



**“EVALUACIÓN DE IMPACTO DEL PROGRAMA URBANO DE
BIENESTAR PARA LA NIÑEZ EN EXTREMA POBREZA”**

**TESIS PARA OPTAR AL GRADO DE
MAGISTER EN POLÍTICAS PÚBLICAS**

Alumno: Karla María Bayres Mejía

Profesor Guía: Rodrigo Herrera Pinto

Santiago, enero de 2017

EVALUACIÓN DE IMPACTO DEL “PROGRAMA URBANO DE BIENESTAR PARA LA NIÑEZ EN EXTREMA POBREZA”

Resumen

Este trabajo evalúa el impacto de los servicios de guardería en Centros de Cuidado Infantil Comunitario (CICO) y visitas domiciliarias ofrecidos por el Programa Urbano de Bienestar para la Niñez en Extrema Pobreza en Nicaragua. El programa fue ejecutado por el Ministerio de la Familia, la Adolescencia y la Niñez Nicaragüense entre inicios del año 2014 y mediados del año 2016. Se utilizan datos recopilados por el Centro de Investigación y Acción educativa Social (CIASES) en noviembre del año 2013 y noviembre del año 2015. Se estima el efecto del programa en la reducción de niños con problemas en su desarrollo. El impacto del programa es estimado mediante dos estrategias. Primero, se mide el impacto del tratamiento de los tratados en un contexto de aleatorización de los servicios y luego se reestima el impacto del programa mediante el método del emparejamiento. Los principales resultados señalan que, tanto por estimaciones paramétricas como no paramétricas, el programa reduce el porcentaje de niños con problemas en su desarrollo en alrededor de 7 y 12 puntos porcentuales.

Palabras claves: políticas públicas de primera infancia, extrema pobreza, evaluación de impacto, desarrollo integral de la niñez

Contenido

1	Introducción	4
2	Descripción del Programa	8
	a. ¿Cuál es el objetivo del Programa?	9
	b. ¿Cómo lo hace?	10
	c. ¿Cómo se asignaron los beneficios?	11
3	Evidencia internacional: Impacto de intervenciones de apoyo a la primera infancia	14
4	Metodología	18
	Estrategia 1: Aleatorización del Programa	18
	Estrategia 2: Diseño cuasi-experimental: Emparejamiento	20
5	Los Datos	23
	a. Descripción de encuestas.....	23
	b. Instrumento de Medición de Impacto	24
	c. Estadística Descriptiva pre-tratamiento.....	25
6	Resultados	31
	a. Estrategia 1: Aleatorización del Programa	31
	b. Estrategia 2: emparejamiento.....	33
7	Recomendaciones de Política	38
8	Conclusiones.....	42
9	Referencias Bibliográficas	44
10	Anexos	47

1 Introducción

Hoy en día, existe un creciente interés por los Gobiernos de América Latina y el Caribe por implementar y expandir políticas efectivas de desarrollo infantil debido a diversos beneficios que estas intervenciones prometen tener en los niños de hoy y los adultos de mañana. Por un lado, diversas investigaciones neurocientíficas han demostrado que el entorno y los servicios de cuidado infantil que reciben los niños durante los primeros años de su vida son esenciales en su proceso de aprendizaje. Esto se debe a que en los primeros años de vida, se desarrolla gran parte del cerebro y con él, circuitos específicos que sientan las bases para las habilidades presentes y subyacentes de cada niño (National Scientific Council). En Chile, por ejemplo, Novoa-Hidalgo y Urzúa (2012) encontraron mejoras significativas en el área motora (fina y gruesa); habilidades de expresión personal (expresión de las emociones) y hábitos adaptativos en la alimentación de niños que asistieron a salas cunas públicas provistas por la Junta Nacional de Jardines Infantiles (JUNJI) y la Fundación Nacional para el Desarrollo Integral del Menor (INTEGRA).

Adicionalmente, se espera que intervenciones focalizadas en niños que viven en ambientes desfavorables y con severas carencias materiales, reduzcan brechas existentes entre clases sociales y ayuden a romper círculos viciosos de pobreza (Magnuson & Duncan, 2014). En este sentido, Baker-Henningham y López-Boo (2013) comparan la efectividad de más de 30 programas orientados a la estimulación del desarrollo infantil implementados en países de ingresos bajos y medios, y concluyen que las intervenciones más efectivas estaban orientadas a niños que viven en entornos desfavorables, con edades más tempranas y cuyas madres y familias fueron integradas en los programas.

Finalmente, se considera que inversiones en edades tempranas resultan más convenientes y costo-efectivas para los Gobiernos, que intervenciones en edades adultas. Así lo demostraron Heckman y Masterov (2007) quienes afirman que intervenciones en los primeros años de vida pueden incidir en la reducción de indicadores relevantes en el crecimiento de un país, como son, las tasas de crimen, deserción escolar y porcentaje de embarazos juveniles. Existen pues suficientes motivos para que los hacedores de políticas públicas enfoquen su interés y recursos en programas que apoyen el desarrollo temprano en la primera infancia.

Ante dichos argumentos y en un contexto de severas privaciones materiales, el Gobierno de Nicaragua creó el Programa Urbano de Bienestar para la Niñez en Extrema Pobreza. Este programa representa una inversión de 15 millones de dólares para el país, fue financiado por el Banco Interamericano de Desarrollo y ejecutado por el Ministerio de la Familia, la Adolescencia y la Niñez Nicaragüense (MIFAN). El principal objetivo del programa es mejorar el desarrollo integral de los niños menores de 6 años que viven en barrios urbanos en extrema pobreza, mediante la entrega de diversos servicios de atención infantil: apoyo para inscripción en el registro civil del niño, construcción de guarderías infantiles, visitas domiciliarias por educadoras de párvulos, entrega de vitaminas, etc.

El Programa fue diseñado para ser implementado en dos etapas, de manera tal que ciertos niños de barrios intervenidos fuesen beneficiados en una primera fase; mientras otros esperarían por recibir los beneficios del programa en un futuro. Para determinar la población elegible de los servicios del programa, primero se eligieron a los municipios que tuviesen las tasas más altas de pobreza y que además concentraran el 50 por ciento de los hogares urbanos en dicha condición. Luego, se seleccionaron aquellos barrios, dentro de dichos municipios, cuya tasa de pobreza fuese igual o mayor al doble de la tasa promedio de pobreza urbana a nivel nacional. Se eliminaron de la muestra los barrios que no cumplían con condiciones de urbanidad ni viabilidad para ser intervenidos. Como resultado de este proceso, se contó con 72 barrios con sus respectivas familias y niños elegibles para ser beneficiados por el Programa en sus distintas fases; los cuales, después fueron designados aleatoriamente para ser grupos intervenidos o de control.

Es importante mencionar que se ofrecieron los servicios aleatoriamente a nivel de barrios por razones éticas, pues todos los niños que habitaban en los barrios priorizados, eran igualmente vulnerables y no se podía excluir a unos y a otros no. Los servicios del programa fueron gratuitos para las familias intervenidas, no era obligatorio participar y se ofrecieron con suficientes cupos para cubrir la potencial demanda de los hogares.

Así pues, el presente estudio tiene por objetivo evaluar qué impacto tuvieron los servicios de guardería y visitas domiciliarias, ofrecidos por el “Programa Urbano de Bienestar para la Niñez en Extrema Pobreza”, sobre el desarrollo integral de niños menores de 6 años. Para ello este estudio utiliza datos de panel, es decir datos de línea base y seguimiento para un conjunto de familias

elegibles para ser intervenidas por el Programa. El Programa empezó a ejecutarse en el año 2014, la Línea Base fue levantada en noviembre del año 2013; mientras que los datos de seguimiento, en noviembre del año 2015.

El instrumento de medición utilizado para evaluar los resultados del programa es el cuestionario CLAP de Desarrollo Infantil. Éste evalúa el desempeño del niño en diversas tareas adecuadas para su edad. El instrumento CLAP fue seleccionado por el MIFAN por dos motivos: i) requiere relativamente poco tiempo para implementarse (15-20 minutos por niño) y ii) es el instrumento que dicha institución ha utilizado desde hace años en diversos centros de cuidado infantil para dar seguimiento al desarrollo de los niños.

Dadas las características de selección de beneficiarios, así como la implementación final que tuvo el programa, el impacto se encuentra a partir del uso de dos metodologías. En un primer momento, en un escenario de aleatorización, el impacto se busca mediante la simple diferencia de medias entre niños de barrios intervenidos que participan del Programa y niños de barrios no intervenidos. Adicionalmente se controla la estimación por diversas variables que pueden incidir en el desarrollo del niño: características de la madre, características del jefe de hogar y condiciones socioeconómicas del hogar.

Luego, y dado que existen diferencias significativas entre los niños tratados y no tratados en términos de características observables, especialmente las relacionadas a las condiciones socioeconómicas del hogar, se utiliza el método de emparejamiento para hacerse cargo del potencial sesgo de selección presente. Esta metodología permite revisar la robustez de la estimación anterior ante potenciales problemas del diseño de evaluación experimental, como el no cumplimiento del protocolo en la aleatorización o pérdida de la muestra por bajo take-up. Las variables pre-tratamiento especificadas en el Propensity Score corresponden a dos tipos: características socioeconómicas del hogar y características relacionadas con la madre del niño.

Los principales resultados muestran que el impacto estimado mediante la simple diferencia de medias entre niños tratados y no tratados no es estadísticamente significativo. Sin embargo, el análisis de estadística descriptiva permite reconocer diferencias en características observables entre ambos grupos. Una vez que se controla por dichas diferencias se encuentra un impacto

estadísticamente significativo del Programa de alrededor de 7 puntos porcentuales. Luego, se reestima el impacto del Programa por medio del método de Propensity Score Matching (PSM) o emparejamiento. En este caso se encuentra un impacto significativo del Programa entre 7 y 12 puntos porcentuales. Se observa que, en la medida en que se usan menos controles, los grupos son más parecidos en términos de variables observables, lo que permite usar los niños más parecidos entre los grupos tratados y no tratados, por lo que se reduce el sesgo del estimador.

El estudio que se presenta a continuación, está estructurado de la siguiente forma: la sección 2 presenta una descripción del Programa en relación a lo que hace, cómo lo hace y cómo entregó sus beneficios. En la sección 3 se hace un breve resumen de resultados encontrados en evaluaciones similares a esta. La sección 4 presenta la metodología utilizada en este estudio; mientras que la sección 5 describe los datos utilizados. Luego se entregan los resultados estimados. La siguiente sección presenta dificultades encontradas durante el diseño e implementación del Programa y entrega recomendaciones de política pública para futuras intervenciones. Finalmente, en la sección 8 se encuentran las conclusiones.

2 Descripción del Programa

Nicaragua es un país con altos niveles de vulnerabilidad, pobreza y retrasos en el desarrollo infantil. Para el 2009, según la Encuesta de Medición de Nivel de Vida, el 42.5 por ciento de la población nicaragüense era pobre, el 35.9 por ciento de ellos residía en áreas urbanas y casi el 50 por ciento de niños entre 0 y 5 años vivía en condiciones de pobreza general (INIDE, 2011). Por otro lado, resultados preliminares de un estudio del BID, evidencian que para el año 2008, el 85% de niños nicaragüenses de 5 años pertenecientes a comunidades rurales pobres mostraban un retraso de 21 meses en su desarrollo cognitivo (Shady, y otros, 2014). Por su parte, la población infantil que residía en áreas rurales venía siendo atendida desde el año 1996 a través del Programa de Atención Integral a la Niñez (PAININ) en sus diversas fases. Como respuesta a la preocupación del Gobierno de Nicaragua para incidir en la realidad que experimentaban miles de niños y niñas menores de 6 años que estaban siendo desatendidos en las áreas urbanas, se crea entonces el Programa Urbano de Bienestar para la Niñez en Extrema Pobreza.

