

DOI: <http://dx.doi.org/10.15517/rce.v33i1.19964>

EVALUACIÓN DE IMPACTO DE LA IMPLEMENTACIÓN DE TRANSFERENCIAS MONETARIAS CONDICIONADAS PARA EDUCACIÓN SECUNDARIA EN COSTA RICA (AVANCEMOS)

Catherine Mata¹
Karla Hernández²

Recibido: 02/02/2015

Aprobado: 08/04/2015

RESUMEN

En este trabajo se evalúa el impacto del programa de transferencias monetarias condicionadas para secundaria, Avancemos, en Costa Rica. Específicamente, se mide el impacto sobre la deserción estudiantil para el primer año en el que se implementó el programa, mediante un panel creado con datos de las Encuestas de Hogares de Propósitos Múltiples para los años 2006 y 2007, elaboradas por el Instituto Nacional de Estadística y Censos. Por medio de herramientas econométricas y metodologías de impacto cuasi-experimentales (*Propensity Score Matching* y Diferencia en Diferencias), se encuentra un impacto positivo asociado al programa, tanto sobre la deserción como sobre la reinserción. Específicamente, entre un 10% y 16% de los estudiantes que no desertaron del sistema educativo, lo hicieron única y exclusivamente debido a Avancemos, de otra forma hubiesen abandonado sus estudios. Por lo tanto, este es un programa con impactos positivos sobre sus objetivos planteados: evitar la deserción y procurar la reinserción.

PALABRAS CLAVE: POLÍTICA GUBERNAMENTAL, BECA DE ESTUDIOS, ENSEÑANZA SECUNDARIA, DESERCIÓN ESCOLAR, POLÍTICA EDUCACIONAL.

ABSTRACT

This paper evaluates the impact of Avancemos, a conditional cash transfer program in Costa Rica. Specifically, this paper measures the impact on student desertion for the first year of the program using a panel created with the Household Surveys for Multiple Purposes for the years 2006 and 2007, elaborated by the National Institute of Statistics and Census. Using econometric tools and quasi-experimental methodologies such as Propensity Score Matching and difference-in-differences, we find a positive impact associated to the program for desertion and reinsertion. Specifically, for between 10 and

1 Universidad de Costa Rica. Instituto de Investigaciones en Ciencias Económicas, Código Postal 11501-2060 - Costa Rica; catherine.mata@ucr.ac.cr

2 Barcelona Graduate School of Economics. Estudiante de M.Sc. Economics of Public Policy, Código Postal 08005 Barcelona, España; karla.hernandez@barcelonagse.eu

16 percent of the students who did not leave high school, it was only due to Avancemos, meaning that without the program they would have abandoned their studies. This is why we can conclude that Avancemos had a positive impact according to its planned objectives of preventing dropouts and ensuring their reinsertion.

KEYWORDS: STATISTICAL ANALYSIS, POLITICAL PARTICIPATION, DEPUTIES, PRESIDENT, POLITICS, K-MEANS.

I. INTRODUCCIÓN

La deserción escolar en Costa Rica se ha caracterizado por ser más problemática a nivel de secundaria que en primaria. Durante los últimos 14 años, la tasa de deserción en secundaria ha pasado de 11,9 % en el año 2000, a 10,7 %, en 2012; mientras que en primaria pasó de 4,1 % a 2,5% en el mismo período. Esto explica por qué, a diferencia de otros países de Latinoamérica, en Costa Rica las transferencias monetarias condicionadas (TMC) se dirigen a la secundaria y no a la primaria³.

Avancemos es un programa social del Gobierno de Costa Rica ejecutado por el Instituto Mixto de Ayuda Social (IMAS), cuyo objetivo es promover la permanencia y reinsertión en el sistema educativo formal de adolescentes y jóvenes pertenecientes a familias que tienen dificultades para mantener a sus hijos en el sistema educativo por causas económicas. El programa consiste en una TMC directa a la madre o el encargado del estudiante; a cambio la familia debe firmar un contrato donde se compromete a cumplir con ciertas condicionalidades previamente definidas por el IMAS.

Este programa, desde sus inicios, ha sido foco de la academia y entes gubernamentales que buscan determinar la eficiencia y eficacia de dichas TMC, tanto en sus procesos como en su objetivo final: disminuir la deserción en educación secundaria. La Contraloría General de la República (CGR), como organismo encargado de velar por los recursos públicos que utilizan las instituciones del Estado Costarricense, ha señalado la necesidad de realizar evaluaciones de impacto del programa Avancemos: “Al Ministro de Educación Pública se le solicitó evaluar el Programa Avancemos, en coordinación con el IMAS, de manera que se conozca el efecto del programa en la permanencia o deserción de los estudiantes beneficiados con la TMC” (CGR, 2012, p. iii).

Las TMC no son una política social nueva en Costa Rica, antes de Avancemos, existió a cargo del IMAS un programa llamado Superémonos (ver Duryea & Morrison, 2004), cuya transferencia monetaria a los hogares pobres iba condicionada a la matrícula de los hijos (as) en la escuela sin embargo, era un programa a muy pequeña escala, comparado con Avancemos. Este programa fue cerrado y en el 2006 aparece el plan piloto de Avancemos, como transferencia monetaria condicionada a la permanencia de los jóvenes de los hogares pobres beneficiarios en el colegio o institución de educación similar.

Desde entonces, las becas de secundaria que brindaba el Fondo Nacional de Becas (FONABE), se convierten en transferencias monetarias condicionadas, ejecutadas al inicio por FONABE y el IMAS, y –en la actualidad–, exclusivamente por el IMAS. Estas transferencias nacen bajo la premisa de ser una política de superación de pobreza, basada en varios ejes de colaboración de distintos actores, entre ellos el Ministerio de Educación Pública (MEP) y la Caja Costarricense del Seguro Social (CCSS). Sin embargo, con el pasar de los años, el programa fue dejando de lado algunos de sus componentes iniciales y adquiriendo nuevos matices. No obstante, la condición más importante de superación de la pobreza, la inversión en capital humano, sigue siendo el eje central de la política.

3 Sin embargo, en primaria se cuenta con las becas del Fondo Nacional de Becas (FONABE). Estas constituyen una ayuda económica al estudiante para sufragar sus costos educativos, para lo cual los estudiantes deben mantenerse asistiendo al centro educativo. Además, está dirigido a aquellos estudiantes en condición de pobreza, medida a través de la línea de pobreza establecida por el Instituto Nacional de Estadística y Censos (INEC).

Como ha sido señalado por la CGR (CGR, 2008), el programa no fue evaluado antes de tomar las decisiones de los cambios en sus parámetros. Específicamente, no se realizó una estimación temprana de impacto de su puesta en marcha. Por esta razón, resulta esencial evaluar si el programa, bajo su diseño original, cumplió con su objetivo inmediato: evitar la deserción en secundaria de jóvenes que se encontraban en situación de pobreza o vulnerabilidad social. De esta forma, se puede analizar si los cambios eran necesarios de realizar y, mediante una posterior evaluación de impacto, para un periodo más reciente, evaluar si los cambios en sus reglas provocaron alguna mejora en sus resultados e impacto.

Para llevar a cabo la evaluación de impacto, y aislar el efecto directo de Avancemos sobre la deserción en secundaria, se utilizaron principalmente dos metodologías cuasi-experimentales: el propensity-score matching (PSM) y diferencia en diferencias. De esta forma, el presente trabajo pretende llenar el vacío existente con respecto a las evaluaciones de impacto de la implementación del programa Avancemos.

Los datos para el periodo 2006-2007 muestran que Avancemos tuvo un impacto positivo, tanto sobre la deserción como sobre la reinserción (aunque en el segundo caso, la muestra disponible es más pequeña que la de deserción). Esto significa que algunos estudiantes permanecieron estudiando gracias a la ayuda económica que recibieron; por otro lado, el programa también influyó en el regreso de los jóvenes que ya habían desertado.

Algunas de estas personas se mantuvieron en el sistema educativo como consecuencia de otra serie de factores, aunados al de contar con la TMC; pero, entre un 10% y 16% de ellos, lo hicieron única y exclusivamente debido a Avancemos. Si bien, quienes se reintegraron a la educación secundaria, lo hicieron por diversos motivos, un porcentaje bastante alto (77% o más) se reincorporaron debido a la transferencia que se les entregó. A pesar de que los datos derivados de estas encuestas constituyen muestras pequeñas respecto a la población beneficiaria del programa, se considera una muestra representativa, al no mostrar sesgos respecto a la distribución de las características de los beneficiarios reportados por el IMAS y, por consiguiente, se posiciona como un buen primer acercamiento del impacto del programa, el cual no había sido abordado en el país.

Por lo tanto, el programa Avancemos es un programa con impactos positivos sobre sus objetivos planteados, evitar la deserción y procurar la reinserción. A pesar de que la muestra en el caso de reinserción es mucho menor que en el caso de deserción, no se puede dejar de lado que el impacto positivo de la transferencia es mucho mayor en reinserción.

Antecedentes

Las transferencias monetarias condicionadas (TMC), según Fiszbein y Schady (2009), consisten en otorgar efectivo a hogares pobres con el requisito de que estos realicen inversiones –previamente especificadas– en el capital humano de sus integrantes. Lo usual es que las condiciones estén relacionadas con temas de salud y educación, que son verificadas con cierta periodicidad. Estas iniciativas de protección social han ganado gran aceptación en los países en desarrollo; ya que han logrado “ayudar a los hogares a salir del círculo vicioso que transmite la pobreza de una generación a la siguiente; de promover la salud, la nutrición y la escolarización de los niños, y de ayudar a los países a cumplir con los objetivos de desarrollo del milenio” (Fiszbein y Schady, 2009, p. 1).

