

Análisis del impacto de las transferencias condicionadas en educación del programa Familias en Acción Fase III, desde enfoques econométricos y de machine learning: Bogotá D.C. Colombia

Mónica Tatiana Barahona Varón¹

Clara Inés Molina de Barbosa²

Claudia Gimena Rincón³

Resumen

Esta investigación se orienta a realizar un análisis crítico del impacto de las transferencias condicionadas del programa Familias en Acción Fase III en centros urbanos, como es el caso de Bogotá D.C, frente al componente de educación, específicamente, en asistencia y permanencia en el sistema educativo. Los resultados de la investigación permitirán evidenciar el cumplimiento de las metas y objetivos del programa, de tal forma que esta información sirva de insumo para el diseño de agendas de política pública en el futuro. El marco de referencia que orienta el presente estudio se fundamenta en conceptos tales como Transferencias Monetarias Condicionadas (TMC), asistencia y permanencia en el sistema educativo, y modelos de evaluación de impacto. La metodología aplicada tiene un enfoque cuantitativo de alcance correlacional, de tipo no experimental y con diseño transeccional. Se toma como referencia los datos de la Encuesta Nacional de Calidad de Vida (ECV) de los años 2012 y 2020 con muestras representativas para Bogotá D.C. A partir de los datos se aplican técnicas econométricas y de Machine Learning para determinar el impacto del programa. Los resultados de los avances obtenidos hasta el momento consisten en la generación de un modelo de regresión logística, que implicó la aplicación de técnicas de balanceo de datos, selección de variables y validación sobre un conjunto de datos de testeo; para clasificar los registros de las bases de datos y conformar los grupos de tratamiento y control que se emplearán en la evaluación de impacto mediante un modelo de diferencias en diferencias.

Conceptos Clave: 1. Políticas públicas, 2. transferencias condicionadas, 3. impacto, 4. educación, 5. centros urbanos.

Introducción

La identificación y la necesidad de análisis de políticas públicas y de programas, permiten generar agendas para la solución de problemas sociales como la educación y la caracterización de sus brechas. Por lo tanto, cobra importancia la realización de un análisis desde una visión crítica de las políticas, como en este caso del programa de transferencias condicionadas en educación “Familias en Acción”, implementado por el gobierno colombiano en respuesta a la asistencia y permanencia en el sistema educativo de los hijos de las familias de menores ingresos del país. Uno de los principales retos para los analistas de políticas y

¹ MSc en Analítica para la Inteligencia de Negocios. Universidad de Bogotá Jorge Tadeo Lozano. monicat.barahonav@utadeo.edu.co

² MSc en Ciencias de la Educación. Universidad de San Buenaventura, Bogotá. cmolina@usbbog.edu.co

³ Especialista en Finanzas Públicas. Universidad de Bogotá Jorge Tadeo Lozano. claudiag.rincon@utadeo.edu.co

programas es que mediante sus investigaciones se puedan apoyar a los procesos de construcción de las políticas públicas. Por esto es importante determinar la influencia de las políticas públicas en la acumulación del stock de capital humano, que se evidencia a través de la investigación económica en aspectos como la educación en su calidad, constituidos por la asistencia y permanencia en el sistema educativo y la calidad de los conocimientos adquiridos.

El logro educativo es una dimensión crítica del bienestar humano que contribuye a la equidad, no solo por derecho propio, sino también como un aporte importante al empoderamiento y al desarrollo de las capacidades de una persona, para lograr su participación plena en la sociedad. Además, la educación es un medio para el mejoramiento de los ingresos y de la calidad de vida de las personas, por consiguiente, si estas carecen de educación o habilidades básicas de alfabetización se enfrentan a mayores riesgos de desempleo o informalidad en el trabajo y son más propensas a vivir en la pobreza de manera permanente.

Dentro de los principales programas implementados en la lucha contra la pobreza, los gobiernos han implementado las Transferencias Monetarias Condicionadas-TMC, con el fin de reducir la pobreza estructural en el largo plazo, mediante la inversión en capital humano que permite aumentar la cualificación para el trabajo en las personas beneficiarias, con el fin de llevarlos a obtener mejores oportunidades laborales. Dallorso (2013), explica desde la teoría del capital humano cómo las TMC buscan reducir la pobreza estructural en el largo plazo, partiendo de la definición de inversión en capital humano como aquellos costos en que incurren las personas para incrementar sus propias capacidades, que aumentarán su cualificación para el trabajo y finalmente las conducirá a obtener mejores oportunidades laborales. El Banco Mundial (2022), en su último informe sobre pobreza y prosperidad compartida recalca la necesidad por parte de los responsables de la formulación de políticas públicas, de implementar medidas que permitan crecer las economías en beneficio de los ingresos de los más pobres, señalando las transferencias de efectivo como mecanismo eficaz para apoyar a los grupos pobres y vulnerables.

En Colombia, tras sufrir drásticamente cambios sociales, económicos y políticos a finales de los años 1990, se diseñó y ejecutó el Programa Familias en Acción (FA) como respuesta a los efectos causados por dicha crisis. La principal razón por la que surgió FA, según indica Llano (2014), fue que gran cantidad de hogares estaban en situación de pobreza dado que su fuente de recursos era inestable y sus ingresos se veían afectados constantemente; es por ello, que como estrategia para mitigar esta problemática el gobierno decidió implementar programas enfocados a aumentar la inversión de capital humano, ofreciendo subsidios condicionados que aseguraran la asistencia de los niños y jóvenes al colegio con el fin de aumentar los años de escolaridad y reducir la deserción escolar. Desde su implementación, el programa ha tenido cuatro fases de acuerdo con los cambios en la metodología SISBEN, que es el principal instrumento de focalización de las potenciales familias beneficiarias del programa, la fase 1 implementada entre los años 2001 y 2006, la fase 2 estuvo vigente desde 2007 hasta 2011, la fase 3 abarcó el período de 2012 a 2020 y actualmente se encuentra vigente la fase 4, que inició su implementación en 2021. Las fases del programa Familias en Acción evolucionan paralelamente con respecto a la evolución del índice SISBEN, que en lo relativo al cubrimiento de territorios se pasa de una cobertura de