El Programa Urbano de Bienestar para la Niñez en Extrema Pobreza – de ahora en adelante denominado sólo “El Programa” – es un proyecto piloto implementado en Nicaragua, financiado por préstamo del Banco Interamericano del Desarrollo (BID) y ejecutado por el Ministerio de la Familia, la Adolescencia y la Niñez Nicaragüense (MIFAN). Éste surge como iniciativa de Proyecto de Ley en el año 2009 y se circunscribió dentro del Sistema Nacional de Bienestar Social (SNBS) como Política Nacional¹ de Atención a la Primera Infancia por parte del Gobierno de Nicaragua.

El equipo formulador del programa identificó que las políticas existentes para promover el desarrollo infantil a nivel urbano tenían muy poca cobertura, baja participación de las madres y padres y cero mecanismos de focalización en su implementación. Este programa se creó como un aporte al SNBS para realizar una intervención focalizada a los niños pobres extremos urbanos que expandiera la cobertura de servicios sociales y promoviera la demanda de dichos servicios por las familias más vulnerables, para mejorar el desarrollo integral de los niños intervenidos (BID, Perfil de Proyecto. Programa Urbano de Desarrollo Integral para Familias en Extrema Pobreza., 2009).

¹ El SNBS fue creado en el 2008 mediante decreto presidencial como un espacio de coordinación de los programas sociales del país. El programa Amor es un componente del SNBS que se focaliza en garantizar servicios de bienestar social para la niñez nicaragüense (0 a 18 años) y el Programa Urbano se registra bajo este componente.

El Programa fue diseñado para ser implementado escalonadamente en dos fases: La primera, resultó ejecutándose entre inicios del 2014 y finales del año 2016, en formato de piloto. De los resultados de la evaluación de la primera fase se esperaba obtener la evidencia necesaria que permita tomar la decisión de expandir o no el Programa en su segunda versión. Este estudio se enmarca justamente en dicho proceso evaluativo.

a. ¿Cuál es el objetivo del Programa?

El Fin último del Programa es mejorar el desarrollo integral de los niños menores de 6 años que viven en barrios urbanos en extrema pobreza, a través de la ampliación del acceso y mejora de calidad de servicios comunitarios de atención infantil. Así pues, la lógica vertical del programa se basa en la teoría de que intervenciones institucionales en el cuidado infantil y en la promoción de capacidades parentales, mejoran el desarrollo integral en los niños menores de 6 años (BID, 2016).

En este sentido, se seleccionaron cuatro áreas claves en el desarrollo del niño en las que se deseaba incidir: i) el desarrollo físico y motor, ii) las destrezas lingüísticas, iii) las habilidades cognitivas y iv) el desarrollo socio-emocional, las cuales son evaluadas mediante una sola variable de desarrollo global del niño. Este es el objetivo primordial del Programa y el componente en el que se invierte más dinero en relación a los otros servicios que el Programa ofrece como son entrega de multi-vitaminas, seguimiento de inscripción en el registro civil, entre otros. Por lo tanto, secundariamente, el programa también se propuso reducir la desnutrición crónica infantil, promover la inscripción pertinente a la escuela primaria y mejorar los conocimientos parentales de cuidado infantil.

Consecuentemente, los indicadores de impacto establecidos para el Programa, son:

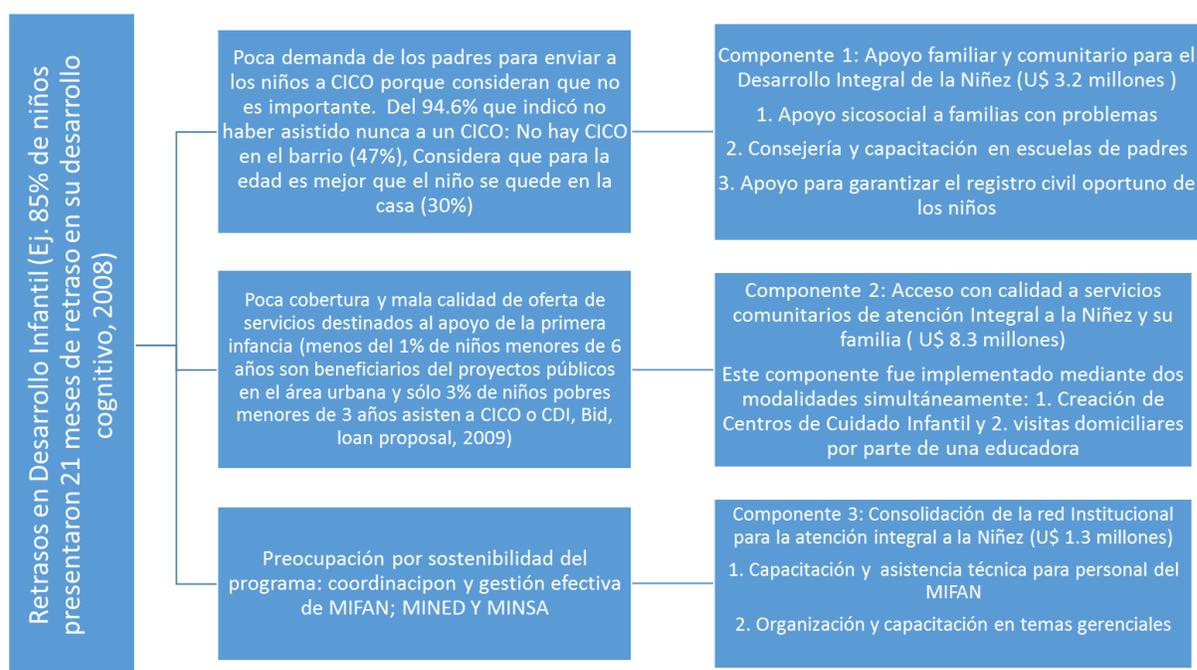
- i) reducir el porcentaje de niños con retrasos en su desarrollo global
- ii) aumentar la tasa de inscripción de niños en la escuela primaria.
- iii) reducir el porcentaje de niños menores de 6 años con desnutrición crónica.
- iv) aumentar el porcentaje de padres con conocimientos en cuidado infantil.

La información disponible para efectos de este documento, permite sólo evaluar el impacto del Programa en el desarrollo global del niño y no en los otros indicadores que responden a otras encuestas y datos administrativos que no fueron facilitados para este estudio.

b. ¿Cómo lo hace?

Para lograr su objetivo, el Programa intervino las comunidades elegidas mediante tres componentes (ver figura 1). El primero de ellos apuntó a fomentar buenas prácticas de crianza y aumentar los conocimientos de cuidado infantil responsable por parte de los padres, mediante la creación de escuelas de padres. El segundo, se desarrolló para mejorar la oferta de servicios infantiles mediante la construcción y atención de niños menores de 6 años en Centros de Cuidado Infantil Comunitario (CICO) o visitas domiciliarias por una educadora; y el tercer componente se creó para ofrecer las herramientas técnicas al MIFAM para implementar y monitorear el Programa mismo.

Figura 1. Diagnóstico y Componentes del Programa Urbano de Bienestar para la Niñez en Extrema Pobreza



Fuente: Elaboración propia basada en (BID, 2009)

En este documento se evaluará el impacto del acceso de los servicios de cuidado infantil, en el desarrollo integral de los niños menores de 6 años que se inscribieron y participaron de cualquiera de las dos modalidades ofrecidas por el Programa. Se elige este indicador por tres razones: i) es el fin último del programa, ii) este componente abarca dos tercios del presupuesto del programa y iii) sólo se cuenta con las encuestas que permiten evaluar este indicador. La primera de estas modalidades consistió en la construcción y el equipamiento de Centros Infantiles Comunitarios

(CICO) que atenderían a niños menores de 6 años en media jornada diaria durante 5 días a la semana. El centro educativo ofreció a los niños actividades contempladas en el Currículo Nacional orientado por Ministerio de Educación para pre-kinder, organizado en tres áreas de aprendizaje: socio-personal, comunicación y comprensión del mundo. Las actividades incluían juegos, cantos y lecturas. Adicionalmente, el Centro garantizaba, por medio del Programa, una merienda básica para los niños, con la opción de ser complementada por los padres de los niños participantes.

La segunda modalidad que ofreció el Programa consistió en la realización de visitas domiciliarias cada quince días, durante 45 minutos, por parte de una educadora de primera infancia para enseñarle a las madres o cuidadores del niño/a actividades educativas similares a las que se realizan en el CICO. Para dicho cometido, la educadora contaba con material gráfico demostrativo y utilizaba recursos del hogar como medios de apoyo para estimular al niño (ej. Ollas, tapas de botella, piedras, bombillas, etc.). Inicialmente, el Programa había considerado asignar a ciertas comunidades la modalidad CICO y a otras comunidades la modalidad de visitas domiciliarias; sin embargo, por problemas operativos con la legalización y construcción de los CICOs y presiones de la comunidad, se tuvo que iniciar el Programa indistintamente en todas las comunidades que habían resultado elegidas para ser intervenidas.

c. ¿Cómo se asignaron los beneficios?

En vista de que los beneficios del Programa no se podían negar a una fracción de los pobladores de un mismo barrio en el cual se construyera un Centro de Cuidado Infantil, se decidió ofrecer los servicios a ciertos barrios y a otros no, por medio de la aleatorización. Esto se realizó únicamente por razones éticas y para garantizar que los pobladores de los barrios no tratados aceptaran participar del estudio, con la posibilidad de que en un futuro los beneficios del Programa podrían extenderse a sus hogares. El proceso de aleatorización fue realizado por consultores externos al BID, quienes además prepararon el cálculo de la muestra necesaria para el levantamiento de encuestas (Gertler, Paul; Galiani, S; Sánchez, E & Vivo, S., Octubre 2009).