Los programas de TMC nacieron en la década de los noventa, en países latinoamericanos esencialmente. De allí su propagación se hace de manera rápida en la misma región, así como también se difunde en países no latinoamericanos. Para Rawlings (2004), los programas de TMC, originalmente desarrollados en América Latina y el Caribe, buscan proveer asistencia social y mejorar el capital humano de los niños. De esta forma, Rawlings (2004) establece tres grupos de TMC: las que

abarcan la salud y nutrición, las que incluyen la educación solamente y aquellas que abarcan todos los componentes mencionados.

Los programas de TMC Progresar en México y Bolsa Escola en Brasil fueron pioneros en este tipo de ayuda social. Progresar inició con unos 300 mil hogares en 1997, pero a finales de la década pasada ya alcanzaba los 5 millones de beneficiarios. Así mismo, Brasil con su programa Bolsa Escola inició a una escala municipal en 1997 y en la actualidad es un programa federal (Bolsa Familia) que atiende a unos 46 millones de personas.

Cada país ha ajustado el programa a su realidad y a la tipología de sus grupos más vulnerables, sustentados sobre la base de romper con el ciclo de la pobreza a través de los ejes de educación y salud. La mayoría de los programas de TMC se basan en atacar problemas relacionados con educación y salud. En términos educativos, mientras que algunos países luchan por evitar la deserción en secundaria a través de las TMC, la mayoría todavía lucha por evitar la deserción en educación primaria. Además, algunos países dirigen las TMC a evitar el matrimonio temprano en las mujeres y a promover su educación; así como otros deciden invertir en educación temprana. En temas de salud, reducir tasa de mortalidad materno-infantil, acabar con la desnutrición infantil y la salud de las mujeres embarazadas son los principales objetivos que asumen estos programas.

Algunos de los países luchan actualmente contra la pobreza a través de varios programas de TMC. Otros han ido cerrando, reformando y renovando los programas para llegar a programas nuevos o fusionados que sean más depurados para lograr los fines propuestos. En todos los casos, para los que se encuentra información, la inversión en este tipo de programa no supera el 1% del PIB de cada país. Su magnitud dentro de cada país, en cuanto a recursos y personas beneficiarias, es muy variable. “En algunos países, incluyendo Brasil, Ecuador y México, las TMC se han convertido en el mayor programa de asistencia social, cubriendo millones de hogares” (Fiszbein y Schady, 2009, p. 47).

En el año 2006, Costa Rica decidió emular este tipo de políticas. Para ese año, mediante el Decreto Ejecutivo No. 33154 (2006), se crea el programa de transferencia monetaria condicionada Avancemos. Este se llevó a cabo considerando que, en el país, también existe el reto de la universalización de una educación secundaria de calidad. Así mismo, la evidencia empírica disponible señala una relación inversa entre educación y pobreza, de modo tal que la acumulación de años de escolaridad reduce la probabilidad de estar bajo la línea de pobreza. Una temprana inserción en el mercado laboral conspira contra el derecho a la educación de un contingente importante de adolescentes, especialmente en las regiones periféricas del país.

Avancemos es un programa social del Gobierno, ejecutado por el Instituto Mixto de Ayuda Social (IMAS). Su objetivo es promover la permanencia y reinserción en el sistema educativo formal de adolescentes y jóvenes pertenecientes a familias que tienen dificultades para mantener a sus hijos en el sistema educativo por causas económicas. Dada la naturaleza condicional del programa, la familia firma un contrato donde se compromete a cumplir con las condiciones definidas por el IMAS. El estudiante debe estar matriculado en cualquiera de las modalidades educativas de secundaria del MEP.

Entre los principales requisitos para tener acceso a este tipo de ayuda, se encuentran contar con la Ficha de Información Social (FIS) y calificar en condiciones de pobreza según los parámetros institucionales⁴, así como información de los padres y sus ingresos y firmar un contrato con los compromisos que adquiere la familia. En este último punto, la familia se compromete a apoyar y

4 El IMAS ha desarrollado una clasificación de las familias por nivel de prioridad, mediante tres categorías o rangos de puntaje. El grupo 1 es el de los más pobres; el 2 es de pobreza moderada y el grupo 3 es gente con un nivel menor de pobreza. (Viquez, 2005). El grupo 3 es catalogado como un grupo vulnerable por el IMAS. De esta forma, la institución atiende a personas de los grupos pobres y vulnerables. A diferencia de la medida de línea de pobreza o pobreza por ingresos, el IMAS contempla además del ingreso otras características del hogar como el estado de la vivienda, el número de miembros y sus condiciones laborales, de educación, de salud, entre otras.

asegurar la asistencia puntual y permanente de quienes estudian en el centro educativo, para que aprueben el curso lectivo.

La transferencia de Avancemos se paga mensualmente y se comienza a entregar una vez la familia ha sido aceptada por la entidad ejecutora para recibir el beneficio y se mantiene durante el resto del año, siempre y cuando el estudiante se mantenga en el sistema educativo. Desde 2006 y hasta 2014⁵, los montos eran escalonados (aumentando de conformidad con el año cursado) debido a que se consideraba un incentivo para el éxito o aprovechamiento educativo. Los montos vigentes en el periodo de estudio se muestran en el cuadro 1.

CUADRO 1
AVANCEMOS: MONTO MENSUAL DE TRANSFERENCIA SEGÚN AÑO CURSADO
(EN COLONES)

Nivel Educativo	Monto de Transferencia (mensual)
7° Año	15,000
8° Año	20,000
9° Año	25,000
10° Año	35,000
11° Año	45,000
12° Año	50,000

Fuente: IMAS. Ayuda social: Avancemos. Tomado de: http://www.imas.go.cr/ayuda_social/avancemos.html

En un inicio, Avancemos exigía dos requisitos primordiales: primero –como es tradicional en estos programas–, la principal condición era la asistencia al colegio y el desempeño académico. En segundo lugar, existía un requerimiento en el tema de salud, el cual consistía en recibir al menos una valoración integral de salud por parte de la CCSS.

Además, el programa incluía tres componentes: el primero era la transferencia monetaria; el segundo era el incentivo para fomentar el ahorro, que pretendía que a los beneficiarios del programa se les creara un fondo de ahorro para poder continuar con los estudios universitarios, estudiar idiomas o iniciar su propio negocio y; el tercero era un apoyo a la oferta educativa y formativa.

II. METODOLOGÍA

La medición del impacto que el programa tuvo durante su implementación se realiza con base en una combinación de métodos que permite estimar el contrafactual de la política, es decir, que habría pasado en este caso si el joven no hubiese sido parte del programa. La diferencia entre lo acontecido efectivamente y la estimación de este contrafactual define el impacto del programa. Dado que los individuos beneficiados por esta política y quienes no han recibido la ayuda pueden ser distintos, se busca eliminar el efecto de esas diferencias sobre la estimación para que no sea sesgada, logrando así extraer el efecto exclusivo de la transferencia monetaria condicionada sobre la deserción de los estudiantes de secundaria.

5 En el año 2015 estos montos son modificados, estableciendo dos únicos montos, 22.500 colones para el tercer ciclo (séptimo, octavo y noveno) y 35,000 colones para cuarto ciclo (décimo, undécimo y duodécimo).

El propensity score-matching (PSM) consiste en buscar observaciones no tratadas que sean similares a las tratadas, con base en un conjunto de variables que determinen la probabilidad de estar en el programa. Para esto, se utiliza el propensity-score que es estimado a través de la probabilidad de ser tratado. Esta probabilidad se calcula con base en características de la Ficha de Información Social del IMAS (identificadas en las encuestas de hogares del INEC); así como con la población pobre definida por el INEC en dichas encuestas, mediante el método de línea de pobreza. Una vez identificados el grupo de control y de tratamiento, se procede a estimar el efecto de recibir la TMC sobre la deserción y reinserción, utilizando el método de diferencia en diferencias.

Para desarrollar estas metodologías se utilizaron los datos de la Encuesta de Hogares de Propósitos Múltiples (EHPM) que realizó el INEC para los años 2006 y 2007, y con ellas se elaboró un panel para este periodo de dos años. A continuación, se explica el método de matching que precede al PSM, y posteriormente, se ahonda en los dos métodos utilizados para este estudio: PSM y diferencia en diferencias.

Matching o emparejamiento

La idea principal detrás del método de *matching* o emparejamiento es establecer un grupo de control para aquel grupo de tratamiento que no fue seleccionado aleatoriamente. Es decir, forma parte de los métodos cuasi-experimentales de evaluación de impacto. Este método econométrico ha sido ampliamente desarrollado y discutido por diversidad de autores, por lo que en la actualidad se considera bastante estándar en su aplicación. Por esta razón, en el presente trabajo se explica con base en lo expuesto por Gertler et al (2011).

A grandes rasgos, la idea original del *matching* es establecer un vector de características X del grupo de tratamiento, y buscar individuos con las mismas características dentro del grupo de individuos que no fueron tratados, para formar un grupo de control con el cual comparar el resultado de los tratados. Este método requiere de dos supuestos principalmente: que haya 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 individuo⁶. Es debido a estos supuestos del método que el *matching* se suele combinar con algunos otros métodos de evaluación de impacto.

“La idea es encontrar, dentro de un gran grupo de no participantes, individuos que sean similares a los participantes en términos de sus características observadas no afectadas por el programa (estas pueden incluir características antes del programa⁷, por ejemplo, porque estas claramente no son afectadas por la subsecuente participación en el programa) [...] El problema es identificar de forma creíble grupos similares entre sí. La identificación es un problema porque incluso si los hogares son emparejados a lo largo de un vector X de diferentes características, rara vez uno puede encontrar dos hogares que son exactamente iguales en términos de muchas características. Debido a que muchas posibles características existen, una manera común de empatar los hogares es el PSM” (Khandker et al, 2010: p. 54).

6 Es decir, después de controlar por X , las observaciones que se trataron y las que no se trataron tienen la misma esperanza del resultado si se hubieran tratado o si no se hubieran tratado.

