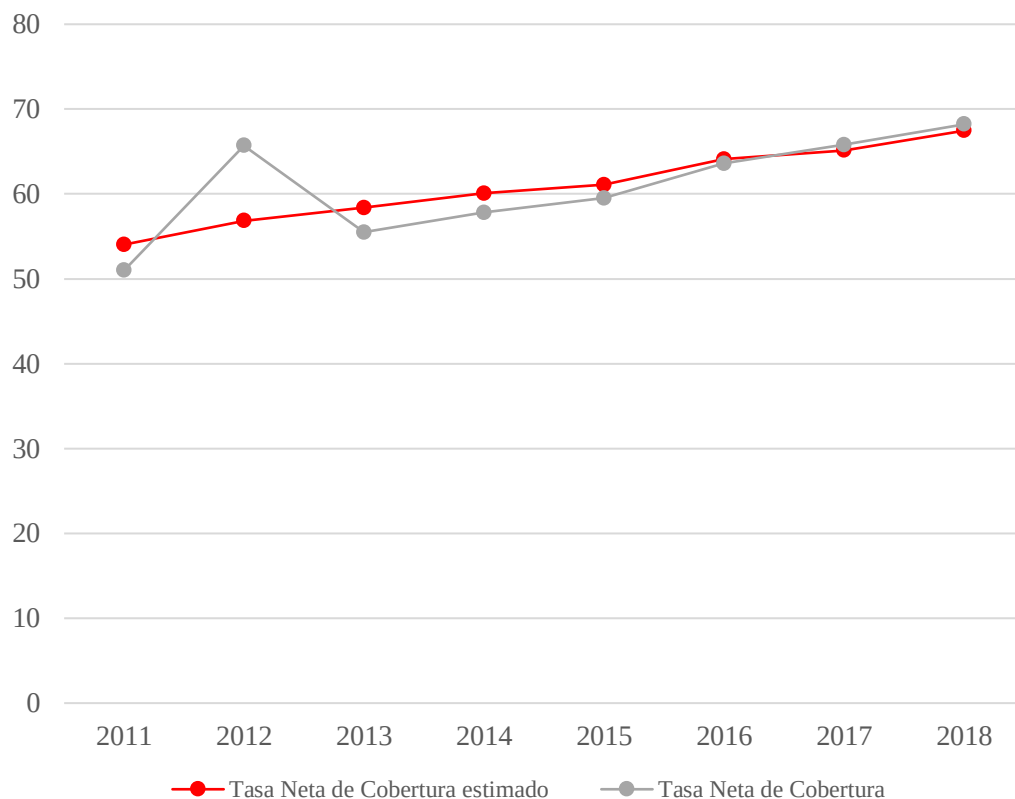
Para la estimación 2 se observa que el coeficiente es estadísticamente significativo al 5%. La magnitud del efecto parcial es de 0.0024114. Existe una relación positiva entre la Tasa Neta de Cobertura y el Programa de Gasto Federalizado Total. El valor de R-cuadrado que es de .5595, esto significa que el Programa de Gasto Federalizado explica hasta un 56% la tasa neta de cobertura.

Se observa que los valores estimados de la Tasa Neta de Cobertura no son muy diferentes de los valores reales y se acercan más a éstos en comparación con la estimación 1 para los años 2011, 2014, 2015, 2016, 2017 y 2018.

Estimación 3



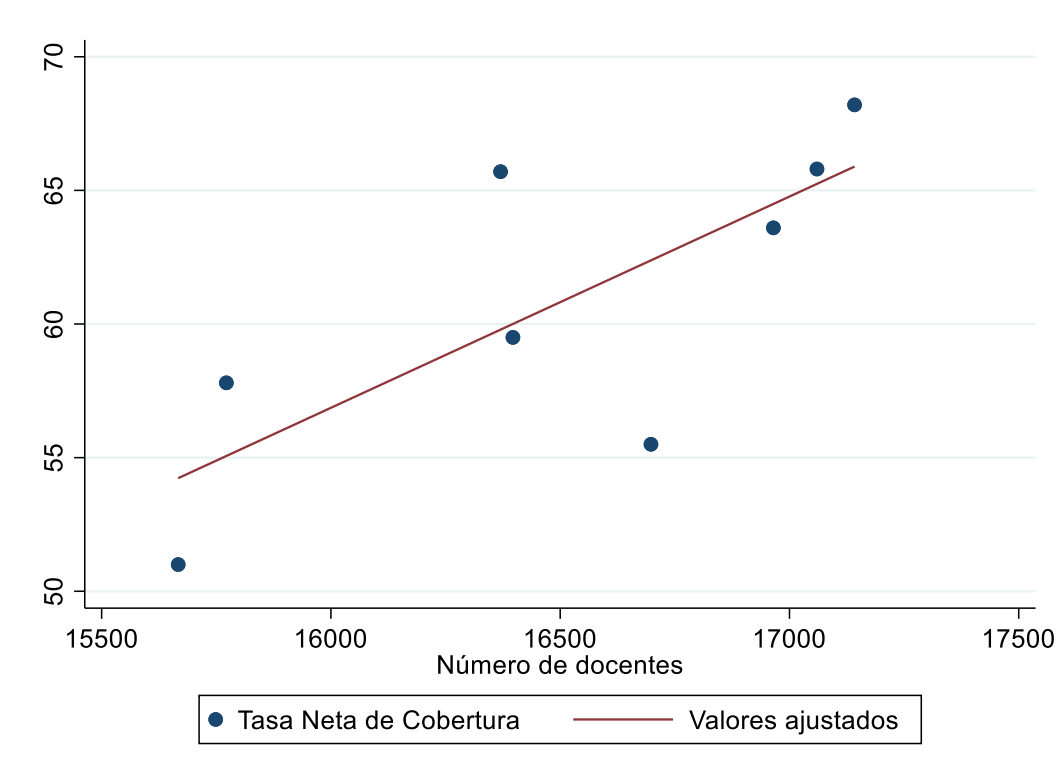
Año	Tasa Neta de Cobertura estimado	Tasa Neta de Cobertura	Número de alumnos
2011	54.03561	51	233599
2012	56.85327	65.7	246778
2013	58.37488	55.5	253895
2014	60.07629	57.8	261853
2015	61.06789	59.5	266491
2016	64.08588	63.6	280607
2017	65.15018	65.8	285585
2018	67.456	68.2	296370