municipios con poblaciones hasta 100,000 habitantes en su primera fase a coberturas por descomposición del índice de SISBEN para los 32 departamentos de Colombia y sus respectivas zonas urbanas y rurales en las tres últimas fases. Lo anterior ha surgido a partir del reconocimiento de diferencias regionales en las condiciones socioeconómicas de la población, con el fin de reflejar mejor la realidad de cada territorio. “A través de Familias en Acción, Prosperidad Social entrega un incentivo monetario a las familias inscritas, para fomentar la asistencia y permanencia escolar de los niños, niñas o adolescentes menores de 18 del hogar beneficiario, así como el acceso al sistema de salud, previa verificación de cumplimiento de requisitos por parte de la entidad” (Departamento Nacional de Planeación-DNP, 2021).

La calidad de la educación a través de la política pública debe lograr eficiencia y eficacia en el desarrollo de programas y proyectos. La eficacia en las políticas es el grado de alcance de los objetivos propuestos, por lo tanto, estos deben ser claros, delimitados en el tiempo que se deben cumplir, para generar determinado producto y la calidad esperada en el impacto de la ejecución de las políticas (Unesco, 2007). En este sentido resulta interesante analizar si efectivamente el programa de transferencias condicionadas para la educación presenta resultados que son acordes a las metas establecidas en las fases de ejecución y por esto la evaluación de un centro urbano como es el caso de Bogotá D.C. que es la capital de Colombia, constituye un eje orientador de cómo este programa a nivel de centros urbanos ha presentado impactos en la población afectada. Es importante aclarar que la ciudad de Bogotá ingresó al programa en la segunda fase, pero se evalúa la tercera fase porque se supone que la ejecución del programa ya se encuentra más consolidada. Con la aplicación de la tercera fase del Programa Familias en Acción, lo que se esperaba en primer lugar era reducir las brechas regionales urbano-rurales persistentes y de ese mismo modo, llegar de manera diferencial a cada una de las poblaciones según sus necesidades.

Adicional a lo anterior, el programa seguiría manteniendo la misma estructura de programa TMC a través del cumplimiento de los objetivos de formación de capital humano. Considerando el interés específico en determinar los impactos del programa en el aspecto educativo como eje central de formación de capital humano, se esperaría encontrar un progreso en las condiciones educativas de los niños y jóvenes entre 7 y 18 años pertenecientes al programa, en cuanto a la asistencia y permanencia en el sistema educativo de las familias de más bajos ingresos de la ciudad de Bogotá, por lo cual, se plantea la siguiente pregunta de investigación: ¿Cuál ha sido el impacto del programa Familias en Acción en el incremento de la asistencia y permanencia en el sistema educativo en la ciudad de Bogotá en la tercera fase del programa comprendida entre 2007 y 2019?

De acuerdo con lo anterior, el objetivo de esta investigación consiste en analizar el impacto de las transferencias condicionadas en educación del programa Familias en Acción Fase III desde enfoques econométricos y de analítica de datos en Bogotá D.C. Colombia.

Antecedentes

En desarrollo de la investigación, se inicia con una revisión de literatura de los conceptos y teorías que fundamentan esta investigación y que se abordan a continuación:

- **Transferencias Monetarias Condicionadas**

Desde la política pública se han identificado herramientas para reducir el riesgo de deserción escolar, dentro de las que se encuentran los programas de TMC, estos se encuentran condicionados a que los hogares garanticen la permanencia de sus hijos en el sistema escolar, y ayudan a resolver el problema de la deserción escolar asociada con la falta de recursos para enviar a los hijos a la escuela. Los programas de TMC plantean un mecanismo de condicionalidad para la entrega de una transferencia monetaria a una familia, a cambio de la escolarización y del cumplimiento de controles de salud de niños y adolescentes. Es importante enfatizar en su objetivo indirecto de reducción de la pobreza en dos temporalidades: en el corto plazo busca aliviar la pobreza monetaria con un ingreso complementario al hogar que a su vez impacta el consumo global, y en el largo plazo busca romper con la transmisión intergeneracional de la pobreza gracias a las condiciones que plantea sobre el capital humano, que permiten el aumento de los niveles de escolaridad y nutrición, lo cual se asume aumentará las posibilidades de mejores ingresos futuros de los niños, llevando así a la superación de las trampas de la pobreza (Ayola B., 2018).

Estos programas presentan diferencias entre un país y otro, pero en esencia comparten algunas características: los criterios para su asignación se basan en el nivel de ingreso u otras medidas de pobreza, las transferencias se pagan de manera periódica en el transcurso del año, y los pagos se condicionan al cumplimiento de requisitos en asistencia a la escuela, asistencia de los menores a controles de salud y participación en reuniones comunitarias, entre otros. Por otra parte, los pagos se hacen a las madres de familia y están sometidos a controles con el fin de evitar desviaciones hacia población que no corresponda o que no cumpla las condiciones para acceder al mismo (Cecchini, S. y Madarriaga A., 2011).

En Colombia, el Programa Familias en Acción se diseñó inicialmente como un programa temporal enfocado en municipios rurales y pequeños, es decir aquellos con menos de 100,000 habitantes. Sin embargo, a partir de 2007 se expandió a municipios con más de 100,000 habitantes y actualmente es un programa permanente y de cobertura nacional, sustentado por la ley 1532 de 2012 (Medellín, 2015). Este programa tiene como objetivo “contribuir a la reducción, superación y prevención de la pobreza y la desigualdad de ingresos, a la formación de capital humano y al mejoramiento de las condiciones de vida de las familias pobres y vulnerables mediante un complemento al ingreso” (DPS, 2013).