En un primer momento, se ordenaron y eligieron los municipios con las tasas más altas de extrema pobreza² y que concentraran hasta el cincuenta por ciento del total de los hogares urbanos en tal

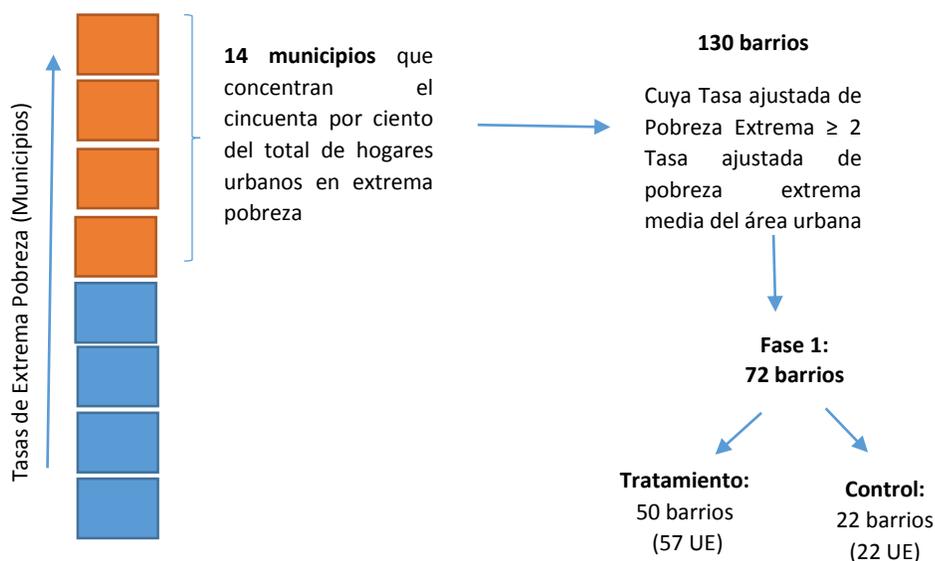
² Como indicador de Pobreza se utiliza el "Mapa de Pobreza Extrema Departamental" elaborado a partir de datos del VIII Censo Nacional de Población y IV de Vivienda 2005 y en base a la Metodología de Necesidades Básicas Insatisfechas (NBI).

condición socioeconómica. Bajo estos criterios, se seleccionaron 14 municipios. Luego se generó un índice por cada hogar y se agregó a nivel de barrio, indicando la proporción ponderada de las Necesidades Básicas Insatisfechas (NBI) de cada hogar. Los barrios cuya concentración ajustada de hogares en extrema pobreza fuese mayor o igual al doble del promedio urbano³ (del mismo índice) quedaron seleccionados para ser intervenidos por el Programa (130 barrios). Seguidamente se descartaron barrios que tras visitas en terreno no cumplieron con las características de localidades urbanas y quedó una sub muestra de 100 barrios con sus 17,956 familias y sus 16,540 niños menores a 6 años para ser intervenidos por el Programa, tanto en su fase 1 como en su fase 2. Dado que se encontró que algunos de dichos barrios presentaban dificultades operativas para ser intervenidos en el corto plazo; de la sub muestra de 100 barrios que iban quedando, se dejaron 72 barrios operativamente viables para ser intervenidos por el Programa. En este sentido, algunos de las restricciones operacionales presentadas por los barrios fueron limitaciones de acceso a agua y saneamiento para construir los CICOS o contextos con alta inseguridad ciudadana

Posteriormente, los 72 barrios elegidos fueron ordenados según los Segmentos Censales definidos en el Censo de Población del año 2005 y fueron designados aleatoriamente a Unidades Tratadas y No tratadas. Luego, dichas unidades fueron denominadas por el Programa como Unidades de Evaluación (UE), que en la mayoría de los casos corresponden exactamente a un barrio urbano y sólo en algunos casos, una sola UE puede estar conformada por dos barrios. Así pues, para esta fase, el Programa contempla para su estudio un total de 79 UE, ubicadas en 72 barrios de 14 municipios en 8 departamentos de Nicaragua. De éstas UE, 22 son de control y equivalen a 22 barrios; mientras que 57 son de tratamiento y equivalen a 50 barrios (ver gráfico 2). En vista de que, en esta fase, los recursos no eran suficientes para intervenir a toda la población, el Programa se propuso intervenir los barrios de control en su segunda fase.

³ Los barrios priorizados presentaron un desempeño inferior en estadísticas básicas de pobreza, salud y educación en relación al resto de la zona urbana; no así en relación a las estadísticas nacionales cuyos promedios empeoran debido al contexto de las zonas rurales del país.

Figura 2. Proceso de Selección de unidades tratadas y de control



Fuente: Elaboración propia basada en Gertler, P; Galiani, S; Sánchez, E y Vivo, S. (2009)

Los beneficios ofrecidos por el Programa son totalmente gratuitos y la participación en los CICOs o en las visitas domiciliarias no es obligatoria. Es decir, en los barrios intervenidos hay familias que participan del programa y otras que no lo hacen. El único requisito para beneficiarse del programa es presentarse al Centro Infantil con la cédula de identidad de los padres e inscribir al niño o inscribirse para recibir las visitas de la educadora.

3 Evidencia internacional: Impacto de intervenciones de apoyo a la primera infancia

Las autoridades gubernamentales están cada vez más interesadas en tener evidencia acerca de los impactos logrados por sus intervenciones públicas, es por ello que a través del continente se han propagado diversos estudios que buscan estimar la incidencia de dichos programas. Se ha demostrado que las inversiones en políticas de apoyo a la primera infancia son deseables y justificables en términos de beneficios a largo plazo para las sociedades.

Uno de los argumentos a favor de estas intervenciones, es que el 80 por ciento del cerebro se desarrolla en los primeros tres años de vida, y junto con esto se abre una “ventana de oportunidades” para potenciar las funciones ejecutivas y de autorregulación de los niños; por lo que garantizar un desarrollo integral en estos primeros años, es un prerrequisito para tener bienestar individual, productividad laboral y mejor manejo de situaciones emocionales en la vida adulta (Shokoff, Ritcher, Van Der Gaag, & Bhutta, 2012).

Luego, se ha encontrado que el desarrollo humano en la infancia provee una base relevante para mejorar los ingresos económicos futuros y reducir las desigualdades culturales y sociales de origen de los niños (Magnuson & Duncan, 2014). Finalmente, el laureado economista James Heckman ha destacado en diversos trabajos (Cunha, Heckman, Lochner & Masterov (2006), Heckman & Masterov (2007); Heckman (2000, 2008)) que políticas que promueven y apoyan la crianza en infantes, especialmente en familias vulnerables, reducen inequidad y son costo efectivas en el futuro. El autor destaca que las habilidades generan más habilidades (“skill begets skill), o, en otras palabras, que las inversiones tempranas mejoran las inversiones futuras; por lo que las sociedades deben invertir más en los más jóvenes y menos en adultos ya maduros y con bajas habilidades. Así pues, a largo plazo, la cadena de resultados de estos estudios implica que realizar intervenciones que promuevan el desarrollo integral de niños pequeños; en principio, mejorará sus habilidades motoras, lingüísticas, cognitivas y socio-emocionales; para en un futuro incidir en mejoras en aspectos académicos, económicos y personales de dichos niños.

En el corto y mediano plazo, se vuelve relevante entonces, evaluar el impacto directo de estas intervenciones en primera infancia en el desarrollo integral de los niños. En este sentido, los pocos

estudios que existen a nivel latinoamericano presentan evidencia variada, tanto en sus metodologías, como en sus resultados.

En Chile, por ejemplo, se realizó un estudio no experimental (Novoa-Hidalgo & Urzúa, 2012) con datos longitudinales y administrativos del impacto de participar en salas cunas públicas provistas por la Junta Nacional de Jardines Infantiles (JUNJI) y la Fundación Nacional para el Desarrollo Integral del Menor (INTEGRA). El estudio implementó como estrategias de evaluación estimaciones por medio de Mínimos Cuadrados Ordinarios, Variable Instrumental y Funciones de Control. Los autores encontraron mejoras significativas en el área motora (fina y gruesa); habilidades de expresión personal (expresión de las emociones) y hábitos adaptativos en la alimentación. Sin embargo, también encontraron efectos negativos en los procesos de interacción entre el niño y el adulto. La investigación confirma la existencia de beneficios de la educación temprana en aspectos de desarrollo motor y de auto-regulación de los niños; pero destaca la importancia de garantizar la calidad de los procesos en dichas iniciativas para evitar efectos negativos no deseados, especialmente cuando se prioriza programas que expanden cobertura.

En Colombia, se realizó una evaluación de impacto cuasi-experimental entre participantes y potenciales participantes del Programa de Hogares Comunitarios de Bienestar (HCB), para medir los efectos en el estado nutricional, de salud, desarrollo cognitivo y psicosocial de los niños (Bernal, y otros, 2009). Los HCB consisten en hogares en los que se cuidan a niños menores de 6 años que viven en condiciones de pobreza y vulnerabilidad. En este programa los niños están a cargo de una madre escogida por la comunidad que debe tener al menos 10 años de nivel educativo y es certificada y capacitada por las autoridades pertinentes.

El programa ofrece cuidado infantil, complemento alimentario, actividades de desarrollo psicosocial y promoción de la salud. El estudio realiza una evaluación del programa al comparar niños tratados con niños no tratados por medio del método del emparejamiento; así como una evaluación de la intensidad del programa a través de la comparación de niños tratados por más tiempo en el programa, con niños tratados por menos tiempo. En el área de desarrollo psicosocial, la evaluación del programa reflejó una diferencia de 0,011 en el índice de desarrollo socio-emocional (aproximadamente 0,04 de una desviación estándar), a favor de los niños tratados. Sin embargo, al realizar el análisis de intensidad se encontró que los niños más expuestos al programa empeoran

sus comportamientos agresivos. Este último hallazgo, según los autores, está condicionado a la calidad del centro de cuidado al que asiste el niño.

Por otro lado, en el aspecto de desarrollo cognitivo, el estudio encontró diferencias en contra del programa en todos sus indicadores medidos; aunque el análisis de intensidad, arrojó que existen efectos positivos sobre el desarrollo cognitivo de los niños que participaron al menos durante 15 meses en el programa. Finalmente, los investigadores analizaron el impacto de los HCB en el desarrollo cognitivo de los niños en el mediano plazo comparando los resultados de la prueba cognitiva SABER aplicada en el 5to grado de primaria. Los autores encuentran un efecto positivo y significativo de 14.5 puntos porcentuales en dicha prueba.

En Bolivia, la evaluación de impacto del Proyecto Integral de Desarrollo Infantil (PIDI) brinda evidencia que programas de cuidado infantil incrementan los puntajes de habilidades cognitivas y sicosociales para aquellos niños que participan en al menos 7 meses. El programa evaluado ofrece cuidado infantil durante el día, servicios de nutrición y educación para niños pobres, entre los 6 y 72 meses, que residen en áreas urbanas. El resultado de la evaluación mediante métodos no paramétricos evidenció que las estimaciones del impacto dependen altamente de la edad del niño y del tiempo de exposición al programa (Behrman, Cheng, & Todd, 2004).

En Ecuador, por su lado, se evaluó el impacto de visitas domiciliarias semanales durante una hora y de servicios de cuidado infantil diarios en un centro institucional, mediante el uso de regresión discontinua tomando como umbral el puntaje que recibe el proveedor del servicio para recibir o no fondos nacionales. El estudio encontró que las visitas domiciliarias tiene un impacto positivo en el desarrollo cognitivo y motor de los niños; contrario a las salas cunas que reflejaron impacto nulo o incluso negativo en las mismas variables de interés (Rosero & Oosterbeek, 2011). Las visitas domiciliarias, además mostraron un efecto positivo en la reducción de depresión y el estrés de la madre y mejora en la conciencia de maternidad responsable; mientras que los centros infantiles comunitarios, lograron lo opuesto. Los resultados deben ser considerados a la luz de la premisa de que el impacto deseado en el desarrollo de los niños depende de la cantidad y la calidad del tiempo invertido con la madre o con el programa.