7 Debe recordarse que el matching debe realizarse usando las características de la línea basal. Además, el método de matching es tan bueno como lo sean las características empleadas para el pareo, por eso disponer de un gran número de características generales es crucial (Gertler et al, 2011).

Propensity-Score Matching (PSM)

El *propensity-score* (PS) es un valor numérico entre cero y uno, que resume la probabilidad de ser parte del programa, con base en todas las características observadas de los individuos de ambas muestras, tanto los tratados como los no tratados. De acuerdo a Gertler et al (2011), para que las estimaciones del impacto de un programa, a través de métodos observacionales, sean externamente válidas, todas las unidades tratadas deben poder empatarse exitosamente con una unidad no tratada. Sin embargo, es posible que no haya área de traslape entre ambas poblaciones en los extremos de la distribución de sus PS, por lo que las unidades con un PS muy cercano a 1 no pueden ser empatadas con ninguna de las unidades no tratadas. Intuitivamente, las unidades con alta probabilidad de estar en el programa son muy diferentes a las unidades no tratadas, por lo que no se puede encontrar un buen empate para ellas.

De acuerdo a Khandker et al (2010), en el método de PSM básicamente se construye un grupo estadísticamente comparable basado en un modelo de probabilidad de participar en el tratamiento T condicional a las características observables X, o el *propensity score*: $P(X)=Pr(T=1|X)$. Rosenbaum y Rubin (1983) muestran que, bajo ciertos supuestos, hacer el emparejamiento en P(X) es tan bueno como hacerlo en X. Los supuestos necesarios para la identificación del efecto del programa son: a) independencia condicional (es decir, los factores no observables no afectan la participación) y b) considerable traslape del PS entre las muestras de participantes y no participantes.

La independencia condicional establece que dado un set de variables observables X que no son afectadas por el tratamiento, los resultados potenciales Y son independientes de la asignación del tratamiento T. Es decir,

$$(Y_i^T, Y_i^C) \perp T_i | X_i$$

Este supuesto es también llamado “*unconfoundedness*” (Rosenbaum y Rubin, 1983) e implica que ser parte del programa depende enteramente de características observables. Para estimar el efecto del tratamiento sobre los tratados⁸ se necesita un supuesto suavizado:

$$(Y_i^C, Y_i^T) \perp T_i | X_i$$

Si se tiene una buena base de datos antes del programa, esto ayudará a cumplir con el supuesto de independencia condicional, permitiendo controlar las variables observables que estén afectando la participación en el programa (asumiendo que la selección no observada es limitada).

El supuesto de traslape⁹ asegura que las observaciones tratadas tienen observaciones comparables cercanas en la distribución del PS. Solo en esta área de traslape se puede inferir acerca de causalidad.

$$0 < P(T_i=1 | X_i) < 1$$

Algunas pruebas de balance pueden llevarse a cabo para comprobar la riqueza de tal área de traslape, las cuales se explicarán más adelante. Para estimar el efecto del tratamiento sobre los tratados, es suficiente con un supuesto suavizado:

$$P(T_i=1 | X_i) < 1$$

8 ATT Average treatment on the treated o TOT Treatment over treated.

9 Common support overlap condition.

Algunas unidades no tratadas deberán ser excluidas de la muestra para asegurar comparabilidad. Sin embargo, en ciertas ocasiones, también deberá prescindirse de las unidades tratadas de la muestra si no existen unidades similares contra las cuales compararlas (Ravallion, 2008). Esta última situación es más problemática porque puede crear un posible sesgo de muestra en el efecto del tratamiento.

Khandker et al (2010) exponen la aplicación del método de PSM a través de tres pasos esenciales: la estimación de un modelo de participación en el programa, la definición de la región de traslape y las pruebas de balance, y finalmente el *matching* entre participantes y no participantes.

Paso 1. Estimar un modelo de participación en el programa

Primero se unen las muestras de los participantes ($T=1$) y los no participantes ($T=0$), para proceder a estimar la participación T con base en todas las variables X que se cree determinan la participación en el programa. Esta probabilidad T es estimada a través de un modelo probit o logit, modelos con respuesta cualitativa, específicamente, de probabilidad con respuesta binaria. En este trabajo se utiliza el modelo logit.

Después de haber estimado la ecuación de participación en el programa, se tienen los valores predichos de T , para todos los individuos. Este es el PS o probabilidad de participar en el programa, $\hat{P}(X|T=1) = \hat{P}(X)$. Debe tomarse en cuenta que la ecuación de participación no es un modelo determinante, por lo que los resultados de la estimación como los t -estadísticos y el R^2 ajustado no son muy informativos y pueden ser engañosos. Para esta etapa del PSM, la causalidad no es de gran interés como lo es la correlación de X con T .

De acuerdo a Heckman et al (1997), el sesgo en el programa de las estimaciones de PSM puede ser bajo, tomando en cuenta tres previsiones: primero, si es posible, utilizar la misma fuente de datos, la misma encuesta, tanto para la muestra de participantes como la muestra de no participantes; de esta forma se asegura que las variables X fueron medidas bajo los mismos parámetros y significan lo mismo en ambas muestras. Segundo, tener una muestra representativa, tanto en el grupo de los participantes como en el grupo de los no participantes, puede mejorar la precisión del PSM; inclusive, entre más grande sea la muestra de los no participantes más se facilitará el *matching*. Por último, si se utilizan encuestas diferentes, estas deben ser lo más parecidas posible y los individuos se deben enfrentar a las mismas condiciones, para ello se puede decidir utilizar participantes y no participantes de la misma zona geográfica (Ravallion, 2008).

En todo caso, incluir demasiadas variables en la ecuación de participación también debe evitarse, una sobre-especificación del modelo puede llevar a errores estándar mayores para el PS estimado $\hat{P}(X)$ y puede llevar a una predicción perfecta de participación para muchos hogares ($\hat{P}(X)=1$).

Paso 2. Definir la región de traslape y pruebas de balance

Al obtener un PS para cada una de las observaciones, tanto las participantes como las no participantes, se obtienen también sus correspondientes distribuciones del PS. Existirá un área en donde estas dos distribuciones se traslapen, es decir, un área común entre ambas. Las observaciones que no se encuentren dentro de esta área de traslape deberán ser retiradas de la muestra a utilizar. El sesgo de muestra o de selección puede ocurrir, si las observaciones no tratadas que son expulsadas de la muestra son sistemáticamente diferentes de las observaciones tratadas que se conservan en la muestra, en términos de las características observadas; estas diferencias deben ser monitoreadas de manera cuidadosa en aras de interpretar correctamente el efecto del tratamiento.

Algunas pruebas de balance se pueden ejecutar en cada uno de los cuantiles de la distribución del PS, para verificar si el promedio del PS es igual a la media de X. El balance implica que las distribuciones de ambos grupos, tratados y no tratados, son similares. Formalmente se necesita verificar que $\hat{P}(X|T=1) = \hat{P}(X|T=0)$.

Paso 3. Matching entre participantes y no participantes

Existen diferentes métodos para emparejar los individuos participantes con los no participantes del programa, con base en el PS estimado. Esto conlleva calcular un *peso* para cada emparejamiento entre tratado y no tratado. La escogencia de la técnica de *matching* puede afectar la estimación del impacto del programa de acuerdo a los pesos asignados. Algunas de las técnicas más usadas o conocidas, de acuerdo a Khandker et al (2010), son el *matching* con el vecino más cercano, *matching* dentro de un caliper o radio, *matching* por estratificación o intervalos y el *matching* de diferencias en diferencias.

Para efectos de este trabajo, se utiliza el *matching* con el vecino más cercano (Nearest-neighbor matching), en donde cada unidad tratada se empareja con la unidad no tratada que tenga el PS más cercano. Para ello, se establece una cantidad *n* de unidades de control que serán emparejadas con cada una de las tratadas. Además, se debe decidir si el *matching* se hace con o sin reemplazo, es decir, si cada unidad de control solo puede emparejarse una vez con la tratada o si puede emparejarse con varias tratadas.

Como todos los demás métodos de evaluación, el *matching* posee algunas bondades y otras limitaciones. Gertler et al (2011) establecen dos limitaciones principalmente. Primero, que se requieren grupos extensos de datos con grandes muestras de unidades y, aun cuando están disponibles, la falta de traslape entre el grupo tratado y el grupo de no tratados puede aparecer. Segundo, el *matching* únicamente puede ser llevado a cabo con base en características observables, por definición no se pueden incorporar características no observables en el cálculo del PS, por lo que es generalmente menos robusto que otros métodos de evaluación. Por ejemplo, los métodos de selección aleatoria no requieren supuestos no comprobables de la falta de variables no observables que explican tanto la participación en el programa como los resultados. Inclusive no necesitan grandes muestras o gran cantidad de características generales, como lo requiere el PSM.

Algunas otras limitaciones, consideradas por Khandker et al (2010), son que, en comparación con otros métodos tradicionales de regresión, la varianza estimada del efecto del tratamiento al utilizar PSM puede incluir la varianza atribuible a la derivación del PS, a la determinación del área de traslape, y (si se utilizó *matching* sin reemplazo) al orden en el cual las unidades tratadas fueron emparejadas. Esto puede ocasionar que los errores estándar estén estimados de forma incorrecta.

Sin embargo, Khandker et al (2010) argumentan que una solución propuesta por varios autores es el uso de *bootstrapping*, en donde repetidas muestras se obtienen de la muestra original y las propiedades de las estimaciones (como el error estándar y el sesgo) se vuelven a estimar con cada muestra. Cada estimación de una muestra de *bootstrap* incluye los primeros pasos de la estimación que derivó el PS, el área de traslape, y así sucesivamente. Debido a que los estimadores son asintóticamente lineales, el método de *bootstrapping* lleva a errores estándar e intervalos de confianza válidos.