La población objetivo son todas las familias del país en condición de pobreza y vulnerabilidad con integrantes menores de 18 años, dicha condición de pobreza se define utilizando un índice multidimensional que estima el estándar de vida de las familias. El principal instrumento de focalización del programa es el Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales -SISBEN-, son elegibles para el programa aquellas familias calificadas como pobres de acuerdo con su puntaje en el SISBEN. Adicionalmente, son elegibles las familias en situación especial de vulnerabilidad que incluyen pobreza extrema (Red Unidos), desplazamiento forzado (Registro Único de Víctimas) y pertenencia a comunidades indígenas (Medellín, 2015b).

Familias en Acción orienta sus acciones a la formación de capital humano en los menores pertenecientes a los hogares beneficiarios, mediante la entrega de dos tipos de incentivos monetarios en cumplimiento de los compromisos en salud y en educación. El incentivo en salud se entrega a las familias con niños y niñas menores de 6 años, con el fin de

complementar el ingreso familiar, dirigido a mejorar la salud de niños y niñas, durante la etapa crítica de su crecimiento. Se entrega un solo incentivo por familia, independientemente del número de niños que se encuentren en este rango de edad. El incentivo en educación se otorga a las familias con niños, niñas y adolescentes en edad escolar, que cursen los grados de transición a undécimo, con el fin de estimular la asistencia escolar, aumentar los años de escolaridad y reducir la deserción escolar. Se entrega un incentivo por cada niño, niña y adolescente, con un máximo de tres beneficiarios por familia y su entrega está condicionada al cumplimiento de compromisos de corresponsabilidad en educación, consistentes en la matriculación de los niños, niñas y adolescentes en los establecimientos educativos, así como el cumplimiento de asistencia mínimo al 80% de las clases programadas por bimestre escolar. La periodicidad de entrega de los incentivos es bimestral, sujeto al cumplimiento de la corresponsabilidad en salud y educación de las familias participantes. Es importante mencionar, que el valor de los incentivos que se entregarán a las familias para su uso en salud y educación se define con base en dos criterios: por grupos de municipios de intervención y para el incentivo de educación, de acuerdo con el grado escolar que cursa el estudiante (Medellín, 2015c).

Para el proceso de inscripción por medio del cual las familias focalizadas como potenciales beneficiarios del programa hacen efectiva su vinculación voluntaria a Familias en Acción, deben quedar registradas en el Sistema de Información de Familias en Acción-SIFA. Algunos de los requisitos para participar en el proceso operativo de inscripción es encontrarse incluido en la base de datos empleada para la focalización, según grupo poblacional: SISBEN, SIUNIDOS, RUV o listados censales indígenas; que la titular haya sido seleccionada en el proceso de focalización y que además cuente en su núcleo familiar con beneficiarios menores de 18 años. Los titulares serán los responsables directos del cumplimiento de compromisos para el acceso a los incentivos de salud y educación otorgados por el Programa (DPS, 2017).

- **Evaluación de impacto**

La revisión de literatura de evaluaciones de impacto realizadas a los programas de TMC a partir de su implementación en diferentes países de América Latina se realiza con el fin establecer las diferentes metodologías utilizadas de medir los impactos que han tenido en aspectos de la calidad de vida como es el caso de la educación en el corto y mediano plazo, lo que a su vez debe contribuir a la superación de las barreras de pobreza en el largo plazo. Las mismas abordan diferentes metodologías dependiendo del enfoque de la evaluación, el horizonte temporal a evaluar, la disponibilidad de información y los objetivos planteados. Para los propósitos del presente trabajo, los antecedentes revisados corresponden a metodologías y resultados obtenidos para variables de educación.

Una evaluación de impacto realizada para un período reciente a su implementación es la realizada por Mata y Hernández (2015) al programa Avancemos en Costa Rica, enfocada en evaluar si el programa cumplió su objetivo inmediato de evitar la deserción en secundaria de jóvenes que se encontraban en situación de pobreza o vulnerabilidad social durante el período 2006-2007, en ésta se utilizaron dos metodologías cuasiexperimentales: Propensity-Score Matching (PSM) y diferencias en diferencias, concluyendo que el impacto de

Avancemos fue positivo en las variables deserción escolar y de reinserción de los jóvenes que regresaron a estudiar.

Otra evaluación al programa Avancemos que abarca un período de mayor madurez del programa fue realizada por Muñoz (2016), orientada a determinar si el beneficio monetario del programa influía en la decisión de abandonar la educación formal en estudiantes de 12 a 25 años que eran hermanos o hermanas de una persona beneficiaria del programa para el año 2013. En el desarrollo de ésta se aplicaron modelos de regresión logística, los cuales permitieron encontrar cinco variables significativas para explicar el abandono educativo: la zona en el sentido de que si el hogar beneficiario se pasaba a vivir de una zona rural a urbana, la probabilidad de deserción disminuía en 0.072 puntos porcentuales-pp., la calidad de la vivienda por cuanto si ésta empeoraba, la probabilidad aumentaba en 0.163 pp., la calidad de los servicios básicos, pues si ésta empeoraba, la probabilidad de abandono aumentaba en 0.081 pp., la edad, pues por cada año cumplido la probabilidad de deserción escolar aumentaba en 0.087 pp., y en cuanto a la condición de ser mujer, la probabilidad de abandono escolar era menor en 0.206 pp.