Otro estudio en Jamaica (Gertler, Paul; Galiani, S; Sánchez, E & Vivo, S., Octubre 2009), evidencia que las visitas domiciliarias a niños vulnerables que viven en áreas urbanas resulta efectivo para mejorar los indicadores de salud y desarrollo psicoemocional (Griffiths Mental Development Scales) en niños menores de 30 meses, en la medida en la que se incrementa el grado de periodicidad de las visitas. Este estudio se realizó mediante un diseño experimental en el cual se asignó aleatoriamente diferentes tratamientos: un grupo visitas domiciliarias una vez al mes, otro grupo visitas domiciliarias dos veces al mes y otro grupo de control que no recibió visitas. El grupo con visitas más seguidas mostró efectos pequeños pero significativos en el aumento de la escala Griffiths de Desarrollo Mental; mientras que el grupo que recibió visitas mensuales no mostró ninguna mejora en las escalas.

Baker-Henningham y López-Boo (2013) comparan la efectividad de más de 30 programas orientados a la estimulación del desarrollo infantil (servicios de consejería, visitas domiciliarias y centros de cuidado infantil) en niños entre 0 y 3 años, implementados en países de ingresos bajos y medios, y concluyen que las intervenciones que tuvieron impactos positivos fueron aquellas que presentaron mayor intensidad y duración. Adicionalmente, aquellas intervenciones que mejoraron su efectividad estaban orientadas a niños que viven en entornos desfavorables, con edades más tempranas y cuyas madres y familias fueron integradas en los programas. Así, pues los autores encontraron que las intervenciones de estimulación temprana, pueden resultar positivas en el desarrollo mental de los niños y en las pautas de crianza de las madres en el corto plazo. Los investigadores concluyen que aún queda pendiente la tarea de determinar la edad óptima para intervenir a los niños y las modalidades más costo-efectivas para lograr los resultados esperados.

De la revisión de la literatura, se puede concluir que los estudios de evaluaciones de impacto para Latinoamérica muestran resultados diversos en cuanto al efecto de dichas intervenciones en el desarrollo de los niños; pero convergen en destacar la importancia de complementar la evaluación del programa sobre indicadores del niño con la evaluación de la calidad de los servicios ofrecidos por las intervenciones; sobretodo aquella concerniente a la calidad los procesos; de lo contrario, los esfuerzos que se llevan a cabo con dichos programas son limitados.

Adicionalmente, un factor determinante en la efectividad de los programas es el grado de exposición de los niños a dichas intervenciones y la efectiva integración de las familias en los programas. Por supuesto esto está sujeto a las limitaciones con la que cuentan los gobiernos en cuanto a disponibilidad de recursos y capacidad para monitorear y evaluar la calidad en dicho sentido.

4 Metodología

Toda evaluación de Impacto busca determinar si la intervención realizada tuvo los efectos esperados en la variable de interés en la que se quería incidir; para el caso particular del Programa Urbano, y para este estudio, se desea estimar si el Programa tuvo efecto en reducir los problemas en el desarrollo de los niños que participan de la atención de cuidado infantil institucional o itinerante.

La evaluación de impacto desea analizar los cambios en la variable de interés, en este caso la probabilidad de que un niño tenga problemas en su desarrollo o no, que son realmente atribuibles a la intervención y no a otras variables que puedan incidir en el resultado. Así pues, las evaluaciones de impacto buscan efectos causales y atribuibles a un programa. El problema básico de una evaluación es que no se puede contar con un mismo niño que participe o no del programa simultáneamente; el niño puede participar o no del programa; pero nunca se encontrará en ambos escenarios al mismo tiempo; esto es lo que se denomina “el dilema del contrafactual”. Para poder estimar el efecto causal de un programa es necesario encontrar un grupo válido de comparación o control lo más parecido al grupo intervenido, tanto en sus características observables como no observables, y así poder estimar lo que habría ocurrido con los participantes del Programa en ausencia del mismo (Gertler, Martínez, Premand, Rawlings, & Veermeersch, 2011).

Tomando en consideración el proceso de asignación del programa, y también como una manera de revisar la robustez de los resultados encontrados, se realizarán diferentes estrategias de evaluación, las que se comentan a continuación:

Estrategia 1: Aleatorización del Programa

El sistema de asignación de los beneficios del Programa Urbano corresponde a un diseño experimental en el que se “aleatoriza en fases”; esto significa que un grupo de individuos igualmente elegibles para ser beneficiarios del programa recibirán los beneficios del mismo en distintos momentos del tiempo: primero, un grupo de niños será tratados, mientras otro grupo esperara a ser beneficiado del Programa en su siguiente fase. Este enfoque de aleatorización ayuda a lidiar con temas éticos en cuanto a la decisión de a quien excluir y a quién beneficiar, siendo que todos los pobladores de los barrios tienen la misma necesidad de recibir los servicios ofrecidos por el Programa. Esto además incrementa la probabilidad de que áreas de control y áreas intervenidas

sean similares en características observables y no observables, siendo que la aleatorización equilibra los sesgos y/o diferencias que puedan existir entre los pobladores de un barrio y otro (Khandker, Koolwal, & Samad, 2010).

En este caso, el grupo de control estará compuesto por niños que habitan en barrios donde el programa no fue implementado. La aleatorización a este nivel permite que los individuos que habitan en los barrios (UE) de control y que son igualmente elegibles⁴ para beneficiarse del Programa no participen en el mismo sólo por el hecho de que los servicios no fueron ofrecidos en sus barrios; siendo que dicha decisión es exógena a los habitantes del barrio, esto los convierte en un contrafactual adecuado. Por otro lado, el grupo de tratamiento estará compuesto por los niños que pertenecen a barrios en donde se ofrecieron los servicios del Programa y que participan del Programa.

La aleatorización de los servicios a los barrios, en teoría, permite que el tratamiento esté distribuido de manera independiente de otros determinantes de impactos y asegura que tanto la variable de resultado como otras características sean estadísticamente iguales entre los grupos de tratamiento y de control; por lo tanto, la primera estrategia para estimar el impacto será aprovechar la manera en cómo se implementó el programa, y medir el “Impacto Promedio del Tratamiento en los Tratados”. Esto se realiza mediante el siguiente modelo probit:

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

Donde, Y_i toma valores de 1 si el niño tiene problemas en su desarrollo y 0 si no los tiene, D_i es el indicador del tratamiento que toma valores de 1 si el niño participa del programa siendo que su barrio fue intervenido y 0 si el niño no es tratado y u_i , que es el término de error y que incluye determinantes no observados en los datos y que pueden influir en el desarrollo de los niños. En este caso el coeficiente β_1 recoge el impacto del programa ya que si el i -ésimo niño participa del Programa ($D_i=1$), entonces su probabilidad de tener problemas en su desarrollo será $\beta_0 + \beta_1 + u_i$. Si no participa ($D_i=0$), su probabilidad de tener problemas en su desarrollo global será $\beta_0 + u_i$.

⁴ Recuérdese que los controles son tan pobres y vulnerables como los beneficiarios según el indicador por el que fueron seleccionados.

Así pues, el resultado de dicha diferencia en la probabilidad de que un niño tenga problemas en su desarrollo, entre un niño intervenido y uno no intervenido, es β_1 .

Sin embargo, pueden existir otras variables que pudiesen estar correlacionadas con la variable de resultado además del tratamiento, y que presentan diferencias estadísticas entre ambos grupos. Si no se controla por estas variables el estimador del impacto no sería insesgado; por lo tanto, otra manera de estimar el impacto será controlar por estas eventuales diferencias estimando la siguiente ecuación:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + u_i$$

Donde, adicional a las variables ya señaladas, el vector X representa un set de características que inciden en el desarrollo del niño y que presentan diferencias que son estadísticamente significativas entre los grupos tratados y control.

Estrategia 2: Diseño cuasi-experimental: Emparejamiento

Adicionalmente, a la estimación anterior y como ejercicio para analizar la robustez de los resultados, se estimará el impacto del Programa a través del método del emparejamiento. Esta metodología permite simular un diseño experimental entre los que participan en el programa y los que no participan en el programa. Se elige esta metodología para revisar la robustez de los resultados anteriores y para equilibrar las diferencias que persistan en los grupos tratados y controles, debido al “sesgo de selección” que se observa en base a las características observables. Así pues, el “Propensity Score” es la probabilidad condicional de un tratamiento particular dado un vector de covariables observables (Rosenbaum & Rubin, 1984). Se utiliza esta metodología cuando se sospecha que la estimación del impacto está “sesgada” debido a que se observan diferencias sistemáticas entre un grupo tratado y uno no tratado, y una manera para controlar dicho sesgo, y que ambos grupos sean estadísticamente comparables, es elegir a las unidades tratadas y controles agrupando dichas unidades en subgrupos, basados en características observables, y luego

comparando sólo tratados y controles que resultan “similares” o que se encuentran en el mismo subgrupo.

La ventaja del método de emparejamiento es que reduce la dimensionalidad de los observables condicionantes del tratamiento a un solo escalar: $P(x)$; por lo tanto, el estimador por medio de PSM es simplemente la diferencia media en las variables de resultado entre el grupo de tratamiento y el grupo de control en el soporte común, ponderada por la distribución de la probabilidad de participación (o propensity score) en el programa de los participantes (Bernal & Peña, 2011). Así, pues, se puede usar Propensity Score Matching (PSM) para emparejar participantes con miembros en el grupo de control similares en las mismas características observables por las que se autoseleccionan para participar, es decir, se pueden comparar los niños de las madres que eligieron enviar a sus hijos con los niños de las madres que enviarían a sus hijos, si tuviesen el servicio en su barrio. Bajo los supuestos de la metodología del PSM, es posible estimar los efectos de la intervención del Programa, sólo comparando las diferencias de los resultados entre participantes y no participantes observacionalmente similares.

Es importante señalar que uno de los temas relevantes para la correcta implementación del PSM es modelar o al menos justificar bajo que consideraciones se están asumiendo similares los individuos a compararse, por lo que deben tomarse en cuenta un buen rango de variables por las que los participantes y no participantes puedan diferir en aras de mitigar el sesgo potencial. Por lo tanto, una de las primeras condiciones para poder aplicar el PSM es disponer de suficientes características observables que permitan estimar el proceso de autoselección. Para tales efectos, este Programa cuenta con una rica base de datos de la Línea Base con preguntas de las características de la madre de los niños, tanto educacionales como de estados de ánimo, y siendo que la decisión de enviar o no al niño al CICO recae principalmente en la madre, esto brinda una buena oportunidad para estimar el Propensity Score. Adicionalmente, se cuenta con características socioeconómicas del hogar y características del niño participante.

La segunda condición para poder implementar PSM es que la probabilidad de encontrar un individuo en el grupo de control para cada valor de X debe ser positiva, esto es que pueda existir un traslape en las densidades de las probabilidades de participación de los individuos participantes y no participantes del Programa, o en otras palabras que debe existir deben existir pseudo-controles con

un P-Score similar al de los tratados, o tener niños en ambos grupos con el mismo valor del P-Score. El Programa posee una muestra rica en variables y en observaciones, siendo que se cuenta con la información con más de 500 niños y hogares en condiciones pre-tratamiento, lo cual representa una ventaja para la estimación, pues se cuenta con suficiente densidad en la muestra para mejorar la precisión del estimador.

Finalmente, se eligen como algoritmos para estimar el PSM, al Vecino Más Cercano con reemplazo y al emparejamiento por distancia máxima (radius matching); sólo para comparar la sensibilidad de las estimaciones y para comparar los resultados en base a los mejores niños no tratados de los que se disponga.