Así mismo, consideran que el

PSM es un método semiparamétrico, que impone menos restricciones en la forma funcional del modelo del tratamiento, así como menos supuestos acerca de la distribución del término de error. Además, al sacar de la muestra aquellas observaciones no tratadas que no fueron

empatadas, se disminuye el sesgo en el efecto del tratamiento. Sin embargo, si el área de tras-lape no es suficientemente grande y las unidades no tratadas expulsadas de la muestra son sistemáticamente diferentes de las que se conservaron en la muestra, se induce a un sesgo de selección. (Khandker et al, 2010, p.64).

Finalmente, el método de *matching* puede producir estimaciones consistentes bajo condiciones débiles. Una práctica ventaja del PSM sobre el método de MCO es que se reduce el número de dimensiones en las cuales empatar los participantes con las unidades de control. No obstante,

cuando están disponibles los datos de la línea basal, el matching basado en las características generales de dicha línea basal puede ser muy útil combinado con otras técnicas, por ejemplo, diferencia en diferencias, que representa el invariante en el tiempo, la heterogeneidad no observada. Matching es incluso más útil cuando la regla de asignación del programa es conocida, en cuyo caso el matching se puede llevar a cabo sobre esa regla (Gertler et al, 2011, p.115).

Por esta razón, se considera que el PSM más la técnica de Diferencia en Diferencias, logra simular una situación en la cual el tratamiento se hubiese asignado de forma aleatoria. El método de Diferencia en Diferencias se explica a continuación.

- Diferencia en Diferencias

En el tema de evaluaciones de impacto existen métodos como asignación aleatoria y regresión discontinua que producen estimaciones del contrafactual para programas con reglas específicas de asignación que los evaluadores conocen y comprenden. No obstante, métodos como diferencia en diferencias y métodos de emparejamiento ofrecen a los evaluadores distintas herramientas que pueden ser aplicadas en situaciones en las cuales las reglas de asignación del programa no son tan claras. Ambos métodos son estadísticamente muy poderosos y, como ya se mencionó, se pueden utilizar juntos; sin embargo, ambos requieren supuestos más fuertes que otros métodos de evaluación.

Una evaluación de impacto, como indican Bernal y Peña (2011), pretende medir el impacto de un programa o una política social sobre un conjunto de variables de resultado en un conjunto de individuos. Estas variables de resultado son sobre las cuales se espera que el programa tenga un efecto en los individuos beneficiarios del programa evaluado.

La principal dificultad que presenta una evaluación de impacto es que, para construir el efecto de tratamiento, se necesita conocer la diferencia entre la variable resultado del individuo beneficiario –una vez que se ejecutó– el programa y la variable resultado que habría obtenido ese individuo en el caso hipotético de que no existiera el programa. El primer caso se conoce como una variable observada; ya que la información existe y los datos están a disposición del investigador. El segundo caso hipotético no es observable pues los datos no están a disposición y se conoce como el contrafactual.

El indicador de tratamiento es definido como D_i , en el caso de un tratamiento binario $D_i=1$ si el individuo i recibe el tratamiento y 0 en caso contrario. Las variables de resultados se definen como $Y_i(D_i)$ para cada individuo $i=1,\dots,N$, donde N es la población total. $Y_i(1)$ es la variable de resultado si el individuo i es tratado y $Y_i(0)$ si el individuo no es tratado.

El efecto de tratamiento (o impacto del programa) para un individuo i se puede escribir como:

$$\tau_i = Y_i(1) - Y_i(0)$$

Este impacto del programa se refiere a un momento dado en el tiempo, cuyo resultado observado se puede escribir como:

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0) = \begin{cases} Y_i(1) & \text{si } D_i = 1 \\ Y_i(0) & \text{si } D_i = 0 \end{cases}$$

Como se mencionó anteriormente, debido a que uno de los resultados de la ecuación no es observable para cada individuo i , no es posible observar el efecto individual del tratamiento τ_i . Es por ello que el análisis se debe concentrar en el impacto promedio del programa en la población o subconjuntos de la población.

En primera instancia, se puede observar el impacto promedio del programa en la población (ATE)¹⁰:

$$\tau_{ATE} = E[\tau_i] = E[Y_i(1) - Y_i(0)]$$

Como se mencionó anteriormente, debido a que uno de los resultados de la ecuación no es observable para cada individuo i , no es posible observar el efecto individual del tratamiento τ_i . Es por ello que el análisis se debe concentrar en el impacto promedio del programa en la población o subconjuntos de la población.

En primera instancia, se puede observar el impacto promedio del programa en la población (ATE)¹⁰:

$$\tau_{ATE} = E[\tau_i] = E[Y_i(1) - Y_i(0)]$$

Una representación simple de la variable de resultado con base en el modelo de regresión lineal está dada por:

$$Y_i = \beta_0 + \tau_i D_i + u_i$$

Donde $\tau_i = Y_i(1) - Y_i(0)$ y $Y_i(0) = E[Y_i(0)] + u_i = \beta_0 + u_i$

El efecto τ_{ATE} se interpreta como el cambio promedio en la variable de resultado cuando un individuo escogido al azar pasa aleatoriamente de ser participante a ser no-participante. Este parámetro es relevante en una evaluación de un programa universal. No obstante, en la mayoría de los casos, el alcance del programa no es universal, sino que está disponible para un subconjunto de la población, por lo que se debe utilizar un estimador que únicamente promedie el efecto sobre la población elegible o el impacto promedio del programa sobre los tratados (ATT)¹¹:

$$\tau_{ATT} = E(\tau_i | D_i = 1) = E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 1]$$

El efecto promedio del programa sobre los tratados es relevante para definir si un programa existente debe continuar, modificarse o eliminarse. En este caso, $E[Y_i(1) | D_i = 1]$ es el valor esperado de la variable de resultado en el grupo de tratamiento en presencia del tratamiento y $E[Y_i(0) | D_i = 1]$ es el resultado contrafactual o valor esperado de la variable de resultado en el grupo de trata-

10 Average treatment effect.

11 Average treatment on the treated.

miento en ausencia del tratamiento. Este último es el resultado hipotético que no se observa en la realidad y, por tanto, no queda registrado en los datos.

A su vez, puede calcularse el impacto promedio del programa sobre los no participantes, que consiste en la diferencia entre la media de la variable de resultado que habrían tenido los no participantes si hubieran participado en el programa y la media de la variable de resultado que efectivamente tuvieron los no participantes al no haber participado.

$$\tau_{ATU} = E(\tau_i | D_i = 0) = E[Y_i(1) | D_i = 0] - E[Y_i(0) | D_i = 0]$$

En este caso, el resultado contrafactual $E[Y_i(1) | D_i = 0]$ corresponde al promedio de la variable de resultado de los no participantes si hubieran participado en el programa, dado que no han sido tratados. La importancia del parámetro τ_{ATU} es cuando la evaluación tiene por objetivo investigar si el programa se debe extender o no a otros grupos de población.

Para ambos estimadores τ_{ATT} y τ_{ATU} el reto es elegir el contrafactual adecuado, ya que este es hipotético y no observable. Una opción es utilizar el promedio de la variable de resultado entre los individuos no participantes pero elegibles para participar en el programa; a este se le conoce como el grupo de control o de comparación. Es decir, se podría utilizar el resultado de los no-participantes, pero elegibles, como una aproximación del resultado que habrían tenido los participantes si el programa no hubiera existido. No obstante, esto puede generar un “sesgo de autoselección” ya que las variables de resultado del grupo de tratamiento y del grupo de control podrían ser diferentes, aún si la política no existiera.

Otro riesgo es que existan diferencias no observadas o no medidas entre los grupos. La comparación entre el promedio de la variable de resultado de los no participantes está dada por:

$$E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 0]$$

Adicionalmente se sabe que el ATT está dado por:

$$\tau_{ATT} = E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 1]$$

$$\tau_{ATT} + E[Y_i(1) | D_i = 1] = E[Y_i(0) | D_i = 1]$$

Restando $E[Y_i(0) | D_i = 0]$ a ambos lados de la ecuación se obtiene que:

$$\tau_{ATT} + E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 0] = E[Y_i(0) | D_i = 1] - E[Y_i(0) | D_i = 0]$$

Donde el lado derecho de la ecuación es la diferencia entre el promedio de la variable de resultado de los participantes y el promedio de la variable de resultado de los no participantes. Es decir, se puede utilizar $E[Y_i(0) | D_i = 0]$ como aproximación del contrafactual $E[Y_i(0) | D_i = 1]$, lo que permite recuperar el τ_{ATT} si y solo si:

$$E[Y_i(0) | D_i = 1] - E[Y_i(0) | D_i = 0] = 0$$

Esta condición implica que la variable de resultado en ausencia del programa debería ser idéntica para el grupo de individuos tratados y para el grupo de control. Si se cumple este supuesto que indica que el valor promedio de la variable de resultado en ausencia del programa es idéntico para el grupo de individuos tratados y no tratados, el efecto del programa está dado por:

$$\tau_{ATT} = E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 0]$$

Y el estimador estaría dado por:

$$\tau_{ATT} = (\bar{Y} | D = 1) - (\bar{Y} | D = 0)$$

Donde $(\bar{Y} | D)$ es el promedio muestral de la variable de resultado, dado el estatus de participación D . Esta comparación de medias se puede interpretar en la siguiente regresión:

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

Donde Y_i es la variable de resultado para el individuo i , D es una variable binaria o dummy que toma el valor de 1 si el individuo participa en el programa y de 0 si el individuo es elegible pero no participa, mientras que u_i es el término de error de la regresión que contiene las variables observadas y no observadas del individuo i , sin contar D_i , que afectan la variable de resultado. Si el supuesto se cumple, no existe correlación entre la participación del individuo y las características contenidas en u_i , ya que el supuesto indica que los individuos participantes son idénticos a los individuos no participantes. Es decir, la participación de los individuos es completamente independiente de las características de los mismos y por tanto $E [Y_i(0) | D_i = 0]$ es una aproximación adecuada del contrafactual $E [Y_i(0) | D_i = 1]$, lo que formalmente significa que:

$$E (u_i | D_i) = 0$$

Lo anterior se conoce como el supuesto de independencia condicional e indica que conocer D_i no añade información adicional acerca de u_i . Esto lleva a que el estimador de β_1 por Mínimos Cuadrados Ordinarios (MCO) sea consistente¹² e insesgado¹³.