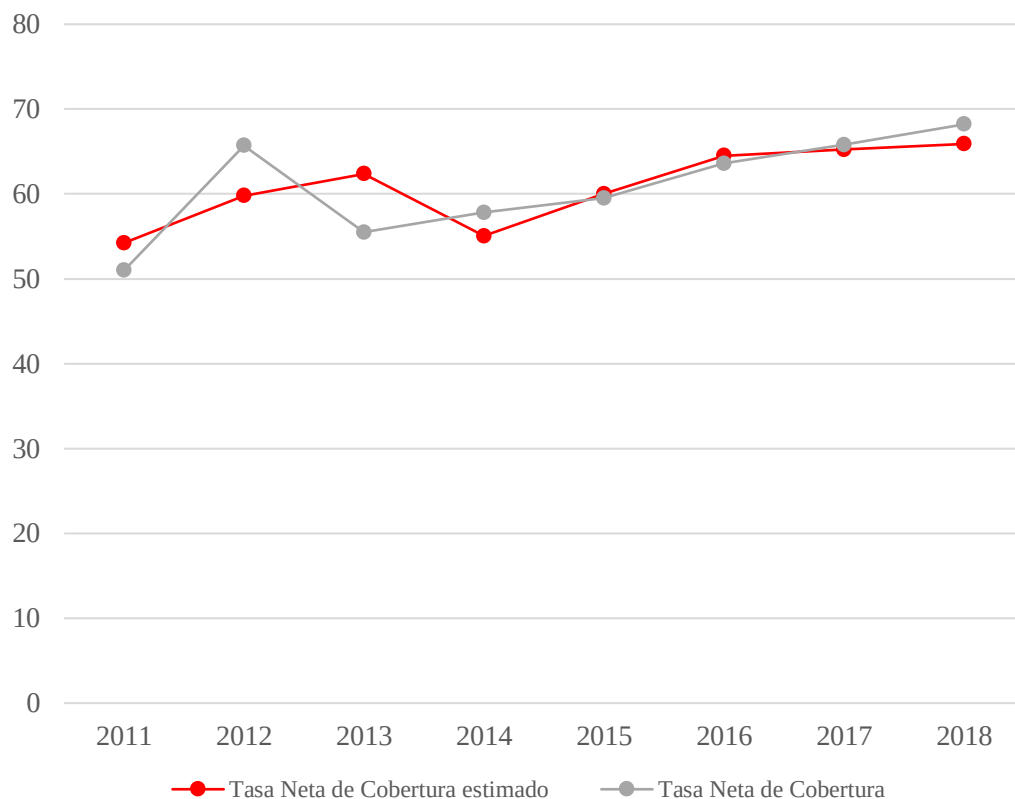
Para la estimación 3 se observa que el coeficiente es estadísticamente significativo al 5%. La magnitud del efecto parcial es de 0.000075. Existe una relación positiva entre la Tasa Neta de Cobertura y el Número de alumnos. El valor de R-cuadrado que es de .5754, esto significa que el número de alumnos explica hasta un 58% la tasa neta de cobertura.

Se observa que los valores estimados de la Tasa Neta de Cobertura no son muy diferentes de los valores reales y se acercan más a éstos en los años 2013, 2014, 2015, 2017 y 2018 en comparación con la estimación 2.