Otro estudio interesante para su análisis, y que contempla un período de evaluación mayor a cinco años posterior a la implementación del programa, es el realizado por Carriel y Mendoza (2021), quienes analizaron el impacto del Bono de Desarrollo Humano (BDH) sobre la inserción escolar en Ecuador, utilizando un diseño de regresión discontinua sobre los datos de la Encuesta Nacional de Empleo, Desempleo y Subempleo de diciembre de 2009. Por otra parte, en este estudio se calculó el índice de selección de beneficiarios SELBEN a partir de la encuesta de condiciones de vida (ECV) de 1999, con el fin de determinar la elegibilidad para ser beneficiario del programa. Adicionalmente, se realizaron estimaciones diferenciando los hogares de zonas urbanas y rurales, e incorporando la tasa de desempleo provincial como variable de control. De acuerdo con los resultados, no se encontró efecto del BDH sobre la inserción escolar, por lo cual se concluyó que este programa tiene poca incidencia, con lo cual los autores identifican el hecho de una posible ineficacia del BDH para lograr el objetivo de mejorar la tasa de inserción escolar, adicionalmente estos hallazgos coinciden con estudios de impacto anteriormente realizados al programa (Calvas, 2010; Ponce, 2008).

Aguilar y Siza (2010) evaluaron el impacto de Familias en Acción sobre una comuna de Medellín establecida por Acción Social como piloto para la ampliación del programa a centros urbanos, en la que se evaluaron los componentes de educación, salud, nutrición y consumo. Los datos fueron obtenidos mediante encuestas y se empleó una metodología que combinó el pareo por probabilidad de similitud y diferencias en diferencias. Al igual que para la evaluación realizada por DNP, se documentarán únicamente los efectos del programa en el componente de educación, para el cual se concluyó que Familias en Acción no tiene efectos sobre la población que asiste a educación básica primaria, partiendo del hecho de que en ausencia del programa la asistencia en estas edades es superior al 93%, convirtiendo en marginales los efectos del programa. Por el contrario, en la escuela secundaria, el programa incrementó la probabilidad de asistir a un establecimiento educativo entre 4.5 y 5.2 p.p.

- **Asistencia y Permanencia en el sistema educativo**

Como política educativa para acortar la deserción escolar, el gobierno colombiano establece los siguientes factores socio-pedagógicos en la deserción escolar (Montaña, 2015).

- La planeación de la permanencia escolar.
- La universalización de la gratuidad educativa.
- El fortalecimiento de modelos educativos flexibles.
- La mejora en la equidad en la asignación y distribución de los recursos financieros.
- El análisis, seguimientos y evaluación de la permanencia escolar de lo cual se ocupa esta investigación.

Como se indicó anteriormente, una de las dimensiones que componen el IPM para la medición de la pobreza en Colombia es la de condiciones educativas, medida a través de las variables de analfabetismo y bajo logro educativo. Esta última en particular, se identifica con el nivel educativo alcanzado y se asocia directamente con los años de estudio recibidos.

De acuerdo con la ley 115 de 1994, en Colombia existen tres niveles de educación formal: el preescolar, que comprende mínimo un grado obligatorio, la educación básica que se estructura en torno a un currículo común conformado por las áreas fundamentales del conocimiento y de la actividad humana y está compuesta por los ciclos de educación básica primaria con una duración de cinco grados y la educación básica secundaria que tiene una duración de cuatro grados, y finalmente, la educación media cuya finalidad es la comprensión de las ideas y los valores universales y la preparación para el ingreso del educando a la educación superior y a la vida laboral, y tiene una duración de dos grados.

Las condiciones educativas pueden verse afectadas por la deserción escolar, que implica el abandono de la trayectoria educativa de niños, niñas, adolescentes y jóvenes. La deserción escolar es el resultado de un proceso caracterizado por bajos aprendizajes, reprobación de grados, inasistencia reiterada y desmotivación por avanzar y aprender, factores que en su conjunto se denominan fracaso escolar.

En la identificación y comprensión del proceso que lleva a la deserción escolar, Roman C. (2013) caracteriza dos orientaciones teóricas, la primera ubica los factores y causas por fuera del sistema escolar pues considera que las causas del fracaso escolar son el resultado de una estructura social, económica y política que dificulta la asistencia regular y el buen desempeño en la escuela, y se debe a diferentes causas como la existencia de condiciones de pobreza y marginalidad, vinculación laboral temprana o grados de vulnerabilidad social, entre otros.

La segunda considera que los factores que llevan a la deserción escolar son inherentes a la escuela, e identifica algunas condiciones y situaciones que dificultan la permanencia de los estudiantes en la escuela, entre las más recurrentes, el tránsito entre el nivel primario al secundario. A su vez, ubica estas causas en dos perspectivas conceptuales, la primera relacionada con la resistencia a los códigos socializadores que entrega la escuela de una forma autoritaria, impidiendo que los estudiantes con distinto capital cultural puedan integrarse y asimilar la cultura escolar predominante a partir su propia cultura. La segunda

perspectiva se refiere a aspectos propios de los maestros y su interacción con los estudiantes, así como de las prácticas pedagógicas e institucionales al interior de los centros escolares. Adicionalmente se contempla otro factor relacionado con la actitud y las expectativas del educador sobre sus alumnos, que, entre más negativas, resultan tanto en un menor nivel de logro alcanzado como de rendimiento exhibido.

Metodología

El presente estudio se desarrolla bajo un enfoque cuantitativo de alcance correlacional, de tipo no experimental y con diseño transeccional. Los datos considerados hacen referencia a las características socioeconómicas de los hogares bogotanos para los años 2012 y 2020, de acuerdo con la ECV. A partir de la aplicación de métodos econométricos y Machine Learning se espera concluir frente al impacto del programa Familias en Acción en su etapa III.