El parámetro β_1 se interpreta como el efecto del programa o la diferencia de medias de la variable de resultado entre el grupo de tratamiento y el grupo de control, por lo que se obtiene:

$$E [Y_i(1) | D_i = 1] = E (\beta_0 + \beta_1 + u_i | D_i = 1) = \beta_0 + \beta_1 + E (u_i | D_i = 1) = \beta_0 + \beta_1$$

$$E [Y_i(0) | D_i = 0] = E (\beta_0 + u_i | D_i = 0) = \beta_0 + E (u_i | D_i = 0) = \beta_0$$

Debido a que $E (u_i | D_i) = 0$

Entonces:

$$\tau_{ATT} = E [Y_i(1) | D_i = 1] - E [Y_i(0) | D_i = 0] = (\beta_0 + \beta_1) - \beta_0 = \beta_1$$

El estimador de MCO de β_1 estaría dado por:

$$\hat{\beta}_1 = (\bar{Y} | D = 1) - (\bar{Y} | D = 0)$$

Específicamente, el objetivo de la metodología de diferencia en diferencias es establecer el efecto del tratamiento del programa, también conocido como la diferencia entre la variable de resul-

12 El estimador de β_1 por MCO, $\hat{\beta}_1$, converge en probabilidad a β_1 . Es decir, el límite del estimador a medida que aumenta el tamaño de muestra es exactamente el parámetro que se va a estimar.

13 El valor esperado del estimador de MCO es igual al valor del parámetro que se va a estimar, es decir, $E(\hat{\beta}_1) = \beta_1$, independientemente del tamaño de la muestra.

tado del individuo participante en el programa en presencia del programa y la variable de resultado de ese individuo en ausencia del mismo.

La primera diferencia de este procedimiento es la diferencia antes-y-después de los resultados para los grupos, comparando cada grupo con sí mismo y controlando por factores de ese grupo que son constantes en el tiempo. La segunda diferencia pretende controlar los factores que se dejaron de lado en la primera diferencia que varían a través del tiempo, es decir, se mide el cambio antes -y-después en las variables para un grupo que no participó en el programa, pero que fue expuesto a las mismas condiciones ambientales. Con esto, la metodología de diferencia en diferencias incluye comparaciones antes-y-después y comparaciones entre los que están o no están en el programa.

CUADRO 2
INFORMACIÓN REQUERIDA PARA DIFERENCIA EN DIFERENCIAS

	Tratamiento	Control
t=1 (línea de base)	Y1 D=1	Y1 D=0
t=2 (seguimiento)	Y2 D=1	Y2 D=0

Fuente: Bernal y Peña (2011).

Para definir el cambio esperado en Y entre el periodo posterior y el periodo anterior a la implementación del tratamiento en el grupo de tratamiento, menos la diferencia esperada en Y en el grupo de control durante el mismo periodo, se requiere la siguiente información:

Donde $t=1$ es el periodo anterior a la implementación del programa y $t=2$ es el periodo posterior a la implementación del tratamiento. Y_1 corresponde a la observación de Y en el periodo anterior al experimento y Y_2 es igual pero para el periodo después del tratamiento. La condición $|D$ indica si la observación corresponde a un individuo del grupo de tratamiento $D=1$ o al grupo de control $D=0$.

El impacto del programa por el método de diferencia en diferencias estaría dado por:

$$\tau_{dif-en-difs} = [E(Y_2 | D = 1) - E(Y_1 | D = 1)] - [E(Y_2 | D = 0) - E(Y_1 | D = 0)]$$

Por lo tanto, el estimador estaría dado por el análogo muestral de la ecuación anterior, es decir:

$$\tau_{dif-en-difs} = [(\bar{Y}_2 | D = 1) - (\bar{Y}_1 | D = 1)] - [(\bar{Y}_2 | D = 0) - (\bar{Y}_1 | D = 0)]$$

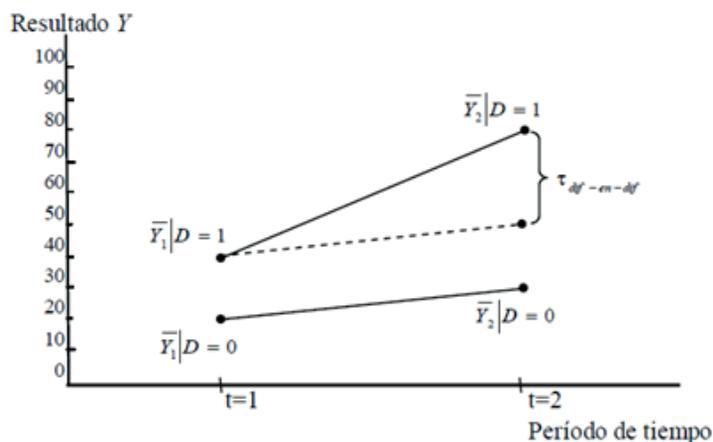
Donde $\bar{Y}_t | D$ es el promedio muestral de Y en el periodo t en el grupo D . El estimador se puede reescribir de la siguiente manera:

$$\tau_{dif-en-difs} = (\Delta \bar{Y} | D = 1) - (\Delta \bar{Y} | D = 0)$$

Donde $\Delta \bar{Y} | D$ es el cambio promedio de Y en el periodo $t=2$ y el periodo $t=1$ en el grupo D .

Gráficamente, se puede observar en la figura 1, en donde el supuesto de tendencia paralela que indica que la tendencia temporal de la variable de resultado entre el periodo $t=2$ y el periodo $t=1$ es la misma para el grupo de control y el grupo de tratamiento es lo que permite utilizar $(\bar{Y}_1 | D=1) - (\bar{Y}_1 | D=0)$ como un control apropiado de las diferencias preexistentes entre estos grupos. Si, por ejemplo, la variable Y crece naturalmente en el tiempo más rápido en el grupo de tratamiento entonces se sobreestimaré el efecto del programa.

FIGURA 1
EL ESTIMADOR DE DIFERENCIA EN DIFERENCIAS



Fuente: Stock y Watson (2003), tomado de Bernal y Peña (2011).

El estimador de diferencia en diferencias se puede obtener con base en un análisis de regresión. Si se define $\Delta \bar{Y}_i$ como el cambio en el valor de Y_i durante el transcurso del experimento, el efecto del programa sería el coeficiente β_1 de la siguiente ecuación:

$$Y_{it2} - Y_{it1} = \beta_0 + \beta_1 D_1 + (u_{it2} - u_{it1})$$

$$\Delta Y_i = \beta_0 + \beta_1 D_1 + v_i$$

Tomando la expectativa condicional, se obtiene que:

$$E(\Delta Y | D = 1) = \beta_0 + \beta_1 + E(v | D = 1) = \beta_0 + \beta_1$$

$$E(\Delta Y | D = 0) = \beta_0 + E(v | D = 0) = \beta_0$$

Donde la última igualdad en ambas ecuaciones se obtiene por el supuesto de independencia condicional $E(u_i | D_i) = 0$.

Entonces:

$$E(\Delta Y | D = 1) - E(\Delta Y | D = 0) = (\beta_0 + \beta_1) - \beta_0 = \beta_1$$

$$\tau_{\text{dif-en-dif}} = \beta_1$$

El estimador de β_1 por MCO ($\hat{\beta}_1$) corresponde al estimador de diferencia en diferencias y es insesgado si $E(v_i | D_i) = 0$.

Finalmente, dado que el objetivo de este trabajo es estimar insesgradamente el efecto del programa sobre la variable dependiente y no lo es maximizar la capacidad de proyección del programa, el indicador de bondad de ajuste R^2 no es relevante para el análisis.

IV. RESULTADOS

Mediante la creación del panel 2006-2007, utilizando las EHPM del INEC, se logró un empate de alrededor de 50% entre las encuestas de ambos años. El diseño de muestra de la EHPM incluye una muestra rotativa, es decir, no todas las viviendas son reemplazadas de la muestra de un año al siguiente, así un 75% de la muestra se repite de un año al otro. Por lo tanto, a la hora de crear un panel con dos años seguidos de la encuesta, se pierde (al menos) un 25% de las viviendas, un porcentaje similar si se compara con otras a nivel de Latinoamérica¹⁴. Algunos individuos adicionales pueden desaparecer del panel si estos se trasladan de lugar de residencia, dado que la encuesta entrevista las mismas viviendas (en lugar de hogares), si una familia se traslada de vivienda, sus integrantes no serían entrevistados nuevamente.

Para realizar este empate se desecharon aquellas observaciones que solamente se encontraron en uno de los años, es decir, las que no pertenecen al panel por no tener información de estos individuos para los dos años de estudio. Se debe destacar que en el año 2006 no se tienen beneficiarios de Avancemos, debido a que la encuesta se efectuó antes de la puesta en marcha del programa, esto permite que en el panel 2006-2007, el primero sea un buen año base.

Dado que el interés se centra en un grupo muy particular de la población, aquellos posibles beneficiarios de Avancemos o bien aquellos beneficiarios de Avancemos en el segundo año (2007), la muestra se restringe de acuerdo a algunos parámetros necesarios para acceder al programa, particularmente, la edad y el nivel educativo necesario para ser un posible beneficiario en el segundo año.

La puesta en marcha de Avancemos en el 2007 abarcaba a los estudiantes de 12 a 21 años. Por esta razón, se conservaron en la muestra las personas con 11 años en el 2006, ya que podían acceder al beneficio en el siguiente año. Por el mismo motivo, se excluyeron de la muestra aquellas personas que, en el 2006, ya habían cumplido los 21 años. Para el 2007, se conservan todos aquellos individuos cuya edad oscila entre 12 y 21 años inclusive, al ser estos la población objetivo del programa.