Estimación 4



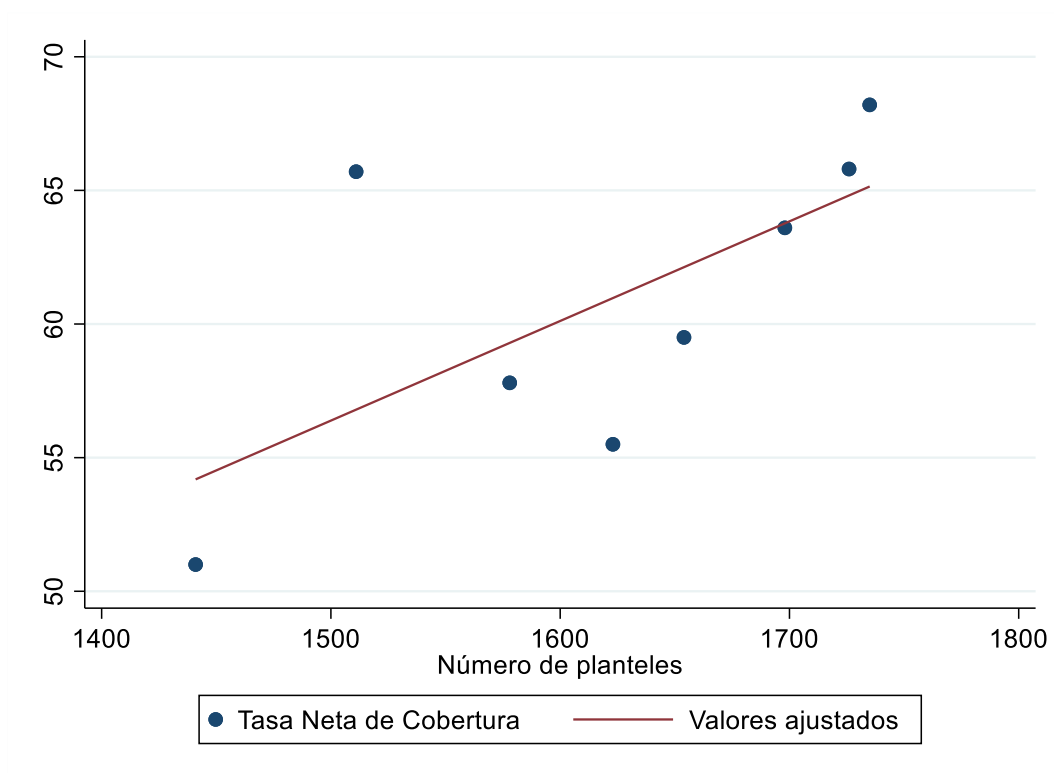
Año	Tasa Neta de Cobertura estimado	Tasa Neta de Cobertura	Número de docentes
2011	54.23325	51	15667
2012	59.78982	65.7	16370
2013	62.38236	55.5	16698
2014	55.06318	57.8	15772
2015	60.00323	59.5	16397
2016	64.49275	63.6	16965
2017	65.24364	65.8	17060
2018	65.89178	68.2	17142



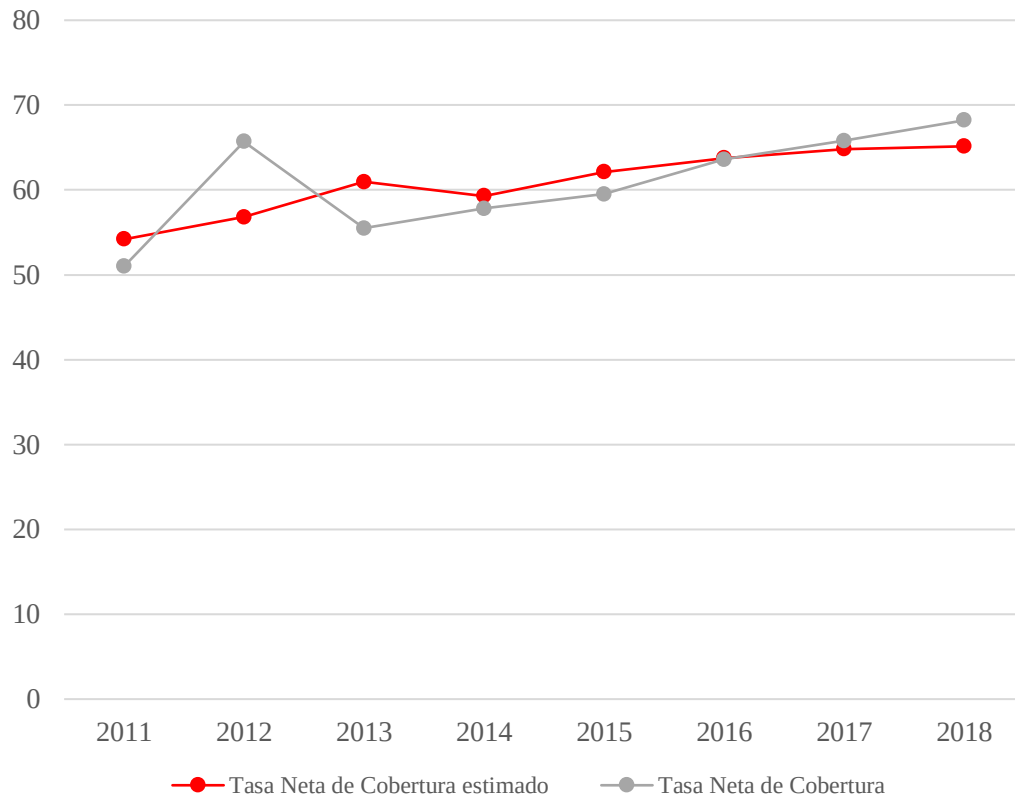
Para la estimación 4 se observa que el coeficiente es estadísticamente significativo al 5%. La magnitud del efecto parcial es de 0.0028257. Existe una relación negativa entre la Tasa Neta de Cobertura y el Número de docentes. El valor de R-cuadrado que es de .5660, esto significa que el número de docentes explica hasta un 57% la tasa neta de cobertura.