La ECV contiene información que caracteriza las condiciones de vida incluyendo variables relacionadas con la vivienda (material de paredes, pisos y servicios públicos), con las personas (educación, salud, cuidado de los niños, fuerza de trabajo, gastos e ingresos) y con los hogares: (tenencia de bienes y percepción del jefe o del cónyuge sobre las condiciones de vida en el hogar). Los criterios de selección de las variables fueron: aporte de información socioeconómica necesaria para verificar el cumplimiento de los requisitos del programa Familias en Acción, y presencia de la variable en los dos años objeto de análisis.

Tabla 1. Variables seleccionadas

Unidad de observación y análisis	Número de variables	Variable
Vivienda	1	1 Estrato
Hogar	2	1 Los ingresos de su hogar alcanzan para cubrir los gastos mínimos. 2 Durante los últimos doce meses, ¿algún miembro de este hogar recibió ayudas o subsidios en dinero o en especie por concepto de: 1. Familias en acción?
Persona	16	1 Sexo 2 Años cumplidos 3 Parentesco con el jefe o la jefa de este hogar 4 El cónyuge vive en el hogar 5 Nivel de educación más alto alcanzado por el padre 5 Nivel de educación más alto alcanzado por la madre 7 Afiliación seguridad social en salud 3 Régimen de seguridad social en salud 9 Actualmente Estudia (Asiste Al Preescolar, Escuela, Colegio O Universidad) 10 Nivel educativo más alto alcanzado 11 Nivel educativo matriculado 12 Grado que cursa 13 El establecimiento donde estudia es oficial o privado

l4	Durante este año escolar, recibió subsidio en dinero o en especie para estudiar – Frecuencia
l5	En qué actividad ocupó la mayor parte del tiempo la semana pasada
l6	Cuánto ganó el mes pasado en su empleo

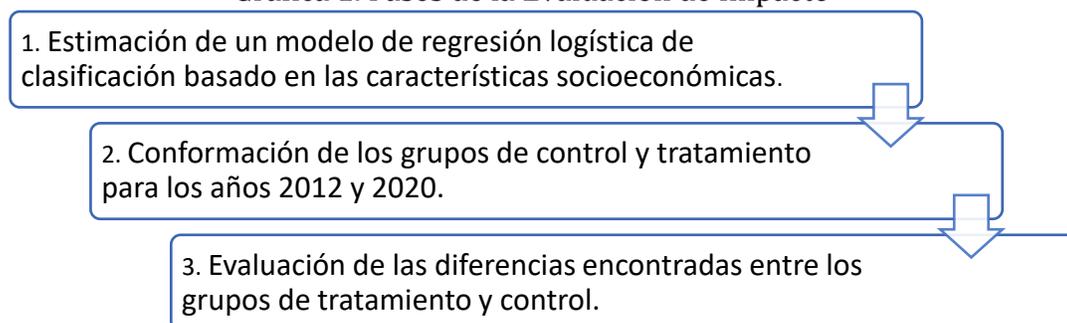
Fuente: Elaboración propia

Como resultado de un análisis descriptivo de cada una de las variables seleccionadas, se tomó la decisión de desistir de aquellas que presentaron un porcentaje superior al 20% de datos faltantes. Es el caso de las variables 3.4, 3.5, 3.6, 3.10, 3.12, 3.14 y 3.16.

Así mismo, se estableció un conjunto de restricciones para evitar registros inconsistentes: encontrarse en edad escolar, estar estudiando actualmente e informar que durante la semana anterior estaban estudiando. Cabe mencionar la data final tiene persona como unidad de análisis.

El método de evaluación de impacto basado en la aplicación de técnicas econométricas y de Machine Learning se encuentra estructurado en 3 fases. La primera fase consiste en la realización de tareas que conllevan a la construcción de un modelo que clasifique correctamente los registros de la base de datos, de tal forma que los resultados de la clasificación sean empleados para la conformación de los grupos de tratamiento y control. Entre las tareas realizadas se encuentra la exploración de los datos, la definición de criterios de éxito del modelo y la estrategia de evaluación, procesamiento de los datos, optimización del modelo a partir de variaciones en los hiperparámetros y la evaluación la capacidad del modelo final.

Gráfica 1. Fases de la Evaluación de Impacto



Fuente: Elaboración Propia

En la exploración de los datos se identificó la necesidad de aplicar técnicas de balanceo de datos, teniendo en cuenta que la clase *beneficiario* tiene una participación menor a 5% en los años analizados. Se aplicaron técnicas las upsampling y ROSE, haciendo uso de librerías del software R. A partir de cada balanceo se propusieron modelos de regresión logística, cuya selección de variables se realizó aplicando la técnica de feature selection Step Wise. Todas las variables categóricas fueron transformadas en variables dummy y fueron procesadas por el algoritmo de machine learning de Modelos Lineales Generalizados. Teniendo en cuenta

este algoritmo no acepta valores perdidos, se tomó la decisión de imputar datos de variables que después de las restricciones continuaban con un bajo porcentaje de datos faltantes.

Como parte de la estrategia de evaluación, las observaciones fueron separadas en un conjunto de entrenamiento, un conjunto de validación y un conjunto de prueba. Los criterios de éxito del modelo se establecieron en términos de las métricas de *Precision*, *Recall* (*exhaustividad*), *F1* y *Balanced Accuracy*. El modelo seleccionado debe tener resultados por encima a 0.6 en cada una de ellas y registrar un área bajo la curva ROC (AUC) superior al 75%. De igual forma, se evalúa la inclusión de variables clave en la identificación de un beneficiario del programa.