5 Los Datos

a. Descripción de encuestas

Para efectos de capturar información sociodemográfica que permitiera realizar un análisis más completo del impacto del Programa, el BID contrató al Centro de Investigación y Acción educativa Social (CIAS) para realizar dos levantamientos de información: el primero corresponde a la Línea Base y el segundo, al levantamiento de los datos de seguimiento. La Línea Base cubrió a 3,691 niños de 3,300 hogares en los 79 barrios urbanos elegibles y fue levantada en noviembre del 2013; antes de la implementación del Programa, mientras que los datos de seguimiento cubrieron a 2,962 niños de 2705 hogares y fue levantada, correspondientemente, en noviembre del 2015.

La estrategia de selección de los niños que formarían la línea base, y más tarde la base de seguimiento, siguió dos esquemas: uno para las UE con las que se contaba con información del CENSO 2005, y a partir de la cual se hizo una selección aleatoria de 50 hogares con niños menores de 6 años (40 hogares a elegir y 10 de reposición); y otra para las UE de las que se carecía de listado de hogares. En este segundo grupo se dividieron las UE en otros dos subgrupos: i. Aquellas UE que tenían menos de 200 viviendas y ii. Aquellas UE que tenían más de 200 viviendas. Para el primer subgrupo (i), se levantó un mini-censo de la UE de los hogares con niños menores de 6 años, se contabilizaron los hogares y se eligieron 50 de ellos aleatoriamente; para el segundo subgrupo (ii), se realizaron visitas hogar por hogar, primero partiendo desde lo más cercano del CICO hasta llegar a tener 20 hogares y 22 niños elegidos y luego, partiendo desde lo más lejos del CICO hasta obtener 20 hogares y 22 niños (CIAS, Febrero, 2014).

Se implementaron tres encuestas adicionales al instrumento de la medición de la variable de interés para contar con información del hogar, de la madre/cuidadora principal del niño y de información antropométrica del niño que se detallan a continuación (CIAS, Septiembre 2015):

1. Encuesta de Hogares: Este instrumento está conformada por 5 secciones. La primera permite identificar el hogar, el control de visitas realizadas y la especificación geo-referencial del hogar; la segunda brinda información de las características físicas de la vivienda como los materiales de los que está hecha, el acceso a servicios básicos o el número de cuartos. La siguiente sección levanta información acerca de las características y la composición de los miembros del hogar como su parentesco con el jefe del hogar, sus edades o la existencia de condiciones de discapacidad o gravidez; la parte 4 del cuestionario cubre información del nivel de educación y la asistencia escolar para los miembros del hogar mayores a 3 años; finalmente, la última

sección recoge características de la situación laboral de los miembros del hogar como son su estado ocupacional, la actividad principal y el sector económico donde se labora.

2. Encuesta de madre/cuidadora principal: Esta encuesta fue diseñada para realizarla a la madre o cuidadora principal del niño para coleccionar información acerca de pautas de cuidado y crianza con el niño; características reproductivas y anímicas de la madre; conocimientos parentales, prácticas del embarazo y parto; así como condiciones anímicas y perceptivas del niño. El cuestionario tiene 10 secciones: i. Fecundidad de la madre biológica; ii. Inmunizaciones de los niños; iii. Lactancia y alimentación; iv. Programas de atención a la primera infancia; v. Conocimiento sobre el Desarrollo Infantil; vi. Pautas de Crianza; vii. Estado de ánimo de la madre; viii. Valoración del niño; ix. Intensidad de talleres y acompañamiento para la madre y x. Relaciones familiares como una aproximación para capturar temas de violencia familiar en el hogar.

3. Medición Antropométrica: Este cuestionario captura las mediciones de peso, talla y longitud de los niños y percepciones de la madre de cómo encuentran a su hijo en relación a otros niños de su edad en relación a dichas medidas antropométricas.

b. Instrumento de Medición de Impacto

Como instrumento de medición de resultados se utiliza el cuestionario CLAP de Desarrollo Infantil. Esta prueba evalúa el desempeño del niño en diversas tareas adecuadas para su edad, por lo que se realizan cuestionarios adaptados a rangos de edad que tiene el niño y a las tareas que éstos deberían estar realizando (ver anexo 1). El instrumento CLAP fue seleccionado para la evaluación de resultados porque requiere relativamente poco tiempo para implementarse (15-20 minutos por niño/a) y sobre todo porque es el instrumento que el MIFAN aplica en sus centros de cuidado infantil (incluyendo los CICO) para dar seguimiento al desarrollo de los niños/as.

Este instrumento permite medir el desarrollo del niño en niveles: 1. Alerta en el Desarrollo del niño, 2. Desarrollo en proceso y 3. Desarrollo Global Normal. Los criterios para clasificar al niño en cada una de las categorías se describen a continuación:

1. Alerta: Si el niño presenta dos o más alertas de desarrollo en cualquier área evaluada (área social, área motora, área lenguaje y área de coordinación)
2. En proceso: Si el niño presenta al menos una alerta de desarrollo en cualquier área evaluada
3. Normal: si el niño presenta tres áreas normales (esto sería la mayoría de sus evaluaciones normales), ninguna área completa en alerta o todas las áreas evaluadas normales.

Para efectos de equilibrar los porcentajes de los niños según sus niveles de desarrollo y para simplificación de este instrumento, se reagruparon las evaluaciones de los niños en dos subgrupos en una variable dicotómica que toma el valor de 1 “si el niño tiene problemas en su desarrollo” y 0 “si el niño tiene un desarrollo normal”; esto es equivalente a agrupar los niveles 1 y 2 explicados anteriormente en “niños con problemas en su desarrollo” y dejar a los niños que tuvieron resultados normales en un nivel por sí mismo, equivalente a un valor de “0”.

La aplicación del instrumento requirió capacitación especializada por una psicóloga especializada en desarrollo infantil, en vista de que requiere trabajo especializado con menores de edad. La utilización del Instrumento CLAP se vio afectada por un problema ligado a la variación en los métodos de capacitación de sus administradores. Aunque la misma organización fue contratada para llevar a cabo el trabajo de campo, tanto la línea de base como para la evaluación final – lo cual se consideró un elemento a favor de la validez interna de las mediciones – se dieron problemas dado que se cambió a la capacitadora. La firma consultora contrató en 2013 a una persona que entrenó con menos rigor a los aplicadores, y en cambio la capacitadora de 2015 era más estricta. Esto provocó diferencias en la administración de algunos ítems de CLAP, así como de sus resultados y esto no permitió la comparabilidad de ambos cuestionarios, como se contemplaba realizar en el diseño experimental de esta evaluación (Castro & Laguna, 2016).

c. Estadística Descriptiva pre-tratamiento

El Programa inició a brindar sus servicios a inicios del 2014 y los datos de Línea Base se levantaron en noviembre del 2013; mientras que las encuestas de seguimiento se levantaron en noviembre del 2015. En la Línea Base se entrevistaron a 3691 niños y la muestra de seguimiento contiene información de 2962 niños, lo que implica que se tuvo una pérdida de muestra de 729 niños, equivalente a una tasa de atrición de 19.7%.

Se eliminaron de esta muestra algunos casos por consideraciones particulares. Primero, se detectó contaminación de la muestra de control pues hubo niños que participaron del otros programas parecidos, por lo que se procedió a eliminarlos de la muestra, quedando información de 2856 niños. Segundo, hubo No respuesta al ítem de la variable de resultado en 66 casos. Siendo que la tasa de no respuesta es baja (2.3%), se prefirió eliminar a estos casos. Probablemente esta tasa de No respuesta al ítem se debió a que los niños no se encontraron en buen estado de ánimo para contestar la evaluación en ninguno de los casos en los que se les visitó. Tercero, para el 2015, se tuvo que suspender los servicios ofrecidos por el Programa en el Barrio Rubén Darío debido a que

éste fue declarado como zona de riesgo luego de un temporal de fuertes lluvias en el país. Se procedió entonces a eliminar de la muestra a los niños que formaban parte de esa Unidad de Tratamiento que, aunque fueron ubicados, experimentaron circunstancias extraordinarias. Finalmente, se cuenta con información de 2758 niños: 2040 en barrios intervenidos y 718 en barrios no intervenidos.

Los anexos 2 y 3 muestran estadística descriptiva y test de medias para evaluar si la aleatorización a nivel de conglomerados (barrios) funcionó y si los grupos de control son estadísticamente parecidos o no. Se observa que tanto en la variable de resultado pre-tratamiento como en la información de los niños, no existen diferencias significativas entre barrios intervenidos y no intervenidos. Sin embargo y, a pesar de la aleatorización, existen diferencias significativas en ciertas características socioeconómicas entre los barrios tratados y controles. Asimismo, existen diferencias significativas en la edad promedio a la que la madre dio a luz, el nivel educativo de las madres y el número de hijos vivos que tienen. Tales diferencias hacen sospechar que la aleatorización a nivel de barrios pudo no ser efectiva por dos posibles razones: primero, la aleatorización se hizo en base a información estadística muy rezagada (Censo 2005) por lo que las condiciones socioeconómicas de los barrios pudieron haber cambiado en el tiempo entre que se diseñó la asignación del tratamiento y el período que se terminó ejecutando el Programa. Cabe señalar que el principal problema para el retraso en la implementación de los beneficios se debió al lento proceso de compra y legislación de los terrenos para la construcción de los CICOs. Otra posible causa por la que se observan diferencias entre ambos grupos, puede deberse a que no haya respetado el protocolo de aleatorización y/o hubo poco rigor en la supervisión del mismo desde el inicio. Esta situación es un problema latente en toda aleatorización.

Como se sabe, el diseño de oferta aleatoria implica que de los 78 UE elegibles para ser beneficiadas del programa, en 22 de ellas no se impartieron ninguno de los servicios ofrecidos por el programa, lo que implica que el tratamiento es exógeno a los habitantes de estos barrios; mientras que al resto de pobladores se le ofreció los servicios del programa a costo nulo, con suficiente capacidad para atender a los niños de los barrios y sin condicionantes de exclusión. Sin embargo, en los barrios intervenidos, de los 2758 niños que podrían haber participado del programa, sólo participaron 716 niños. Una de las principales razones de este bajo take-up se debe que los servicios de cuidado infantil sólo se ofrecieron por medio día, condición que no responde a las necesidades reales de las madres trabajadoras.

La tabla 1 brinda estadística descriptiva de la variable de resultado para la muestra de niños que participan del Programa y para aquellos que no participan por encontrarse en barrios no intervenidos. Se puede observar que pre-tratamiento existen diferencias significativas en los porcentajes de niños con problemas en su desarrollo, siendo mayor el porcentaje de niños con problemas en su desarrollo en el grupo de tratamiento. Adicionalmente, se detectó que el nivel estimado de niños con problemas en su desarrollo para el año 2013 – esto es alrededor del 10.73% – no es consecuente con el porcentaje de niños que presentan retrasos en su desarrollo de acuerdo a diversos estudios realizados en el país. En este sentido, el estudio de PRIDI evidencia que el 25% de los niños de Nicaragua se desempeña entre 5 y 4 desviaciones estándar por debajo del promedio en la escala Engle en las áreas socioemocional, cognitiva, motora y de vocabulario (BID, 2015). Adicionalmente, resultados preliminares de un estudio del BID, evidenciaban que para el 2008, el 85% de niños nicaragüenses de 5 años pertenecientes a comunidades rurales pobres mostraban un retraso de 21 meses en su desarrollo cognitivo (Shady, y otros, 2014). De manera general, los niños nicaragüenses en promedio se desempeñan por debajo de países como Paraguay y Perú en las dimensiones cognitiva, motricidad, de lenguaje y comunicación.