Se observa que los valores estimados de la Tasa Neta de Cobertura no son muy diferentes de los valores reales y se acercan más a éstos en los años 2012, 2015 y 2017 en comparación con la estimación 3.

Estimación 5



Año	Tasa Neta de Cobertura estimado	Tasa Neta de Cobertura	Número de plantilles
2011	54.18707	51	1441
2012	56.79642	65.7	1511
2013	60.97137	55.5	1623
2014	59.29393	57.8	1578
2015	62.12694	59.5	1654
2016	63.7671	63.6	1698
2017	64.81084	65.8	1726
2018	65.14632	68.2	1735



Para la estimación 5 se observa que el coeficiente es estadísticamente significativo al 10%. La magnitud del efecto parcial es de 0.017291. Existe una relación positiva entre la Tasa Neta de Cobertura y el Número de planteles. El valor de R-cuadrado que es de .4365, esto significa que el número de planteles explica hasta un 44% de la tasa neta de cobertura.

Se observa que los valores estimados de la Tasa Neta de Cobertura no son muy diferentes de los valores reales y se acercan más a éstos en los años 2011, 2013, 2014 y 2016 en comparación con la estimación 4.

Análisis factorial

Factor de Educación Básica

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	3.99271	0.88727	0.4991	0.4991
Factor2	3.10544	2.67439	0.3882	0.8873
Factor3	0.43106	0.12764	0.0539	0.9412
Factor4	0.30342	0.19098	0.0379	0.9791
Factor5	0.11244	0.07083	0.0141	0.9931
Factor6	0.04161	0.02829	0.0052	0.9983
Factor7	0.01332	0.01332	0.0017	1.00000
Factor8	0.000000	.	0.000000	1.00000

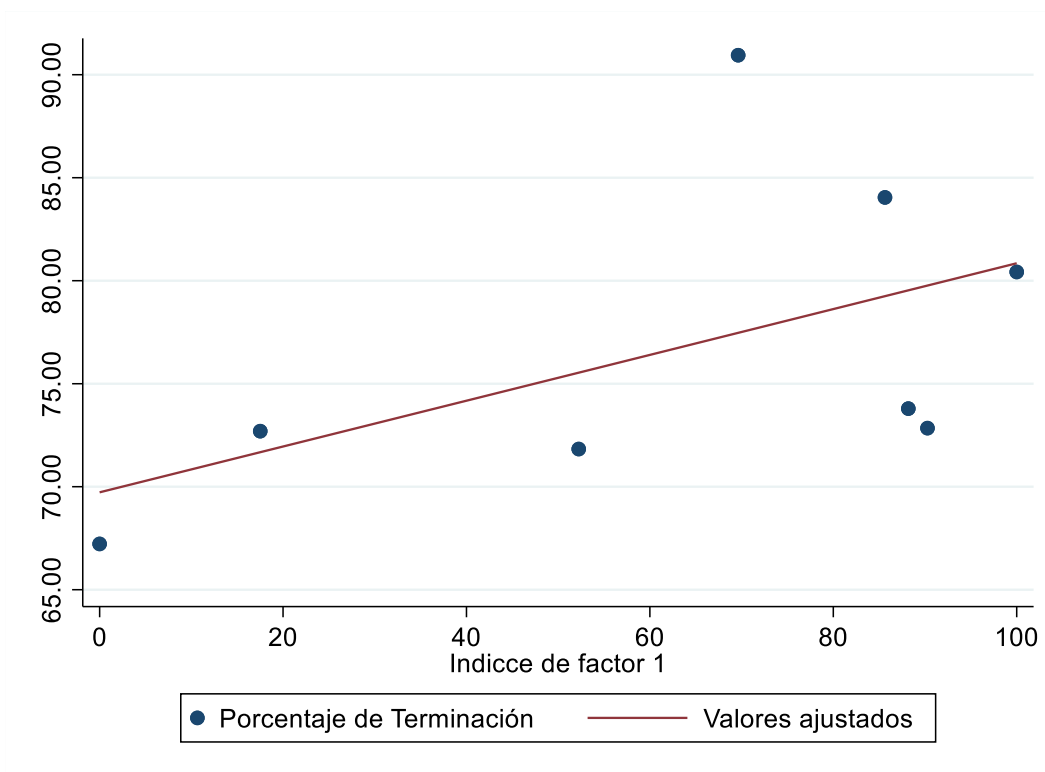
LR test: independent vs. saturated: $\chi^2(28) = 163.04$ Prob> $\chi^2 = 0.0000$