La segunda fase tiene como objetivo establecer un grupo de control para aquel grupo de tratamiento que no fue seleccionado aleatoriamente. Igualmente se creó un vector de características X del grupo de tratamiento para buscar individuos con las mismas particularidades dentro del grupo de individuos que no fueron tratados y de esta forma generar un grupo de control con el que se pueda comparar el resultado de los tratados. Este procedimiento requiere de dos supuestos: que exista traslape entre las características de los grupos y que el valor potencial del resultado condicionado en las X sea independiente de si se le asigna o no el tratamiento al individuo. Lo que permite reducir el sesgo de control entre el grupo control y de tratamiento, que complementa el modelo de diferencias en diferencias. La condición de independencia condicional y supuestos de tendencias paralelas que plantea Aedo (2005) establece que: para enfrentar este problema, se recurre a construir entre los individuos elegibles que no participaron en el programa, un grupo de control, similar al grupo de participantes en algunas variables de interés. Lo anterior requiere disponer de encuestas altamente representativas y comparables de participantes y no participantes, que para este caso se toman las encuestas de Calidad de Vida (ECV), de los años 2012 para establecer la línea de base de la medición del impacto y la de 2020 para determinar el impacto obtenido de la tercera Fase del programa de Familias en Acción, es importante aclarar que las muestras de estas encuestas son representativas a nivel de la ciudad de Bogotá.

La tercera fase tiene como propósito llevar a cabo un análisis econométrico que utilizará un modelo de diferencias en diferencias emparejadas para abordar el impacto del programa de Familias en Acción sobre la demanda educativa, el procedimiento metodológico que se llevará a cabo permite probar el supuesto de tendencias paralelas para un período de inicio y de finalización del tratamiento de la tercera fase del programa. Lo cual conduce a estimar la probabilidad de participación en función de las características socioeconómicas relacionadas, con la participación según la focalización del SISBEN (Lee Benson, 2012).

Según Bernal y Peña (2011), la evaluación de impacto consiste en establecer la diferencia entre la variable de resultado de la persona participante en un programa y la variable resultante de este individuo sin el programa, lo que sería el efecto de tratamiento del programa. Sin embargo, en la mayoría de los programas la elección de participar en un programa no se hace de forma aleatoria, sino que las personas eligen si participan o no, como es el caso de Familias en Acción, por lo tanto, las diferencias entre los participantes y no participantes no son observables por esto es un gran reto de la evaluación de impacto encontrar metodologías que permitan obtener un estimador consistente e insesgado β_1 , en presencia del sesgo de selección.

El modelo de diferencias en diferencias con datos de corte transversal repetido es:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 I [t = 2] + \beta_3 (D_i * I [t = 2]) + u_{it} \quad D_i = 1 \quad (1)$$

Y_i : Variable de resultado

D_i : Tratamiento

Si la observación corresponde a un individuo del grupo de tratamiento o a un sustituto del grupo de tratamiento, si las observaciones de pretratamiento $D_i = 0$ lo contrario $D_i = 1$ si la condición $I [t = 2]$ se cumple $I [t = 2] = 1$ si no se cumple $I [t = 2] = 0$ aplicación del tratamiento $I [t = 2] = 1$ lo contrario $I [t = 2] = 0$ es la interacción entre el indicador de tratamiento D_i y la variable binaria del periodo de seguimiento $I [t = 2]$

El efecto de tratamiento por diferencias en diferencias estaría dado por:

$$[E (Y \setminus D_i = 1, t = 2) - E (Y \setminus D_i = 1, t = 1)] - [E (Y \setminus D_i = 0, t = 2) - E (Y \setminus D_i = 0, t = 1)] \quad (2)$$

Teniendo en cuenta la expectativa condicional de x obtenemos

$$\begin{aligned} E (Y \setminus D_i = 1, t = 2) &= \beta_0 + \beta_1 + \beta_2 + \beta_3 \\ E (Y \setminus D_i = 0, t = 2) &= \beta_0 + \beta_2 \\ E (Y \setminus D_i = 1, t = 1) &= \beta_0 + \beta_1 \\ E (Y \setminus D_i = 0, t = 1) &= \beta_0 \end{aligned} \quad (3)$$

Entonces el efecto del tratamiento por diferencias estaría dado por:

$$\begin{aligned} &[\beta_0 + \beta_1 + \beta_2 + \beta_3] - [\beta_0 + \beta_1] - [(\beta_0 + \beta_2) - (\beta_0)] \\ &= (\beta_2 + \beta_3) - (\beta_2) = \beta_3 \end{aligned} \quad (4)$$

Es decir, el coeficiente asociado a la interacción $D_i * I [t = 2]$ entre la variable binaria de tratamiento y variable binaria de seguimiento, β_3 corresponde al efecto del programa por el método de diferencias en diferencias.

D_i {1 si el individuo pertenece al área geográfica en que se aplica el programa y cumple condiciones de elegibilidad. 0 si el individuo no pertenece al área geográfica en que se aplica el programa y cumple condiciones de elegibilidad}

t = {0 si la observación corresponde al período anterior a la intervención. 1 si la observación corresponde al período posterior de la intervención}

Resultados

Considerando que la investigación se encuentra en curso, los resultados hacen referencia a los avances generados en las dos primeras fases y a los resultados esperados una vez sea culminada.

Fase I

Las observaciones que cumplían con las restricciones fueron sometidas a un proceso de exploración que permitió generar insights. La mayoría de los registros se concentran en los

estratos 1, 2 y 3; pertenecen a hogares que manifiestan que sus ingresos no alcanzan o cubren justo los gastos mínimos; corresponden a hijos del hogar; se encuentran afiliados a seguridad social en salud (principalmente al régimen contributivo) y se encuentran cursando algún grado del nivel de básica secundaria y media. Al analizar la participación de los beneficiarios para casa año, se evidencian datos con un notable desbalance:

Cuadro 1. Observaciones ECV 2012-2020

Año	Datos	Número total de observaciones	Beneficiario		No beneficiario	
			Número de observaciones	Porcentaje	Número de observaciones	Porcentaje
2012	Datos ECV	5,917	259	4.40%	5,658	95.60%
	Datos cumplen criterios	625	45	7.20%	580	92.80%
2020	Datos ECV	8,438	266	3.20%	8,172	96.80%
	Datos cumplen criterios	757	46	6.08%	711	93.92%

Fuente: Elaboración Propia

Se entrenaron modelos sobre los 3 data sets balanceados con las técnicas de upsampling y ROSE para cada año.