Para algunas observaciones, se encontraron incongruencias entre lo reportado en un año y otro, así por ejemplo, se encontraron individuos que reportaban una mayor edad en el 2006 que en el 2007, o bien individuos con más de un año de diferencia entre una encuesta y otra. Estas observaciones fueron excluidas de la base de datos de ambos periodos.

La variable nivel educativo se restringe para incluir solo aquellos individuos que cumplían con el requisito educativo necesario para acceder a Avancemos. Es decir, tener primaria aprobada o bien tener como último año aprobado algún nivel de secundaria académica o técnica. Se eliminaron aquellas personas que no habían terminado la primaria o ya habían finalizado la secundaria.

Posterior a estas filtraciones, el tamaño de la muestra en el panel antes y después de expandir (con el factor de expansión proporcionado por el INEC en la base de datos de la EHPM) queda de la siguiente manera:

CUADRO 3
MUESTRA DEL PANEL

	Panel 2006-2007	
	Sin expandir	Base expandida
Con Avancemos	275	21,164
Sin Avancemos	1,465	151,048

Fuente: Elaboración propia con datos de EHPM.

14 Ver por ejemplo, las encuestas de hogares en Argentina, Brasil y México.

Las observaciones que se indican “con Avancemos” solo recibieron el beneficio en el segundo año (2007), no así en el año base¹⁵. Las observaciones “sin Avancemos” no recibieron la TMC en ninguno de los dos años. Se trabajó con los datos muestrales de las EHPM; ya que, al expandir la encuesta, los resultados no variaban respecto a los obtenidos con la muestra.

Para llevar a cabo las estimaciones, se utilizó una serie de variables propias de las encuestas, así como algunas creadas a partir de estas. Particularmente, ciertas variables fueron recodificadas para convertirlas en dicotómicas. En el cuadro 1 del anexo 1, se presenta una breve descripción de todas las variables utilizadas, así como la distribución de los individuos de acuerdo con estas características; mientras, en el cuadro 2 del anexo 1, se catalogan de acuerdo a si recibieron Avancemos en el año 2007.

Para cada una de las estimaciones de la *propensity score matching*, así como la ecuación de diferencia en diferencias, se utilizaron diferentes especificaciones (ver anexo 2) que combinaban las variables descritas en el cuadro 1 del anexo 1. El tema de la pobreza como condición para acceder a Avancemos se abordó mediante dos visiones. En la primera se midió la pobreza a través de la línea de pobreza que establece el INEC, y en la segunda se incorporaron aquellas variables comunes entre la Ficha de Información Social (FIS) del IMAS y las Encuestas, tomando en cuenta que serían las utilizadas por el IMAS en el SIPO¹⁶ para determinar si un individuo es pobre o no.

El impacto del programa se espera principalmente sobre la deserción, pero también sobre la reinserción. Por este motivo, la evaluación se plantea sobre estas dos variables. Para ello, en ambas evaluaciones se requiere el uso de un grupo de control como contrafactual al grupo tratado. En tratamientos aleatorios este grupo de control es creado desde el inicio del programa, cuando se asigna el tratamiento aleatoriamente entre individuos. Sin embargo, dado que Avancemos no es un programa aleatorio, el grupo de control debe crearse posteriormente, basándose en características comunes con los individuos que sí fueron tratados, pero cuya única diferencia trascendental sea el haber sido parte del programa.

En el cuadro 3, se especificó el tamaño de la muestra, entre individuos que recibieron Avancemos y aquellos que se encuentran en el panel y cumplen los requisitos de educación y edad para ser beneficiarios del programa, pero que no recibieron el beneficio. De este último grupo, se deben identificar aquellos individuos que por sus características económicas y sociodemográficas –entre otras– son lo más parecidos posible a los individuos que sí recibieron la TMC. Así se podrían identificar dos individuos “iguales”, uno que recibió el tratamiento y el otro no, siendo este segundo el grupo contrafactual o individuo de control.

Dada la dificultad de encontrar dos individuos en una muestra o población con exactamente las mismas características, se crea un *pscore* que determina la probabilidad de cada individuo de recibir el programa Avancemos. Para estimar la probabilidad que tiene un individuo de obtener el beneficio se estimó un modelo logit, utilizando las medidas alternativas de pobreza, como se muestra en el cuadro 4. Una breve descripción de las variables utilizadas para el modelo logit puede ser consultada en el anexo 1, cuadro 1. Basados en dicha probabilidad, se busca para cada individuo tratado uno o más individuos no tratados, pero con similar probabilidad de tratamiento, siendo este segundo, el grupo de control a utilizar.

15 Como ya se mencionó, la encuesta 2006 es anterior a la puesta en marcha del programa en el mismo año.

16 Al momento de realizar este trabajo no se contaba con toda la información necesaria para estimar el índice del Sistema de Información de la Población Objetivo (SIPO) directamente.

CUADRO 4
RESULTADOS LOGIT 2006

Variable	Deserción pobreza INEC			Deserción pobreza SIPO			Reinserción pobreza INEC			Reinserción pobreza SIPO		
	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z
Mujer	0.18	0.14	0.20	0.18	0.14	0.19	-0.53	0.59	0.37	-0.36	0.65	0.58
Urbano	-0.88	0.15	0.00	-0.45	0.16	0.01	-1.88	1.11	0.09	-0.87	1.24	0.48
Nacional	0.40	0.50	0.42	0.75	0.52	0.15						
Seguro	0.21	0.23	0.37	0.30	0.24	0.21				0.93	0.84	0.27
Ingreso_percápita				0.00	0.00	0.00						
Estudio_jefe				-0.09	0.02	0.00				0.01	0.11	0.90
Casa_propia							0.36	0.76	0.63	0.19	0.83	0.82
Pobre1	1.12	0.34	0.00				2.23	0.94	0.02			
Pobre2	0.76	0.30	0.01				2.19	0.66	0.00			
Pobre3	-0.21	0.29	0.46									
Hacinado				0.32	0.32	0.33				0.14	0.96	0.89
Agua				-0.23	0.24	0.36				-0.35	0.78	0.66
Banno				0.12	0.69	0.87						
Serv_sanitario										-1.06	0.94	0.26
Paredes										1.63	1.34	0.23
Techo				-0.26	0.15	0.08				-2.26	1.41	0.11
Piso										-2.30	1.39	0.10
Refri				-0.22	0.30	0.45				2.56	1.42	0.07
Tv										-2.16	1.07	0.04
Lavadora										0.12	1.02	0.91
Sexto							-2.42	0.71	0.00	-0.90	1.26	0.48
Séptimo	0.14	0.18	0.43	0.25	0.18	0.17				1.02	1.31	0.44
Octavo	0.11	0.18	0.55	0.18	0.19	0.34						
Noveno	-0.18	0.22	0.41	-0.11	0.23	0.62	-1.22	1.20	0.31			

Continúa...

Continuación...

Variable	Deserción pobreza INEC			Deserción pobreza SIPO			Reinserción pobreza INEC			Reinserción pobreza SIPO		
	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z
Décimo	-1.20	0.48	0.01	-1.25	0.49	0.01	-2.58	0.79	0.00	-2.64	2.06	0.20
_cons	-2.19	0.58	0.00	-0.83	0.88	0.35						

Fuente: Elaboración propia con datos de EHPM.

Nota: Pobreza INEC se refiere a al uso de las variables de pobreza por ingreso (pobre1-pobre3) como medida de pobreza del individuo, mientras que pobreza SIPO se refiere al uso de algunas variables que son parte de las utilizadas por el IMAS para determinar el nivel de pobreza de las personas, además del ingreso.

Las variables del logit y su significancia varían en cada caso. Para las especificaciones que incluyen la medida de pobreza del INEC, la pobreza siempre fue significativa. El sufrir de pobreza o pobreza extrema aumenta la probabilidad de recibir Avancemos. Al contrario, residir en una zona urbana disminuye la probabilidad de recibir la transferencia.

En otras especificaciones del logit, también se incluyeron las variables comunes entre la FIS y la EHPM, que caracterizan al hogar. Aunque muchas de estas variables fueron rechazadas por el modelo o bien resultaron no significativas, sí resultaron significativas las variables de residencia urbana, así como la educación del jefe del hogar, ambas con signo negativo respecto a la probabilidad de recibir Avancemos. En el caso de reinserción, el modelo se enfrenta al problema de contar un menor número de observaciones tratadas en comparación con el caso de deserción (1,050 observaciones tratadas para reinserción en 2007 con la base expandida). Así, pocas variables resultan significativas dentro del modelo probabilístico.

Una vez estimado el modelo logit y calculado el *pscore* para cada una de las observaciones, se realizó el *matching*, estableciendo unidades de control para comparar con las unidades tratadas. En este caso, dada la cantidad de observaciones tratadas, se realizó el emparejamiento de cada unidad tratada con las 4 unidades de control más cercanas según su *pscore*. Además, se permitió el reemplazo de las unidades de control; es decir, una misma unidad puede ser el control de varias unidades tratadas.

Adicionalmente, se realizaron pruebas de balance para determinar si los grupos, tratado y control, son suficientemente similares como para presumir que se asemejan a grupos que hubiesen sido escogidos aleatoriamente. De manera específica, se realiza una prueba de balance para evaluar si las medias de los grupos son iguales. La hipótesis nula de ambos grupos iguales no se rechaza al 95% de confianza. Los resultados de las pruebas muestran que los grupos quedaron efectivamente balanceados. Además, es importante destacar que junto a la prueba de la hipótesis, el programa econométrico utilizado –Stata– también verifica el balance entre grupos. Como la teoría indica, el *pscore* se estima con base en las características de los grupos en el año base, y son estas las que deben estar balanceadas.