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Uniqueness
Porcentaje de becarios a los que se les emitieron apoyos monetarios para becas educativas	0.8778	0.381	0.1708	0.081	-0.1849	-0.1183	0.0191	0.00000
Porcentaje de becarias y becarios a los que se emitió apoyos para útiles escolares	0.3785	0.8644	0.2084	0.2226	-0.0205	0.1272	0.0069	0.00000
Gasto de Programa Federal Total	0.5636	0.5974	-0.5497	0.1497	0.0294	-0.0071	0.0008	0.00000
Programa de Gasto Federalizado Total	0.7823	-0.6037	-0.0426	-0.1004	0.034	0.045	0.0917	0.00000

Número de alumnos	0.5732	0.7451	0.1941	-0.1776	0.2142	-0.0326	-0.0035	0.00000
Número de docentes	0.6788	-0.7146	0.0764	0.1273	-0.0387	0.0607	-0.0365	0.00000
Número de planteles	0.7001	-0.6525	0.0516	0.2361	0.1488	-0.0533	-0.027	0.00000
Tasa Neta de Cobertura	0.9326	-0.012	-0.091	-0.3336	-0.0791	0.0422	-0.0492	0.00000

**Porcentaje de Terminación de
Educación Básica de los jóvenes
atendidos por el programa**

Índice de Factor 1	
Coefficiente	0.111191
Error estándar	0.0735169
Constante	69.72734
R-cuadrado	0.2760



No hay significancia por lo que se optó por no hacer pronósticos para el Porcentaje de Terminación de Educación Básica de los jóvenes atendidos por el programa

Factor de Educación Media Superior

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	4.60937	3.50996	0.6864	0.6864
Factor2	1.09941	0.22023	0.1637	0.8501
Factor3	0.87918	0.75229	0.1309	0.981
Factor4	0.12689	0.11997	0.0189	0.9999
Factor5	0.00692	0.00733	0.001	1.0009
Factor6	-0.00041	0.00523	-0.0001	1.0008
Factor7	-0.00564	.	-0.0008	1.00000

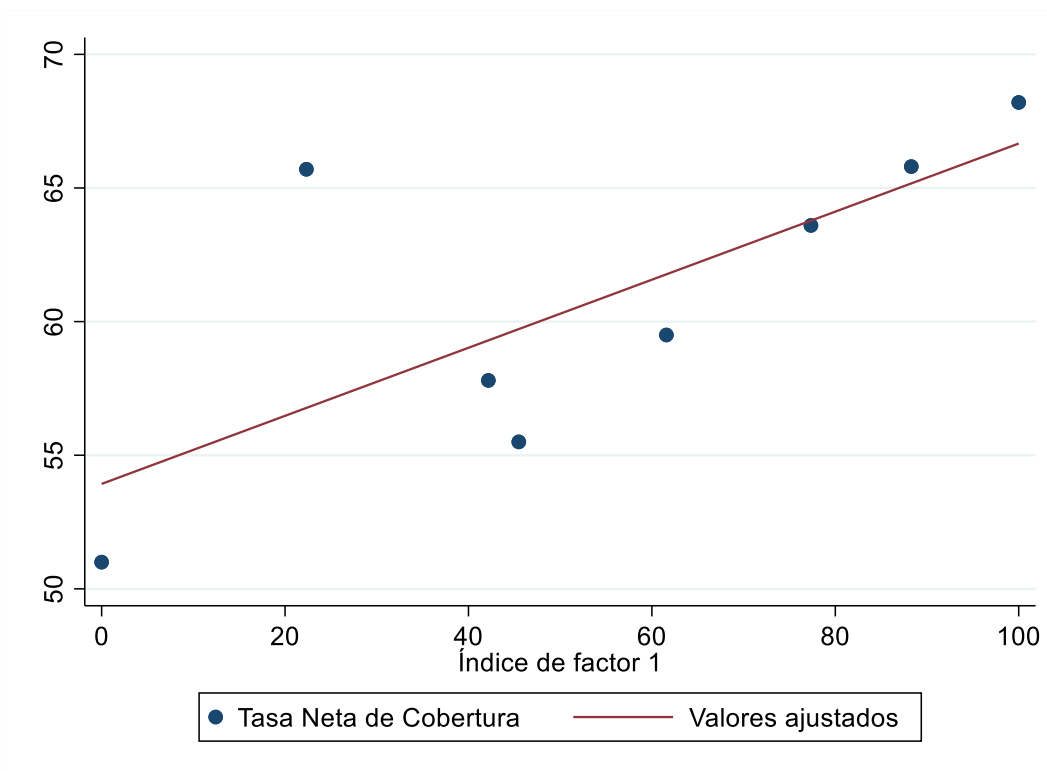
LR test: independent vs. saturated: $\chi^2(21) = 77.66$ Prob> $\chi^2 = 0.0000$