Cuadro 2. Comparativa Técnicas de balanceo ECV 2012-2020

Año	Técnica de balanceo	Beneficiario		No beneficiario		Total
		Número de observaciones	Porcentaje	Número de observaciones	Porcentaje	
2012	Upsampling	134	25.05%	401	74.95%	535
	ROSE	119	23.80%	381	76.20%	500
2020	Upsampling	237	25.00%	711	75.00%	948
	ROSE	192	25.36%	565	74.64%	757

Fuente: Elaboración Propia

A continuación, se encuentran los modelos con mejores resultados en términos de los criterios de éxito:

Cuadro 3. Resultados de los modelos

MODELO	PRECISION	RECALL	F1	BALANCED ACCURACY	CURVA ROC
Upsampling (773=580+193) + step wise	0.6222222	0.7179487	0.6666667	0.7856985	0.8872679
Upsampling (773=580+193) + variables significativas SW	0.8666667	0.3333333	0.4814814	0.65804598	0.8740053
ROSE (625=466+159) + step wise	0.6000000	0.8437500	0.7012987	0.8261303	0.8708444

Fuente: Elaboración Propia

El mejor modelo para el año 2012 se obtuvo sobre el data set upsampling aplicando la técnica de future selection Step Wise. Este modelo tiene *Precision*, *Recall*, *F1* y *Balanced Accuracy* superior al 60% y su tasa de falsos positivos con un Recall de 71.8% es de 14.65%. Asimismo, el área bajo la curva ROC (AUC) es la más alta de los 3 modelos. Este modelo incluye las variables dummy establecimiento oficial, estrato 1, hijo/hijastro, régimen de seguridad social en salud subsidiado, nivel educativo secundaria y media, afiliado a seguridad social en salud y estrato 2.

La precisión indica que el modelo tiene la capacidad de predecir el 62.22% de las personas que realmente cumplen con las características para ser beneficiarias del programa respecto al total de personas que se clasificaron como beneficiarias. Frente a la exhaustividad, el modelo es capaz de identificar el 71.8% de personas que cumplen con las características para ser beneficiarias del programa frente al total de personas beneficiarias existentes. Esto lo hace, asumiendo una tasa de falsos positivos de 14.65%. El valor F1 combina los resultados de las dos métricas anteriores a través de una media armónica, es decir, muestra el rendimiento combinado de precisión y exhaustividad. De acuerdo con el balanced accuracy, de todas las predicciones del modelo el 78.56% serán correctas. El valor del área bajo la curva ROC puede interpretarse como la probabilidad de que, ante dos individuos, uno beneficiario y otro no, la prueba clasifique a los dos correctamente; en este caso, esa probabilidad sería del 88.72%.

Fase II

A partir de los resultados de la fase I se emplea el modelo upsampling step wise para clasificar las observaciones que registran no ser beneficiarios del programa en cada año, con el propósito de identificar cuál de ellos cumple con las características para serlo. Las observaciones clasificadas como positivas por el modelo conforman el grupo de control. Por su parte, el grupo de tratamiento es conformado por las observaciones que registran ser beneficiarios del programa en cada año. En este paso se aplican técnicas de muestreo para que ambos grupos contengan el mismo número de observaciones. Esto permitió identificar cuáles registros cumplen con las características para ser beneficiarios del programa y de esta forma configurar el grupo de control.

Cuadro 4. Grupos Tratamiento-Control

Grupo	Número de observaciones	
	2012	2020
Tratamiento	193	237
Control	193	237

Fuente: Elaboración Propia

Los grupos de tratamiento y control obtenidos para el año 2012 comparten las siguientes características: se encuentran en estratos 1 y 2 principalmente, son hijo o hijastro del(a) jefe(a) del hogar, se encuentran afiliados al Sistema General de Seguridad Social en Salud, principalmente en el Régimen Subsidiado y se encuentran matriculados en el nivel de Básica primaria, en la mayoría de los casos en establecimiento oficial.

Cuadro 5. Características Grupos Tratamiento – Control (Año 2012)

Variable	Clase	Grupo de Tratamiento	Grupo de Control
Estrato	Estrato 1	99	70
	Estrato 2	90	123
	Estrato 3	4	0
Ingresos del hogar	No alcanza para cubrir los gastos mínimos	68	59
	Solo alcanza para cubrir los gastos mínimos	124	128
Sexo	Hombre	96	116
	Mujer	97	77
Edad	Rango de edad	12-17	12-18
Parentesco con el jefe/jefa del hogar	Hijo/hijastro	180	190
	Otro	13	3
Afiliación seguridad social en salud	Si	191	193
	No	2	0
Régimen de seguridad social en salud	Contributivo (EPS)	45	35
	Subsidiado (EPS-S)	148	158
Nivel Educativo Matriculado	Básica primaria (1o a 5o)	193	193
El establecimiento donde estudia es	Oficial	186	188
	Privado	7	5

Fuente: Elaboración Propia

En cuanto a los grupos de tratamiento y control obtenidos para 2020, se encontró que las características comunes son similares a las que comparten los grupos definidos para 2012, es decir pertenecen a estratos 1 y 2 principalmente, son hijos o hijastros del jefe(a) del hogar, la totalidad de individuos se encuentran afiliados al Sistema General de Seguridad Social en Salud, en su mayoría en el Régimen Subsidiado, en este caso se encontraron individuos matriculados en niveles de preescolar y técnico, pero la mayoría se encuentran matriculados en el nivel de Básica Primaria, en establecimiento oficial.