Una vez realizado el *matching*, se procedió a estimar un modelo de diferencia en diferencias, con las dos variantes de medidas de la pobreza. La variable independiente indica si el joven asistía al sistema educativo en el segundo año, midiendo la deserción; ya que en el año base, se partió de una muestra de individuos en que todos asistían, permitiéndonos medir el impacto de Avancemos sobre la deserción interanual.

Para deserción se obtiene que, bajo diferentes especificaciones, el impacto de Avancemos siempre es positivo y superior a 0, 14 pero menor a 0, 16. Lo que significa que el programa combate la deserción, evitando que deserten entre 14 y 16 jóvenes de cada 100 a los que se les entrega la TMC. En los otros casos podría suceder que no desertan gracias a una serie de condiciones, entre ellas Avancemos. La cualidad de diferencia en diferencias es que aísla esos otros efectos, por lo que el impacto aislado de Avancemos oscila entre 0,14 y 0,16.

Estos resultados surgen de varias especificaciones que incluyen diferentes variables de control (ver anexo 2), así como efectos fijos por individuos para controlar por las características innatas, no medibles, de los estudiantes. Las diferentes especificaciones son muestra de lo robusto que es el impacto de Avancemos sobre deserción. En el cuadro 5, se muestra el coeficiente de diferencia en diferencias, que denota el impacto de Avancemos. Se muestran los resultados de las distintas especificaciones, para las dos medidas de pobreza consideradas.

El impacto es similar bajo ambas medidas de pobreza, la línea de ingreso del INEC y la aproximación al SIPO. Mientras que el impacto varía ligeramente cuando se van agregando controles a la regresión, disminuye un punto porcentual (p.p.) cuando se agregan variables de control y otro punto porcentual cuando se incluyen efectos fijos por individuo. No obstante, el impacto se mantiene positivo y bastante similar en su nivel. En todos los casos, el resultado es estadísticamente significativo.

CUADRO 5
RESULTADO DE IMPACTO DE AVANCEMOS SOBRE DESERCIÓN

Medida de pobreza usada	INEC			Proxy SIPO		
	(1)	(2)	(3)	(4)	(5)	(6)
2006-2007	0,16***	0,15***	0,14***	0,16***	0,15***	0,14***
	(0,016)	(0,015)	(0,016)	(0,021)	(0,018)	(0,018)
Controles después del matching						
Variables de control	NO	SI	SI	NO	SI	SI
Efectos fijos	NO	NO	SI	NO	NO	SI

Error estándar entre paréntesis

*** significativo al 1%

Nota: las variables de control utilizadas pueden ser consultadas en el anexo 2

Fuente: Elaboración propia con datos de EHPM.

Para evaluar la reinserción, se dispone de menos observaciones que en deserción; sin embargo, a pesar de esta restricción, se encuentra un coeficiente positivo asociado al programa Avance-mos. Es decir, se encuentra que Avance-mos también logra resultados positivos en reinsertar en el sistema educativo a aquellos jóvenes que ya habían abandonado las aulas.

Se mide reinserción gracias a la variable de “asiste” en el segundo año y partiendo de una base de individuos que no asistían al sistema educativo en el primer año. En el 2007, se encuentra un impacto grande, siendo entre 85% y 89% el porcentaje de jóvenes que regresaron a estudiar gracias al programa. En el cuadro 6, se muestran los resultados sobre reinserción, medido a través de distintas especificaciones y distintas aproximaciones de pobreza, esto para mostrar la robustez del resultado.

Contrario al caso de deserción, en reinserción, el impacto aumenta cuando se agregan controles a la regresión. Así, el impacto aumenta 2 p.p. cuando se incluyen variables de control, pero se mantiene igual cuando, posteriormente, se utilizan efectos fijos por individuos. En la regresión básica, antes de los controles, el impacto es mayor cuando se utiliza la aproximación de pobreza por SIPO. Posterior a los controles, el impacto es igual entre ambas medidas alternativas de pobreza. Todos estos resultados muestran lo constante del impacto positivo, el cual resulta además estadísticamente significativo en las 6 especificaciones expuestas en el cuadro 6.

CUADRO 6
RESULTADO DE IMPACTO DE AVANCEMOS SOBRE REINSERCIÓN

Medida de pobreza usada	INEC			Proxy SIPO		
	(1)	(2)	(3)	(4)	(5)	(6)
2006-2007	0,85***	0,87***	0,87***	0,89***	0,87***	0,87***
	(0,049)	(0,039)	(0,030)	(0,048)	(0,056)	(0,056)
Controles después del matching						
Variables de control	NO	SI	SI	NO	SI	SI
Efectos fijos	NO	NO	SI	NO	NO	SI

Error estándar entre paréntesis

*** significativo al 1%

Nota: las variables de control utilizadas pueden ser consultadas en el anexo 2

Fuente: Elaboración propia con datos de EHPM.

Por lo tanto, los resultados muestran que el programa Avancemos tuvo, en su primer año de implementación, un impacto positivo sobre la deserción de los estudiantes en secundaria, y mayor aún en el caso de reinserción. Los resultados además son robustos, manteniéndose en niveles similares de impacto antes diferentes modelos econométricos.

V. CONCLUSIONES

Este estudio buscaba llenar el vacío en las evaluaciones de impacto del programa de transferencias monetarias condicionadas, Avancemos, en Costa Rica. Como primer punto de análisis, se planteó evaluar la implementación del programa. Dado que el diseño del programa no fue acompañado de un diseño experimental de evaluación, se recurrió a técnicas econométricas cuasi-experimentales para evaluar su impacto.

Se encuentra que los datos del período 2006-2007 muestran la transferencia monetaria condicionada como un programa con un impacto positivo, tanto sobre la deserción como sobre la reinserción. Esto significa que una parte de los estudiantes no abandonan los estudios gracias a la ayuda económica que reciben y, por otro lado, también ayuda a que los jóvenes que ya han desertado regresen a estudiar.

Algunos de estos jóvenes se mantendrán en el sistema educativo gracias a otra serie de factores, aunados al de contar con la transferencia; pero entre un 14% y 16% de ellos, lo hacen única y exclusivamente debido a Avancemos, de otra forma abandonarían sus estudios. Así mismo, los que se reintegran al sistema educativo, lo hacen por diversos motivos, pero un porcentaje bastante alto (alrededor del 87%) lo hizo debido a la transferencia que se les entregó.

Por lo tanto, el programa Avancemos es un programa con impactos positivos sobre sus objetivos planteados: evitar la deserción y procurar la reinserción. A pesar de que la muestra de reinserción es menor que la de deserción, el impacto positivo de la transferencia es mucho mayor en reinserción, planteando una línea de acción para el programa.

Este estudio constituye un primer acercamiento al análisis de impacto del programa, tarea que se encontraba pendiente en el país. Sin embargo, es importante evaluar su impacto en periodos más recientes. Esto permitirá evaluar la evolución del diseño y ejecución del programa, así como identificar los posibles puntos de enfoque para incrementar el impacto positivo, con base en otras evaluaciones de procesos, resultados e impacto.

VI. REFERENCIAS

- Bernal, R. y Peña, X. (2011). Guía práctica para la evaluación de impacto. Colombia: Facultad de Economía, Centro de Estudios sobre Desarrollo Económico. Universidad de los Andes.
- Contraloría General de la República (CGR) (2008). *Informe sobre el Diseño y Ejecución del programa de transferencia monetaria condicionada denominado "Avancemos"*. San José, Costa Rica: Contraloría General de la República (CGR). División de Fiscalización Operativa y Evaluativa, Área de Servicios Sociales. . INFORME No. DFOE-SOC-17-2008.
- Contraloría General de la República (CGR) (2012). *Informe sobre la gestión del programa de transferencia monetaria condicionada Avancemos*. San José, Costa Rica: Contraloría General de la República (CGR). División de Fiscalización Operativa y Evaluativa, Área de Servicios Sociales. INFORME No. DFOE-SOC-IF-10-2012.
- Decreto Ejecutivo N° 33154 (2006). Crea Programa de transferencia monetaria para promover mantenimiento de los y las adolescentes en el sistema educativo formal respecto de educación

- académica como formación técnica en diversas modalidades del sistema educativo no formal que ofrece el INA. Publicado en: *La Gaceta*. N° 96, el día 19 de mayo del 2006.
- Duryea, S. y Morrison, A. (2004). *The Effect of Conditional Transfers on School Performance and Child Labor: Evidence from an Ex-Post Impact Evaluation in Costa Rica*. Washington D.C.: Inter-American Development Bank.
- Fiszbein, A. y Schady, N. (2009). *Conditional Cash Transfers. Reducing present and future poverty*. Washington D.C.: The International Bank for Reconstruction and Development, The World Bank.
- Gertler, P. Martínez, S., Premand, P., Rawlings, L. y Vermeersch, C. (2011). *Impact Evaluation in Practice*. Washington D.C.: The International Bank for Reconstruction and Development, The World Bank.
- Heckman, J., Ichimura, H. y Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluation a Job Training Programme. En: *Review of Economic Studies*, Vol. 64.
- Instituto Nacional de Estadística y Censos (INEC) (s.f.). *Documento metodológico: Encuesta de hogares y propósitos múltiples*. San José, Costa Rica: Instituto Nacional de Estadística y Censos (INEC). Tomado de: http://www.inec.go.cr/ZONA_EHPM/index.htm, el día 30 de mayo de 2010.
- Instituto Nacional de Estadística y Censos (INEC). (2007). Encuesta de Hogares de Propósitos Múltiples 2006. [Archivo de datos].
- Instituto Nacional de Estadística y Censos (INEC). (2008). Encuesta de Hogares de Propósitos Múltiples 2007. [Archivo de datos].
- Khandker, S.; Koolwal, G.; Samad, H. (2010). *Handbook on Impact Evaluation. Quantitative Methods and Practices*. Washington D.C.: The International Bank for Reconstruction and Development, The World Bank.
- Ravallion, M. (2008). Evaluating Anti-Poverty Programs. En: *Handbook of Development Economics*, Vol. 4, ed. Paul Schultz and John Strauss. Amsterdam: North Holland.
- Rawlings, L. (2004). *A New Approach to Social Assistance: Latin America's Experience with Conditional Cash Transfer Programs*. Social Protection Unit, Human Development Network. The World Bank.
- Rosenbaum, P. y Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. En: *Biometrika*, Vol. 70, No. 1. (Apr., 1983), pp. 41-55.
- Viquez, R. (2005). *Sistema de Identificación de la Población Objetivo: SIPO en Costa Rica. Serie de Documentos de Discusión sobre la Protección Social. Banco Mundial*. Tomado de: <http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/0530Spanish.pdf>

VII. AGRADECIMIENTOS

Se agradecen los valiosos comentarios de Juan Diego Trejos, Juan Robalino y Diego Rojas al trabajo previo del que se desprende este artículo.