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Uniqueness
Porcentaje de becarios a los que se les emitieron apoyos monetarios para becas educativas	-0.1808	-0.1713	0.9017	0.0464	0.0105	0.1226
Porcentaje de becarias y becarios a los que se emitió apoyos para útiles escolares	-0.1984	0.9453	0.1028	0.1302	0.002	0.0396
Gasto de Programa Federal Total	0.9407	0.2559	0.1641	-0.1107	-0.057	0.0072
Programa de Gasto Federalizado Total	0.9847	0.1196	0.0636	-0.0924	0.0261	0.0028

Número de alumnos	0.9871	0.1383	-0.0431	-0.0224	0.0513	0.0015
Número de docentes	0.8575	-0.2474	-0.109	0.2941	-0.0119	0.1049
Número de planteles	0.9865	-0.128	0.1038	-0.0008	-0.0104	-0.0004

Regresión de Tasa de Neta de Cobertura de Educación Media Superior e Índice de Factor 1

Tasa Neta de Cobertura	
Índice de Factor 1	
Coefficiente	0.1273364**
Error estándar	0.0491547
Constante	53.92863
R-cuadrado	0.528

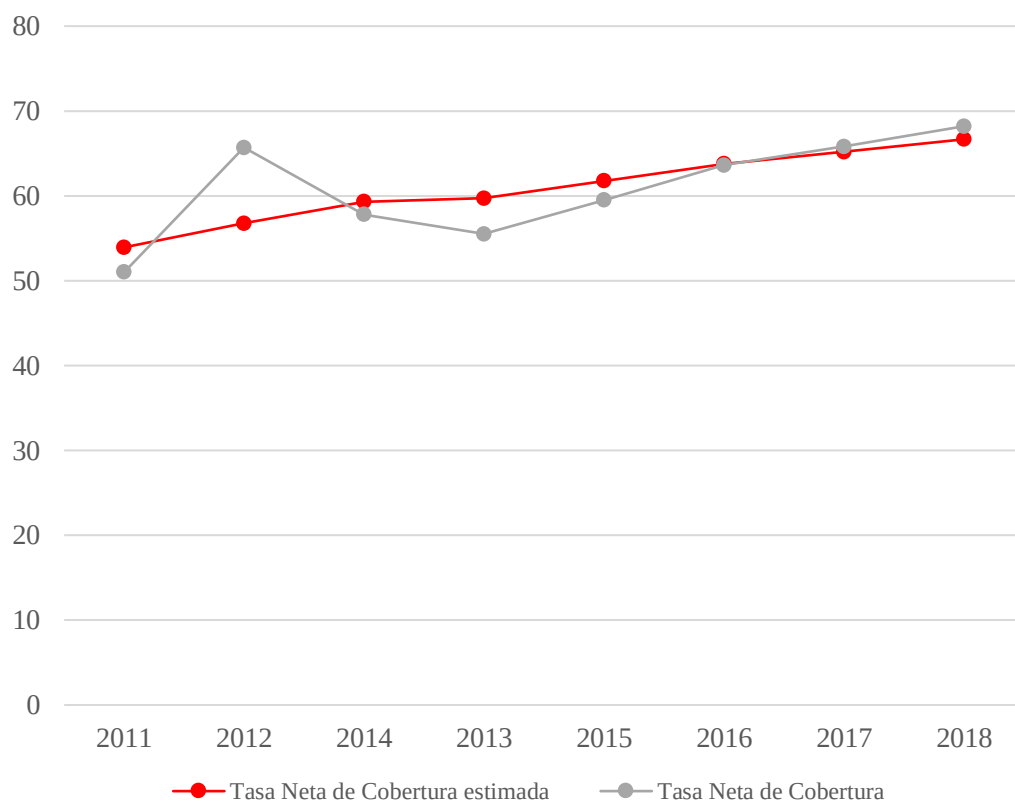


Para la estimación de Índice de Factor 1 se observa que el coeficiente es estadísticamente significativo al 5%. La magnitud del efecto parcial es de 0.0491547. Existe una relación positiva entre la Tasa Neta de Cobertura y

Porcentaje de becarios a los que se les emitieron apoyos monetarios para becas educativas, Porcentaje de becarias y becarios a los que se emitió apoyos para útiles escolares, Gasto de Programa Federal Total, Programa de Gasto Federalizado Total, Número de alumnos, Número de docentes y Número de planteles. El valor de R-cuadrado que es de .528, esto significa que las variables mencionadas anteriormente definen hasta un 53% la tasa neta de cobertura.