Tabla 1. Variable de impacto: Estadística descriptiva y test de medias según niños que participaron en el programa vs niños que no participaron (2013)

Variable de Impacto: Niños con problemas en su desarrollo (en alerta y proceso=1, desarrollo normal=0)				
VARIABLE	Grupo de tratamiento	Grupo de control	Diferencia de medias	Total de la muestra
Porcentaje de niños con problemas en su desarrollo global	12.43	9.05	3.38**	10.73
Edad promedio de los niños (en meses)	20.65	21.54	-0.89	21.10
Observaciones (Número de niños)	716	718		1434

***Diferencias significativas $p < 0.01$, ** Diferencias significativas $p < 0.05$, * Diferencias significativas $p < 0.10$

Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013

Adicionalmente, se observa que la edad promedio de los niños que son considerados en el estudio tiene alrededor de 2 años (aproximadamente 21 meses) y que no existen diferencias significativas entre los niños que participaron del programa y los que no lo hicieron.

Por otro lado, la tabla 2 presenta estadística descriptiva de las madres de los niños que participan en el estudio. Tales características son fundamentales tanto en el desarrollo de los niños como en la probabilidad de participar en el programa; pues la figura materna es generalmente el principal referente del niño por ser quien cuida de él, le brinda seguridad y toma la mayoría de las decisiones relacionadas con su cuidado. Se observa, que a excepción de la edad a la que la madre dio a luz, el número de hijos vivos actualmente y si la madre tiene estudios técnicos o superiores o no, no existen diferencias significativas entre las madres de los niños que participaron y las de los que no participaron. En este sentido, las mamás en el grupo de control son claramente más educadas que el grupo de tratamiento.

De manera general y en el contexto de barrios y hogares pobres, se observa que las madres son bastante jóvenes, tienen alrededor de 27 años de edad y fueron madres por primera vez a los 19 años; además tienen muy bajo nivel educativo: tan sólo el 13.9% de la muestra tiene educación técnica o superior. Esto no se escapa de la realidad que viven muchos hogares en extrema pobreza en Nicaragua.

Tabla 2. Características de las madres: Estadística descriptiva y test de medias según niños que participaron en el programa vs niños que no participaron (2013)

VARIABLE	Características de la madre			
	Grupo de tratamiento	Grupo de control	Diferencia de medias	Total de la muestra
Edad de la madre (en años)	26.89	26.59	0.30	26.74
Edad promedio de la madre biológica al dar a luz (en años)	18.77	19.46	-0.69***	19.11
Número de controles prenatales de la madre durante el embarazo	7.12	7.29	-0.17	7.20
Número de hijos vivos en la actualidad	2.51	2.12	0.38***	2.32
Si la madre tiene educación técnica o universitaria (1=sí)	9.4	18.7	-9.27***	13.9

VARIABLE	Características de la madre			Total de la muestra
	Grupo de tratamiento	Grupo de control	Diferencia de medias	
Si la madre no ha dormido bien en la última semana (1=sí)	43.15	40.66	2.48	41.91
Si el niño colorea en casa (1=sí)	56.16	55.58	0.57	55.87
Si la madre dio lactancia exclusiva al niño (1=sí)	82.78	81.33	1.44	82.06
Promedio de horas de juego con su hijo al día	1.78	1.95	0.16	1.86
Le lee a su hijo alguna vez (1=sí)	84.9	86.1	-1.16	85.5
Observaciones (Número de niños)	716	718		1434

***Diferencias significativas $p < 0.01$, ** Diferencias significativas $p < 0.05$, * Diferencias significativas $p < 0.10$

Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013

La tabla 3 muestra las características de los hogares de los niños que son considerados en este estudio. De manera general, se evidencia que los hogares intervenidos tienen más privaciones económicas en cuanto a infraestructura del hogar se refiere. El 65.16% de los hogares tratados tienen una letrina como tipo de servicio sanitario y el 49.47%, tienen piso de tierra. A pesar de la aleatorización, existen diferencias significativas en características socioeconómicas en los hogares de los niños que participan del Programa y aquellos hogares a los que no se les ofrecieron los servicios de Programa. Nuevamente, ante dichas estadísticas, se sospecha que la aleatorización pudo haber tenido fallas o en su diseño o en su implementación, dado que si el criterio de asignación del tratamiento fue los niveles de pobreza, los barrios no deberían presentar tantas diferencias en aspectos socioeconómicos.

Tabla 3. Características socioeconómicas de los hogares: Estadística descriptiva y test de medias según niños que participaron en el programa vs niños que no participaron (2013)

VARIABLE	Características socioeconómicas de los hogares			
	Grupo de tratamiento (716)	Grupo de control (718)	Diferencia de medias	Total (1434)
Número de miembros en el hogar	5.49	5.61	-0.11	5.55
Número de cuartos de los que dispone el hogar	2.13	2.57	-0.43***	2.35
Porcentaje de hogares con piso de tierra	51.25	37.37	13.8***	44.31
Porcentaje de hogares que tienen letrina	67.31	43.87	23.44***	55.57
Porcentaje de hogares con paredes de concreto	40.92	51.46	10.54***	46.19
Porcentaje de hogares que usan leña como combustible	44.55	34.77	9.77***	39.66

***Diferencias significativas $p < 0.01$, ** Diferencias significativas $p < 0.05$, * Diferencias significativas $p < 0.10$

Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013

6 Resultados

Para analizar el efecto del Programa en su objetivo de reducir el porcentaje de niños con problemas en su desarrollo, se estima primero el “Efecto del Tratamiento en los Tratados” mediante simple diferencia de medias. Luego, se controla por variables que pueden incidir en la variable de resultado del niño para tener una mejor precisión de la estimación. Como se observa que existen diferencias significativas en características observables de los grupos, se reestima el impacto del Programa mediante el método de emparejamiento, lo que permite corregir el sesgo de selección, equilibrando las diferencias en variables observables entre ambos grupos.

a. Estrategia 1: Aleatorización del Programa

La tabla muestra el impacto del Programa en la reducción de porcentajes de niños con problemas en su desarrollo. En primer lugar, la columna 1, estimada como simple diferencia de medias entre niños participantes y niños no participantes, no señala impacto estadísticamente significativo. Como se sabe que existen otras variables que pueden incidir en el desarrollo de los niños y que se observan diferencias en características observables de los grupos, se añaden controles que permitan mejorar la precisión de la estimación. La columna 2 incluye, adicional al tratamiento, características del niño y de la madre; la columna 3, características del jefe de familia y la columna 4, características del hogar como proxy de nivel socioeconómico en el que se desarrollan los niños.

Controlando por distintas características, se encuentra que existe un impacto del programa en los tratados que oscila entre 5 y 7 puntos porcentuales; es decir, el porcentaje de niños con problemas en su desarrollo se redujo en alrededor de 7 puntos porcentuales. Es importante que en este escenario el grupo de control corresponde a niños a los que no se les ofreció el programa por residir en barrios no intervenidos, por lo tanto el tratamiento es exógeno a ellos.

Tabla 5. Estimación del Impacto del Tratamiento en los Tratados

	(1)	(2)	(3)	(4)
	mfx1	Mfx2	Mfx3	Mfx4
VARIABLES	Prob_Desarrollo	Prob_Desarrollo	Prob_Desarrollo	Prob_Desarrollo
Tratados_efectivos	-0.0318 (0.0289)	-0.0556* (0.0296)	-0.0538* (0.0297)	-0.0657** (0.0307)
Mal_global_13		0.0678 (0.0487)	0.0705 (0.0491)	0.0643 (0.0490)
edad_13		-0.0029**	-0.0028**	-0.0030**

	(1)	(2)	(3)	(4)
	mfx1	Mfx2	Mfx3	Mfx4
VARIABLES	Prob_Desarrollo	Prob_Desarrollo	Prob_Desarrollo	Prob_Desarrollo
		(0.0012)	(0.0012)	(0.0012)
lectura		-0.0895**	-0.0901**	-0.0855**
		(0.0428)	(0.0430)	(0.0429)
edad_luz_madre		-0.0045	-0.0048	-0.0041
		(0.0043)	(0.0044)	(0.0044)
niveleducativo_mom		-0.0177*	-0.0137	-0.0106
		(0.0092)	(0.0093)	(0.0095)
mama_sin_dormir		-0.0569*	-0.0587**	-0.0584**
		(0.0292)	(0.0292)	(0.0293)
hijos_vivos		0.0161	0.0154	0.0145
		(0.0098)	(0.0097)	(0.0097)
educ_jefe			-0.0176**	-0.0171**
			(0.0078)	(0.0078)
jefe_trabaja			-0.0372	-0.0425
			(0.0366)	(0.0370)
cuartos_hogar				-0.0214*
				(0.0124)
piso_tierra				0.0066
				(0.0370)
paredes_concreto				0.0100
				(0.0364)
Letrina				0.0228
				(0.0327)
Observaciones	1.024	1,024	1,024	1,024
r2_p	0.0010	0.0243	0.0293	0.0326
chi2	1.21	30.34	36.15	39.86
P	0.2711	0.0002	0.0001	0.0003

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Fuente: Elaboración propia basada en encuesta realizada en noviembre del año 2013 y variable de resultado levantada en noviembre del año 2015

b. Estrategia 2: emparejamiento

De los resultados anteriores se puede concluir que el Programa ha tenido un impacto en los tratados de entre 5 y 7 puntos porcentuales en la reducción del porcentaje de niños con problemas en su desarrollo

Como estrategia alternativa para verificar la robustez de estos hallazgos y para controlar las diferencias encontradas en características observables entre los grupos de niños inscritos y no tratados, especialmente aquellas relacionadas a carencias materiales, se reestimaré el impacto del Programa en los Tratados mediante el método de emparejamiento o Propensity Score Matching.

Primero, se estimó la probabilidad de participar en el Programa usando la información de los niños y hogares participantes del Programa y de los que no participan del mismo (ver anexo 4). En esta estimación de $P(X)$ se incluyeron variables que afectan tanto la decisión de participación en el Programa como a la variable de resultado de manera simultánea, como son las condiciones previas del desarrollo del niño, características de la madre cuidadora y gradientes de riqueza material del entorno del hogar en el que vive el niño. El objetivo de esto es garantizar el supuesto de independencia condicional al tratamiento, como supuesto relevante para realizar estimaciones de impacto por medio del método de emparejamiento, de manera tal que la variable de resultado sea independiente del tratamiento una vez que se condiciona por la probabilidad de participación en el Programa.