Anexo 1

CUADRO 1
VARIABLES UTILIZADAS

Variable	X=0	X=1
Pobre1	Otra condición de pobreza	Pobre extremo
Pobre2	Otra condición de pobreza	Pobre no extremo
Pobre 3	Otra condición de pobreza	No pobre
Pobre 4	Otra condición de pobreza	Ignorado o sin ingreso
Urbano	Zona rural	Zona urbana
Nacional	Extranjero	Costarricense
Seguro	No posee seguro social	Posee seguro social
Asiste	No asiste al sistema educativo	Asiste al sistema educativo
Sexto	Asiste a otro nivel	Asiste a sexto año
Sétimo	Asiste a otro nivel	Asiste a sétimo año
Octavo	Asiste a otro nivel	Asiste a octavo año
Noveno	Asiste a otro nivel	Asiste a noveno año
Décimo	Asiste a otro nivel	Asiste a décimo año
Undécimo	Asiste a otro nivel	Asiste a undécimo año
Casa propia	No posee casa propia	Posee casa propia
Mujer	Hombre	Mujer
Paredes	Paredes en mal estado	Paredes en buen estado
Techo	Techo en mal estado	Techo en buen estado
Piso	Piso en mal estado	Piso en buen estado
Tv	No posee televisión	Posee televisión de algún tipo
Refri	No posee refrigeradora	Posee refrigeradora
Lavadora	No posee lavadora	Posee lavadora
Agua	No tiene acceso a agua	Tiene acceso a agua
Serv_sanitario	No posee servicio sanitario	Posee servicio sanitario
Banno	No posee baño	Posee baño
Elect	No posee electricidad	Posee electricidad
Hacinado	Hogar no hacinado	Hogar hacinado (más de 3.25 individuos por dormitorio)
Avance07	No beneficiario de Avancemos en el 2007	Beneficiario de Avancemos en el 2007
Anno	Si año es 2006	Si año es 2007
Treated	Si es una observación de control	Si es una observación tratada
Ingreso_percápita	Variable continua	
Estudio_jefe	Variable continua	
dd	dd= anno*treated	

Fuente: Elaboración propia con datos de EHPM.

CUADRO 2
DISTRIBUCIÓN DE LA MUESTRA POR CADA VARIABLE UTILIZADA

Características en 2006		Avancemos en 2007			
		0		1	
Total		2059		290	
Pobre1	0	1948	95%	247	85%
	1	111	5%	43	15%
Pobre2	0	1698	82%	198	68%
	1	361	18%	92	32%
Pobre 3	0	628	31%	151	52%
	1	1431	69%	139	48%
Pobre 4	0	1903	92%	274	94%
	1	156	8%	16	6%
Urbano	0	1245	60%	220	76%
	1	814	40%	70	24%
Nacional	0	65	3%	5	2%
	1	1994	97%	285	98%
Seguro	0	358	17%	28	10%
	1	1701	83%	262	90%
Asiste	0	594	29%	15	5%
	1	1465	71%	275	95%
Sexto	0	1215	59%	196	68%
	1	844	41%	94	32%
Sétimo	0	1674	81%	205	71%
	1	385	19%	85	29%
Octavo	0	1653	80%	221	76%
	1	406	20%	69	24%
Noveno	0	1760	85%	253	87%
	1	299	15%	37	13%
Décimo	0	1940	94%	285	98%
	1	119	6%	5	2%
Undécimo	0	2053	100%	290	100%
	1	6	0%	0	0%
Casa propia	0	307	15%	56	19%
	1	1752	85%	234	81%
Mujer	0	1070	52%	138	48%
	1	989	48%	152	52%
Paredes	0	991	48%	182	63%
	1	1066	52%	108	37%
Techo	0	902	44%	161	56%

Continúa...

Continuación...

Avancemos en 2007					
	1	1155	56%	129	44%
Piso	0	820	40%	168	58%
	1	1237	60%	122	42%
Tv	0	85	4%	20	7%
	1	1972	96%	270	93%
Refri	0	129	6%	24	8%
	1	1928	94%	266	92%
Lavadora	0	132	6%	23	8%
	1	1925	93%	267	92%
Agua	0	164	8%	37	13%
	1	1893	92%	253	87%
Serv_sanitario	0	102	5%	32	11%
	1	1955	95%	258	89%
Banno	0	33	2%	3	1%
	1	2024	98%	287	99%
Elect	0	9	0%	0	0%
	1	2048	99%	290	100%
Hacinado	0	1967	96%	268	92%
	1	90	4%	22	8%

Fuente: Elaboración propia con datos de EHPM.

Notas: La distribución incluye tanto la muestra para deserción como para reinserción.

A partir de la variable de paredes se pierden dos observaciones sin avancemos, para las cuales no se tiene información.

Los valores de 0 y 1 representan las características descritas en el cuadro A.1

Anexo 2

Especificaciones utilizadas para realizar el Matching y la ecuación de Diferencia en Diferencias de la Evaluación de Impacto de Avancemos

1. Deserción 2006-2007 utilizando pobreza del INEC

Pscore:

$$avance07 = \beta_0 + \beta_1 pobre1 + \beta_2 pobre2 + \beta_3 pobre3 + \beta_4 urbano + \beta_5 nacional + \beta_6 seguro + \beta_7 sétimo + \beta_8 octavo + \beta_9 noveno + \beta_{10} décimo + \beta_{11} mujer + \beta$$

Diferencia en diferencias:

$$asiste = \beta_0 + \beta_1 dd + \beta_2 anno + \beta_3 treated + \beta_4 pobre1 + \beta_5 pobre2 + \beta_6 pobre3 + \beta_7 urbano + \beta_8 nacional + \beta_9 seguro + \beta_{10} sexto + \beta_{11} sétimo + \beta_{12} octavo + \beta_{13} noveno + \beta_{14} décimo + \beta_{15} mujer + \beta$$

2. Deserción 2006-2007 utilizando pobreza aproximada por variables del SIPO

Pscore:

$$\text{avance07} = \beta_0 + \beta_1 \text{urbano} + \beta_2 \text{nacional} + \beta_3 \text{seguro} + \beta_4 \text{sétimo} + \beta_5 \text{octavo} + \beta_6 \text{noveno} + \beta_7 \text{décimo} + \beta_8 \text{mujer} + \beta_9 \text{techo} + \beta_{10} \text{ingresopercapita} + \beta_{11} \text{refri} + \beta_{12} \text{agua} + \beta_{13} \text{banno} + \beta_{14} \text{hacinado} + \varepsilon$$

Diferencia en diferencias:

$$\text{asiste} = \beta_0 + \beta_1 \text{dd} + \beta_2 \text{anno} + \beta_3 \text{treated} + \beta_4 \text{urbano} + \beta_5 \text{nacional} + \beta_6 \text{seguro} + \beta_7 \text{sexto} + \beta_8 \text{sétimo} + \beta_9 \text{octavo} + \beta_{10} \text{noveno} + \beta_{11} \text{décimo} + \beta_{12} \text{mujer} + \beta_{13} \text{techo} + \beta_{14} \text{ingresopercapita} + \beta_{15} \text{refri} + \beta_{16} \text{agua} + \beta_{17} \text{banno} + \beta_{18} \text{hacinado} + \varepsilon$$

3. Reinserción 2006-2007 utilizando pobreza del INEC

Pscore:

$$\text{avance07} = \beta_0 + \beta_1 \text{pobre1} + \beta_2 \text{pobre2} + \beta_3 \text{urbano} + \beta_4 \text{nacional} + \beta_5 \text{sexto} + \beta_6 \text{octavo} + \beta_7 \text{noveno} + \beta_8 \text{décimo} + \beta_9 \text{casa propia} + \beta_{10} \text{mujer} + \varepsilon$$

Diferencia en diferencias:

$$\text{asiste} = \beta_0 + \beta_1 \text{dd} + \beta_2 \text{anno} + \beta_3 \text{treated} + \varepsilon$$

4. Reinserción 2006-2007 utilizando pobreza aproximada por variables del SIPO

Pscore:

$$\text{avance07} = \beta_0 + \beta_1 \text{urbano} + \beta_2 \text{nacional} + \beta_3 \text{seguro} + \beta_4 \text{sexto} + \beta_5 \text{sétimo} + \beta_6 \text{octavo} + \beta_7 \text{noveno} + \beta_8 \text{décimo} + \beta_9 \text{casa propia} + \beta_{10} \text{mujer} + \beta_{11} \text{paredes} + \beta_{12} \text{techo} + \beta_{13} \text{piso} + \beta_{14} \text{serv sanitario} + \beta_{15} \text{tv} + \beta_{16} \text{refri} + \beta_{17} \text{agua} + \beta_{18} \text{banno} + \beta_{19} \text{hacinado} + \beta_{20} \text{lavadora} + \beta_{21} \text{estudio_jefe} + \varepsilon$$

Diferencia en diferencias:

$$\text{asiste} = \beta_0 + \beta_1 \text{dd} + \beta_2 \text{anno} + \beta_3 \text{treated} + \varepsilon$$