Estimación de Índice de Factor 1

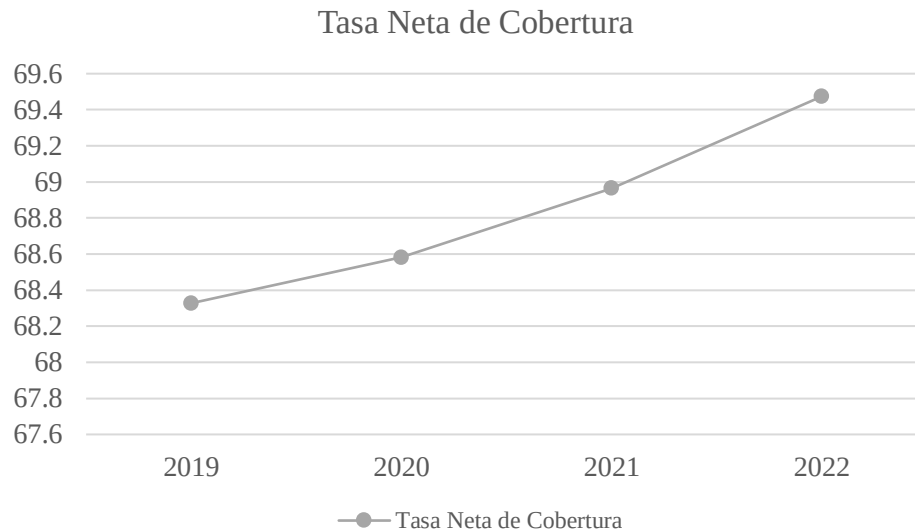
Año	Índice de Factor 1	Tasa Neta de Cobertura estimada	Tasa Neta de Cobertura
2011	0	53.92863	51
2012	22.32863	56.77188	65.7
2014	42.16861	59.29823	57.8
2013	45.48834	59.72095	55.5
2015	61.58233	61.77031	59.5
2016	77.34185	63.77707	63.6
2017	88.28593	65.17065	65.8
2018	100	66.66228	68.2



Pronósticos

Educación Media Superior

Año	Tasa Neta de Cobertura
2019	68.3273364
2020	68.5820092
2021	68.9640184
2022	69.473364



Conclusiones

De las variables que se tomaron en cuenta para el análisis de mínimos cuadrados ordinarios en Educación Básica, se observa que hay significancia entre las variables Programa de Gasto Federalizado Total, Número de docentes y Número de planteles y la variable Porcentaje de Terminación de los Jóvenes atendidos por el programa. La variable Número de Docentes muestra mayor significancia. El número de docentes ha aumentado del año 2011 al 2018 y se puede decir que esta variable explica hasta un 94% del Porcentaje de Terminación. En segundo lugar se encuentra la variable Número de Planteles, que explica un 84% el porcentaje de terminación. En tercer lugar se observa que la variable Programa de Gasto Federalizado Total, cuyo objetivo es mantener la operación de servicios educativos, explica un 80% del porcentaje de terminación.

En el caso de la Educación Media Superior se observa que hay significancia para las variables Programa Federal Total, Programa de Gasto Federalizado Total, Número de alumnos, Número de docentes y Número de planteles y la variable Tasa Neta de Cobertura. La variable Número de alumnos muestra mayor significancia respecto a las 4 variables restantes, seguida del Número de Docentes. El Gasto Federalizado muestra relación positiva y explica un 56% la tasa de cobertura. El

Gasto Federal proporciona subsidios, provee servicios públicos y evalúa procesos administrativos con el fin de apoyar la educación obligatoria. Uno de los propósitos del Gasto Federal es incrementar la tasa de cobertura y se puede observar en la estimación número 1 que hay una relación positiva entre el Gasto Federal y la Tasa Neta de Cobertura y que éste explica hasta un 52% de dicha tasa. La variable Número de Planteles es la menos significativa comparada con el resto y explica un 44% de la tasa de cobertura.

Se observa en la tabla de pronósticos que la Tasa Neta de Cobertura aumentará para los siguientes cuatro años a partir del último periodo del que se tiene información, es decir del año 2018. Esto representa un significado positivo para la Educación Media Superior.

Se observa que el Gasto Federal y Programa de Gasto Federalizado arroja resultados positivos para la Educación Obligatoria en el estado de Puebla, estos se pueden ver reflejados en el aumento de la Tasa de Terminación de Educación Básica y la Tasa Neta de Cobertura de Educación Media Superior.