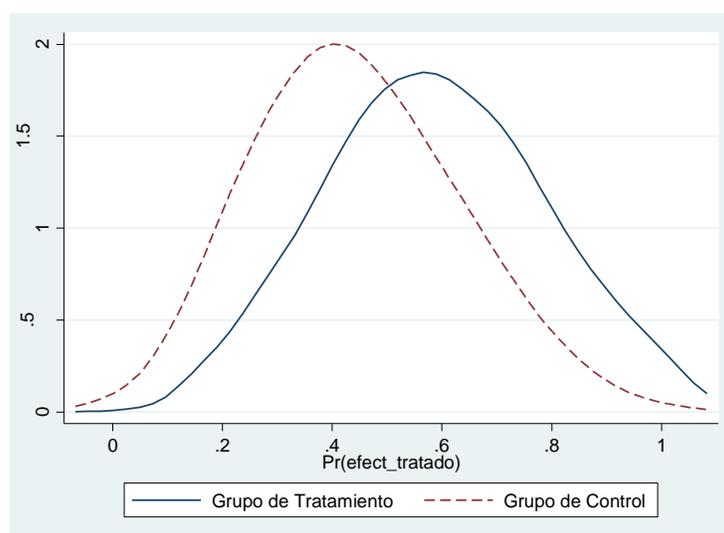
Se observa que las determinantes significativas que aumentan la probabilidad de participación del Programa son el número de hijos que tenga la mamá, la condición de jefa de hogar de la madre, si el hogar tiene letrina como tipo de servicio sanitario, así como la condición del tipo de paredes o de techo que tenga el hogar. Lo anterior hace pensar que aquellos hogares con más niños, en donde la mamá se encuentra sola y cuyos hogares parecen tener las características de casas prefabricadas (losetas de concreto y zinc) son aquellos que integran a sus hijos a los servicios de cuidado infantil. Por otro lado, características del hogar como el número de cuartos disponibles y la condición de tener un piso de cerámica – uno de los tipos de piso más caros y de mejor calidad que un hogar puede tener – resultaron incidir negativamente en la probabilidad de enviar a los niños al recibir los cuidados infantiles de los CICO y el acompañamiento familiar.

Esto refuerza el hallazgo de que los hogares beneficiarios del programa corresponden a aquellos hogares que tienen más carencias materiales, con más niños y con jefaturas monoparentales femeninas; lo que implica que el Programa está bien focalizado a hogares realmente pobres y que más necesitan de este tipo de intervenciones públicas. Sin embargo, esto también revela que existe

sesgo de selección en cuanto a las condiciones socioeconómicas entre los grupos de niños tratados y no tratados.

Una vez modelada $P(X)$, se procedió a predecir las probabilidades de participación de cada niño tanto para los participantes como para los controles. La figura 3, muestra la densidad de la probabilidad estimada de participación en el Programa tanto para los niños tratados como para los controles. Por simple inspección visual, se puede afirmar que la muestra cuenta con suficiente área de soporte común – controles y tratados con el mismo PScore – para poder realizar la estimación del Efecto Promedio del Tratamiento por medio del método del emparejamiento.

Figura 3. Densidad de probabilidad de participación pre-emparejamiento



Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013

Dado lo anterior, se procedió a estimar el Efecto Promedio del Tratamiento (ATT por sus siglas en inglés) por medio de dos algoritmos: 1 vecino más cercano (VMC) con reemplazo y emparejamiento por distancia máxima (radius matching). El primer mecanismo para emparejar cada individuo tratado con su respectivo control, se eligió con reemplazo porque la inspección descriptiva de los datos permite considerar que persisten las diferencias socioeconómicas entre los niños tratados y no tratados; por lo que existen pocos buenos controles para los niños participantes del Programa. Cabe señalar que la utilización de este algoritmo implica que un mismo niño sirve de control para más de un niño tratado, lo que permite mejorar la calidad del emparejamiento y reducir el sesgo; aunque ello implique reducir el número de individuos diferentes que permitan calcular el contrafactual y por tanto aumentar la varianza en las estimaciones. Para efectos de comparación se utiliza a 3 VMC y 5 VMC.

Por otro lado, el emparejamiento de distancia máxima permite emparejar a cada niño participante con el niño control más cercano, siempre y cuando la distancia entre las probabilidades de participación sean menor o igual a la distancia especificada (Bernal & Peña, 2011), lo que mejora la calidad del emparejamiento, en búsqueda de la comparación de los niños más parecidos entre sí, entre el grupo de los tratados y los controles, nuevamente asumiendo el costo de aumentar la varianza en las estimaciones.

La tabla 6 muestra los resultados no paramétricos del impacto del Programa en los tratados. Los resultados se aproximan a las estimaciones presentadas en la estimación anterior. El impacto del Programa resulta ser estadísticamente significativo y a favor de los niños tratados en la reducción de sus problemas de desarrollo. Dicho impacto se encuentra en un intervalo de entre 7 y 12 puntos porcentuales, en dependencia del tipo de algoritmo y la especificación que se utilice para realizar el emparejamiento. Se observa que a medida que se utilizan menos controles (1 VMC y radio de 0.01), los grupos son más parecidos en términos de características observables (ver anexo 5). Si se analizan los tamaños de las muestras, se encuentra que con radio 0.01 se utilizan menos controles y tratados, lo que implica que se están realizando inferencias en base a los niños más parecidos, lo que se comprueba al revisar balance de dichos niños en los observables, tal como muestra el anexo 5. A medida que se usan más controles empiezan a aparecer diferencias en algunas variables de las madres y hogares del niño.

Tabla 6. Resultados no paramétricos del impacto del programa mediante distintos algoritmos de emparejamiento

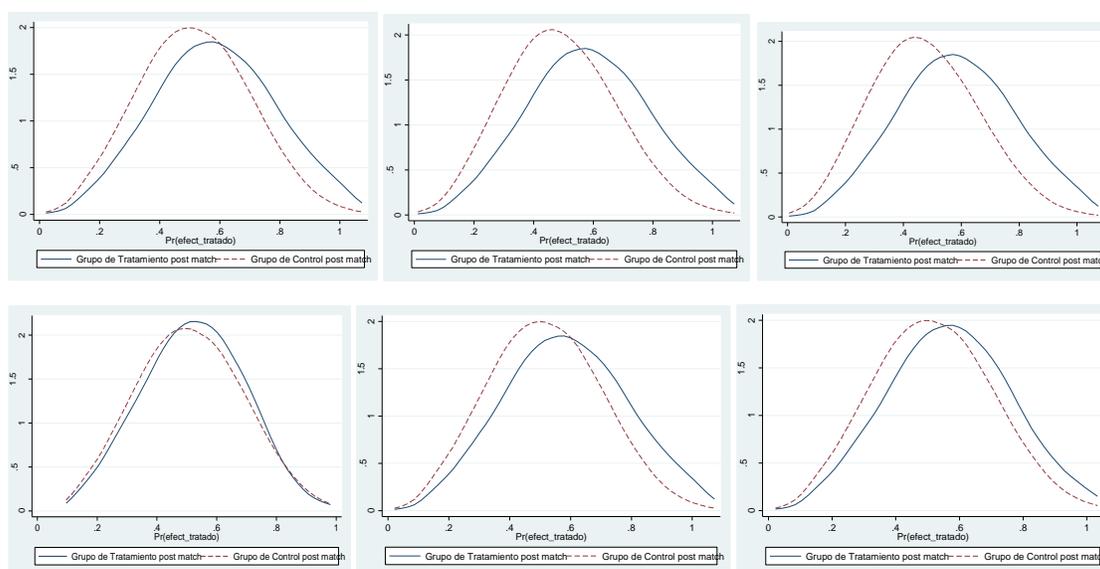
	VMC(1)	VMC(3)	VMC(5)	Radio 0.01	Radio 0.1	Radio 0.5
ATT	-0.1266*** (-2.63)	-0.1272*** (-3.23)	-0.1137*** (-3.00)	-0.0722* (-1.63)	-0.1266*** (-2.63)	-0.1083** (-2.43)
Controles	241	387	433	233	241	241
Tratados	545	545	545	360	545	517
Total	786	932	978	583	786	758

Valores t-test entre paréntesis. Indica significancia estadística: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013 y variable de resultado levantada en noviembre del año 2015

Finalmente, se revisan las densidades de probabilidad de participación predichas para los niños tratados y controles, acotados a las muestras utilizadas en cada estimación de emparejamiento, y la figura 4 (post-emparejamiento) se visualiza que la estimación por medio de 1 VMC y radio de 0.01 se obtienen las estimaciones que mejor se ajustan entre ambos grupos en cuanto al balance en sus probabilidades de participación del Programa. Esto refuerza los hallazgos del impacto encontrados tanto en esta estrategia como en las presentadas anteriormente, aunque para una muestra más pequeña de la que se contaba al inicio en este estudio. Esto resalta la tensión que existe entre el sesgo del estimador del impacto y su varianza. Al comparar individuos “más parecidos” entre sí, se reduce el sesgo en las estimaciones; sin embargo, al utilizar más observaciones para construir el contrafactual de cada participante se aumenta la varianza del estimador y se sacrifica precisión de la misma.

**Figura 4. Densidad de probabilidad de participar posterior al emparejamiento:
VCM (1) con reemplazo, VMC(3), VMC(5), Radius (0.01), Radius (0.5), Radius (0.1)**



Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013 y variable de resultado levantada en noviembre del año 2015

En resumen, el programa tiene un efecto positivo y estadísticamente significativo en el desarrollo de los niños. Los resultados por medio de esta estrategia (matching) son sensibles al número de observaciones que se utilizan, especialmente porque se observan diferencias en características observables entre los grupos de niños tratados y controles. Una vez que se logra balancear las características observables en los grupos, se consigue la reducción del sesgo de selección en

observables, al costo de tener mayor varianza en las estimaciones y realizar conclusiones en una muestra más reducida.

7 Recomendaciones de Política

A pesar de que se logró estimar el impacto en el Programa en los Tratados, por las metodologías que permitieran eliminar el sesgo en observables y que permitieran garantizar mayor validez interna de los resultados, es importante señalar un sinnúmero de dificultades que presentó el Programa tanto en la implementación del diseño de evaluación como en la implementación del Programa mismo, así como realizar las recomendaciones pertinentes para ejecución de programas futuros. Esta sección pretende hacerse cargo de estos temas.

Por un lado, el Programa mismo tuvo dificultades a la hora de ser implementado. Aunque el Programa fue aprobado como proyecto de Ley en la Asamblea Nacional en el año 2009, éste fue implementado hasta el año 2014 debido a dificultades en la legalización de los terrenos en los que se construirían los CICO. Este rezago significativo entre la aprobación y la implementación del Programa pone en evidencia una mala gestión y organización por parte de las autoridades para prever y manejar las dificultades operativas que un Programa de esta envergadura implica.

Otro error importante en el diseño del programa es que éste ofreció servicios de cuidado infantil sólo medio tiempo; lo que resulta ser una decisión contra intuitiva que lógicamente incidió en el bajo take-up que tuvo la oferta del Programa entre los pobladores a ser beneficiados. Esta implementación no responde a las necesidades reales de las madres trabajadoras, por ejemplo. En vista de que no hubo suficiente demanda de matrícula y asistencia a los servicios de medio tiempo de los CICO en los dos primeros años de implementación de este servicio, se recomienda que la planificación de estos aspectos operativos se base en la realización de encuestas de opinión; y que se tomen decisiones en base a tales demandas identificadas, no sólo para mejorar los resultados de la muestra, sino también para responder a necesidades reales de la población a intervenir.

Por otro lado, aunque el diseño de evaluación de este Programa fue realizado por expertos en el tema; su implementación presentó muchas dificultades importantes que deben tomarse en cuenta en evaluaciones futuras.