Cuadro 6. Características Grupos Tratamiento – Control (Año 2020)

Variable	Clase	Grupo de Tratamiento	Grupo de Control
Estrato	Estrato 1	135	142
	Estrato 2	68	64
	Estrato 3	34	31
Ingresos del hogar	No alcanza para cubrir los gastos mínimos	140	156
	hogar Solo alcanza para cubrir los gastos mínimos	97	81
Sexo	Hombre	110	82
	Mujer	127	155
Edad	Rango de edad	12-18	12-18

Parentesco con el jefe/jefa del hogar	Hijo/hijastro	217	221
	Otro	20	16
Afiliación seguridad social en salud	Si	237	237
	No	0	0
Régimen de seguridad social en salud	Contributivo (EPS)	96	88
	Subsidiado (EPS-S)	141	148
Nivel Educativo Matriculado	Preescolar	17	21
	Básica primaria (1.o a 5.o)	204	198
	Tecnológico	16	18
El establecimiento donde estudia es oficial o privado	Oficial	234	237
	Privado	3	0

Fuente: Elaboración Propia

Referencias

- Aedo, Cristian.** (2015). Evaluación de Impacto. Santiago de Chile. CEPAL-Naciones Unidas.
- Aguilar W., Siza O. y Rubio, M.** (2010). "Familias en Acción: Evaluación de impacto de un programa piloto en Medellín, Colombia". Banco Interamericano de Desarrollo, División de Protección Social y Salud.
- Ayola, V.** (2018). "Condicionar, Focalizar e Invertir en Capital Humano: El Programa Familias en Acción o la Receta Internacional de Lucha contra la pobreza para Colombia". Panorama Económico, enero-marzo, 2018 No. 26(1), pp. 37-58.
- Banco Mundial.** (2022). Poverty and Shared Prosperity 2022. Correcting Course. Washington D.C.: Banco Mundial.
- Benson, A.** (2012) "Efectos de Familias en Acción sobre la calidad de la oferta educativa". Revista Desarrollo y Sociedad. pp. 51-91
- Bernal, Raquel y Peña, Ximena.** (2011). Guía práctica para la evaluación de impacto. Colombia. Ediciones Uniandes
- Calvas, G.** (2010). Evaluación de impacto del bono de desarrollo humano en la educación. Maestría en Ciencias Económicas con mención en Economía del Desarrollo; FLACSO sede Ecuador. Quito. 97 p.
- Carriel, V. y Mendoza, C.** (2021). "La inserción escolar en el Ecuador y el papel del Bono de Desarrollo Humano". Revista Economía y Política. 14/01/2021. No. 33.
- Cecchini, S. y Madarriaga A.** (2011). Programas de Transferencias Monetarias Condicionadas. Balance de la Experiencia Reciente en América Latina y el Caribe. Santiago de Chile: Naciones Unidas-CEPAL
- Departamento Nacional de Planeación** (2021). "Guía práctica para conocer el principal instrumento de focalización en el país". Bogotá D.C. Disponible en: <https://www.sisben.gov.co/Paginas/documentos-sisben.aspx>. Fecha de acceso: 5 de octubre de 2022.

- Departamento para la Prosperidad Social.** Rediseño del programa Familias en Acción, Documento operativo técnico DOT No 1 2012-2013
- Departamento para la Prosperidad Social.** Manual Operativo Programa: MÁS FAMILIAS EN ACCION. Bogotá, enero de 2017
- Departamento para la Prosperidad Social.** (2021). Documento Técnico Programa Familias en Acción- TRANSICION FASE III A FASE IV. Bogotá.
- Dallorso, N.** (2023). “La teoría del capital humano en la visión del Banco Mundial sobre las Transferencias Monetarias Condicionadas” enero-abril, 2013, Vol. 31, No. 91, pp. 113-139
- Llano, J.** (2014) “Familias en Acción: La historia a la luz de sus impactos” Coyuntura Económica. Vol. 44 No. 1. 01/06/2014 pp. 77-120
- Mata, C. y Hernández, K.** (2015). “Evaluación de impacto de la implementación de transferencias Monetarias Condicionadas para educación secundaria en Costa Rica (Avancemos)”. Revista de Ciencias Económicas, 33(1), pp. 9-35.
- Medellín, Nadin y Sánchez, Fernando** (2015). “¿Cómo funciona Más Familias en Acción? Mejores prácticas en la implementación de Programas de Transferencias Monetarias Condicionadas en América Latina y el Caribe”. Banco Interamericano de Desarrollo.
- Montaña, D.** (2015). Análisis de los factores Socio-Pedagógicos ante la Deserción Escolar. Bogotá. Disponible en: <https://repository.unimilitar.edu.co/bitstream/handle/10654/7698/VargasAyalaMartaJeannette2016.pdf?sequence=1>. Fecha de acceso: 23 de 10 de 2022
- Muñoz, J.** (2016). Efectos sobre el abandono educativo en el hogar de personas beneficiarias. Revista Electrónica Educare, 01/01/2016 No. 20(1), pp. 1-22.
- Ponce, Juan** (2008). Educational Policy and Performance: Evaluating the impact of targeted education programs in Ecuador. Phd. Dissertation. Shaker Publishing.
- República de Colombia.** (1994). Ley 115 de 1994. Diario Oficial No. 41.320, de 8 de febrero de 1994.
- Román C., M.** (2013). “Factores Asociados al Abandono y la Deserción Escolar en América Latina: Una Mirada en Conjunto”. Revista Iberoamericana sobre Calidad, Eficiencia y Cambio en Educación, No. 11(2), pp. 33-59.
- UNESCO** (2007). Educación de Calidad para Todos, II Reunión Intergubernamental del Proyecto Regional de Educación para América Latina y el Caribe (29 y 30 de marzo de 2007; Buenos Aires, Argentina).