Referencias

- Bárcena, M., Garín, A., Martín, A., Tusell, F., & Unzueta, A. (2017). Un simulador para asistir en la enseñanza del teorema de Bayes. *IN-RED 2017: III Congreso Nacional de Innovación Educativa y Docencia en Red* (págs. 15-23). Valencia: Editorial Universitat Politècnica de Valencia.
- Bernal, H. (Abril de 2017). Obtenido de Proyecto de Inteligencia de Negocios para la Dirección de Recursos Materiales y Financieros de SIEM. Caso Departamento de Almacén. Toluca, México. 2016.:
http://ri.uaemex.mx/bitstream/handle/20.500.11799/67695/TTG_Hugo_Bernal.pdf?sequence=1&isAllowed=y
- Coordinación Nacional de Becas para el Bienestar Benito Juárez. (18 de Julio de 2016). *Gobierno de México*. Obtenido de ¿Qué es PROSPERA Programa de Inclusión Social?:
<https://www.gob.mx/becasbenitojuarez/documentos/que-es-prospera#:~:text=El%20Programa%20de%20Educaci%C3%B3n%2C%20Salud,terceras%20p%20artes%20eran%20comunidades%20ind%C3%ADgenas.>
- Coordinación Nacional de Becas para el Bienestar Benito Juárez. (01 de Marzo de 2021). *Gobierno de México*. Obtenido de Beca Universal para estudiantes de Educación Media Superior Benito Juárez: <https://www.gob.mx/becasbenitojuarez/articulos/beca-benito-juarez-para-jovenes-de-educacion-media-superior-216589>
- Fischer, F., Miller, G., & Sidney, M. (2017). *Handbook of Public Policy Analysis: Theory, Politics, and Methods*. Boca Raton: CRC Press.

- IBM. (2021). *Regresión lineal*. Obtenido de IBM: <https://www.ibm.com/mx-es/analytics/learn/linear-regression>
- Instituto Nacional para la Evaluación de la Educación. (2018). *8.1 Elementos normativos y presupuestales del financiamiento público a la educación*. Obtenido de https://www.inee.edu.mx/medios/informe2018/04_informe/capitulo_0801.html
- Instituto Nacional para la Evaluación de la Educación. (2018). *8.4 Distribución del gasto federal ejercido por tipo de programa y rubro*. Obtenido de https://www.inee.edu.mx/medios/informe2018/04_informe/capitulo_0804.html
- Instituto Nacional para la Evaluación de la Educación. (17 de junio de 2019). *AR03b.2 – Gasto Federal Ejercido en Educación Obligatoria*. Obtenido de <https://www.inee.edu.mx/evaluaciones/panorama-educativo-de-mexico-isen/ar03b-2-gasto-federal-en-educacion/>
- Instituto Nacional para la Evaluación de la Educación. (17 de Junio de 2019). *AT01b – Tasa neta de cobertura*. Obtenido de <https://www.inee.edu.mx/evaluaciones/panorama-educativo-de-mexico-isen/at01b-tasa-cobertura/>
- Johnson, R., & Wichern, D. (2014). *Applied Multivariate Statistical Analysis*. Pearson.
- Khan, R., Nadeem, A., & Ali, A. (2019). Business Analytics; A Framework. *International Journal of Computer Technology & Applications*, 102-108.
- Khandker, S., Koolwal, G., & Samad, H. (2009). *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washintong DC: World Bank Publications.
- Lanot, G. (Enero de 2017). *IZA World of Labor*. Obtenido de Modelado económico y de máxima verosimilitud: <https://wol.iza.org/articles/maximum-likelihood-and-economic-modeling/long>
- Longo, J., & Mcnutt, K. (2018). From policy analysis to policy analytics. En L. Dobuzinskis, & M. Howlett, *Policy Analysis in Canada* (págs. 369-391). Gran Bretaña: Policy Press.
- Osborne, J. (Junio de 2012). *Logits and tigers and bears, oh my! A brief look at the simple math of logistic regression and how it can improve dissemination of results*. Obtenido de Practical Assessment, Research & Evaluation: <http://pareonline.net/getvn.asp?v=17&n=11>
- Secretaría de Bienestar. (30 de Noviembre de 2015). *Gobierno de México*. Obtenido de PROSPERA Programa de Inclusión Social: <https://www.gob.mx/bienestar/acciones-y-programas/prospera-programa-de-inclusion-social-15908>
- Secretaría de Bienestar. (14 de Enero de 2016). *Gobierno de México*. Obtenido de Conoce todo sobre Prospera: <https://www.gob.mx/bienestar/es/articulos/conoce-todo-sobre-prospera>
- Secretaría de Bienestar. (27 de Julio de 2017). *Gobierno de México*. Obtenido de Sedesol y Prospera impulsan la educación de las y los jóvenes mexicanos:

<https://www.gob.mx/bienestar/es/articulos/sedesol-y-prospera-impulsan-la-educacion-de-las-y-los-jovenes-mexicanos>

Taddy, M. (2019). *Business Data Science: Combining Machine Learning and Economics to Optimize, Automate, and Accelerate Business Decisions*. McGraw Hill Professional.

Wooldridge, J. (2013). *Introductory Econometrics: A Modern Approach, Fifth Edition*. California: Cengage Learning.