En primer lugar, los profesionales que diseñaron la evaluación fueron completamente externos al proceso de implementación de la misma evaluación y no realizaron el acompañamiento pertinente

al equipo ejecutor del Programa. A esto se suma que el personal que implementó el Programa contaba con pocas capacidades en relación al dominio teórico y práctico de las evaluaciones de impacto. Este punto pone en manifiesto la importancia no sólo de que una evaluación de impacto sea bien diseñada desde antes de implementado un Programa, sino que también se implemente con el mayor rigor posible y que cuente con el seguimiento técnico adecuado en coordinación con los ejecutores del Programa, cuyos intereses se alejan del cumplimiento de protocolos de investigación. Esto es por lejos, una de las principales recomendaciones – que a pesar de ser bien sabidas – sigue siendo una de las principales dificultades que enfrentan este tipo de estudios.

En segundo lugar, los datos pre-tratamiento evidencian que la aleatorización falló en su implementación. Esto puede deberse a dos razones particulares: i) se realizó la aleatorización en base a datos estadísticos muy antiguos (CENSO 2005) y dado que hubo retrasos entre el diseño del Programa y su implementación, el contexto socioeconómico que describían los datos pudo haber cambiado en ese período de tiempo, de manera tal que los barrios que antes eran pobres, luego ya no lo eran tanto; ii) es posible que la débil institucionalidad que caracteriza al país haya influido en que no se respetara la asignación de los barrios a sus categorías de tratamiento y control, esto especialmente se pudo deber a que intereses políticos fuesen más importantes que el protocolo de la aleatorización. Con esto se desataca la importancia de contar con instituciones públicas que se enfoquen en seguir e implementar políticas de país, imparcialmente a los intereses de agentes particulares. Adicionalmente, esta situación pone de manifiesto la importancia de contar con estadísticas nacionales actualizadas, así como de adecuar los rediseñar las evaluaciones antes de ser implementadas cuando existe tanto rezago en la información disponible en la que se basa la aleatorización.

En tercer lugar, el diseño del Programa consideraba la asignación de las familias a dos modalidades de atención T1 (CICO) y T2 (visita domiciliar) separadamente, pero esto no fue respetado debido a las demandas de los pobladores por ser atendidos lo más pronto posible. Por lo tanto, los ejecutores decidieron ofrecer ambas modalidades de atención a todos los barrios que habían sido seleccionados para ser intervenidos indistintamente y en distintos niveles de exposición al tratamiento, lo que no fue capturado ni adecuadamente monitoreado en los datos y tampoco permitió seguir a cabalidad el diseño de aleatorización diseñado en un principio.

En cuarto lugar, a pesar de haber contratado a la misma consultora para realizar la capacitación y el levantamiento de datos, hubo errores significativos en la manera en la que se capacitaron a los encuestadores, lo que indujo importantes diferencias en los niveles de porcentajes de niños con problemas en su desarrollo en el levantamiento de datos de línea base en relación al levantamiento de datos de seguimiento. Esto limitó la detección rigurosa del estado inicial y final de los niños en este tema, y la posibilidad de realizar una estimación mediante de diferencias en diferencias en la variable de resultado. Nuevamente, es altamente deseable que sea estrictamente riguroso y responsable en la capacitación de los encuestadores y el seguimiento del levantamiento de las encuestas para realizar las correcciones que se deban hacer pertinentemente en el tiempo. Esto es una consideración clave a considerar en cualquier estudio que implique implementación de encuestas, especialmente cuando se hace necesario analizar datos de panel.

Luego, para mejorar la precisión de las estimaciones en términos de impacto de programas futuros, se podría considerar necesario también incrementar el tamaño de la muestra del estudio hasta en un 50% respecto a lo indicado por los cálculos de potencia. Esta precaución podría compensar parcialmente las pérdidas potenciales en los datos del panel durante el periodo de seguimiento. Finalmente, existe un vacío en la información disponible para realizar análisis de la calidad de los servicios prestados en todos sus ámbitos, aun cuando es sabido que la sola prestación de servicios y aumento de cobertura de los mismos no es suficiente para garantizar mejoras significativas a largo plazo. Por lo tanto, se recomienda que para futuras operaciones, el Sistema de Monitoreo se defina en un período previo a su implementación, para asegurar una visión común sobre la definición de calidad del servicio y de cómo medirlo en términos estructurales y de proceso.

La evaluación en términos de calidad hace relevancia a la disposición o carencia de recursos que facilitan el proceso de aprendizaje de los niños e incluye características de la infraestructura de los centros, los equipos disponibles, protocolos de seguridad, características de educadores y de los mismos niños; mientras la evaluación de la calidad de los procesos implica analizar la implementación del currículo, de la forma de interactuar entre niños y educadores, niños y padres, educadores y padres. Así pues, la primera dimensión hace referencia a los recursos materiales y no materiales que son necesarios para que ejecuten los programas; mientras la segunda, evoca las relaciones interpersonales que permiten hacer llegar a los niños y sus familias los beneficios que el programa espera entregar con los recursos estructurales disponibles. La carencia o disponibilidad

de esta información no es en detrimento de los hallazgos encontrados, sino un complemento que permita desentrañar mecanismos por los que ocurren o no ocurren ciertos resultados y que además brinde pautas para mejorar los servicios prestados.

8 Conclusiones

El Programa Urbano de Bienestar para la Niñez en Extrema Pobreza es un proyecto piloto que fue implementado en Nicaragua entre inicios del año 2014 y el año 2016. Su objetivo es mejorar el desarrollo integral de los niños menores de 6 años que viven en barrios urbanos en extrema pobreza mediante la entrega de diversos servicios de cuidado infantil. El Programa ofrece desde acompañamiento a los padres para inscribir a los niños en el registro civil hasta construcción de guarderías para los niños.

Este estudio evalúa el impacto del acceso a servicios de cuidado infantil en guarderías (CICO) y visitas domiciliarias por una educadora de párvulos en el desarrollo integral de los niños menores de 6 años. Se evalúa este componente de la intervención por tres razones: i) es el fin último del programa, ii) este servicio representa dos tercios del presupuesto del programa y iii) las autoridades pertinentes sólo facilitaron los datos de panel con los indicadores de resultado de este objetivo.

El Programa focaliza sus servicios a niños que viven en barrios urbanos en condiciones de extrema pobreza. La estimación del impacto se realiza usando los resultados de la prueba CLAP realizada en niños menores de 6 años en noviembre del año 2015, correspondiente a levantamiento de datos de seguimiento; mientras que se utilizó la información de los niños, las madres cuidadores y los hogares obtenida en noviembre del año 2013, previo a la implementación del Programa, para controlar por variables observables en las estimaciones.

El impacto del Programa es estimado mediante dos estrategias. Primero, valiéndose de la oferta aleatoria de los servicios como mecanismo de asignación de los beneficios del programa, se estima el Impacto Promedio del Programa en los Tratados. Posteriormente, se examina la robustez de los resultados a través de la estimación no paramétrica del impacto por el método del emparejamiento. De este estudio se obtienen dos conclusiones principales. Primero, los resultados muestran que la simple diferencia de medias entre niños tratados y no tratados no es significativa. Sin embargo, el análisis de estadística descriptiva permite reconocer diferencias significativas entre los niños tratados y no tratados en términos de características observables, especialmente las relacionadas a las condiciones socioeconómicas del hogar, una vez controladas estas diferencias encuentra un impacto significativo del Programa de alrededor de 7 puntos porcentuales.

Segundo, en vista de que los datos alarman de posibles problemas en la implementación aleatoria del programa y se sospecha de sesgo de selección, se analiza la robustez de los hallazgos anteriores mediante la estimación del Impacto Promedio del Programa en los Tratados (ATT) por el método del emparejamiento. Las estimaciones no paramétricas señalan que el Programa reduce el porcentaje de niños con problemas en su desarrollo entre 7 y 12 puntos porcentuales. En este análisis, es importante reconocer la tensión que existe entre dos objetivos deseables de la estimación del impacto: tener un estimador insesgado o minimizar la varianza del mismo. En este estudio se utilizaron distintos algoritmos para comparar resultados que alternan tales características del estimador del impacto. En este caso, se observa que a medida que se usan menos controles, los grupos son más parecidos en términos de variables observables, esto permite usar los niños más parecidos entre los grupos tratados y no tratados por lo que se reduce el sesgo del estimador, a costa de tener menor precisión en los resultados.

En resumen, los principales resultados indican que tanto por estimaciones paramétricas como no paramétricas – diferencia de medias una vez controladas las diferencias en variables observables y emparejamiento, respectivamente – el programa tiene un efecto estadísticamente significativo y positivo, en el sentido que reduce el porcentaje de niños con problemas en su desarrollo, en alrededor de 7 y 12 puntos porcentuales.

Estos resultados se interpretan como la existencia de un impacto del Programa en los niños efectivamente tratados en el corto plazo. Si el Programa hubiese continuado en su segunda fase, sería interesante realizar un tercer seguimiento a los niños considerados en este estudio para comparar los resultados a mediano plazo, además de replicar las estimaciones para el grupo que actualmente es no intervenido. Adicionalmente, aún queda como tarea pendiente incluir indicadores que permitan evaluar la calidad de los servicios ofrecidos y mejorar la implementación de los diseños de evaluación para medir intensidades de tratamientos en intervenciones futuras.

9 Referencias Bibliográficas

- Araujo, M., López-Boo, F., & Puyana, J. (2013). *Panorama sobre los Servicios de Desarrollo Infantil Temprano en América Latina y el Caribe*. BID.
- Baker-Henningham, H., & López-Boo, F. (2013). *Intervenciones de Estimulación Infantil Temprana en los Países en Vías de Desarrollo*. BID.
- Barnett, W. (Winter 1995). Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. *The Future of Children. Volumen 5. No.3*, 25-50.
- Behrman, J., Cheng, Y., & Todd, P. (2004). Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach. *The Review of Economics and Statistics*, 108-132.
- Berlinski, S., & Schady, N. (2015). *Los Primeros años. El Bienestar Infantil y el Papel de las Políticas Públicas*. BID.
- Bernal, R., & Peña, X. (2011). *Guía Práctica para la Evaluación de Impacto*. Bogotá: Universidad de los Andes.
- Bernal, R., Fernández, C., Flórez, C., Gaviria, A., Ocampo, P., Samper, B., & Sánchez, F. (2009). *Evaluación de Impacto del Programa Hogares Comunitarios de Bienestar del ICBF*. Bogotá: Universidad de los Andes-Profamilia.
- BID. (2009). *Contrato de Préstamo No. 2264/BL-NI . Programa Urbano de Bienestar para la Niñez en Extrema Pobreza, Fase I*. Managua.
- BID. (2009). *Perfil de Proyecto. Programa Urbano de Desarrollo Integral para Familias en Extrema Pobreza*. Managua: Banco Interamericano del Desarrollo.
- BID. (2015). *PRIDI. Regional Project on Child Development Indicators. Base de Datos*. Managua: BID.
- BID. (2016). *Informe de Terminación de Proyecto. Programa Urbano de Bienestar de la Niñez en Extrema Pobreza*. Managua: Banco Interamericano de Desarrollo.
- Castro, V., & Laguna, J. (2016). *Reporte Evaluación Programa Urbano - Borrador-*. Managua: BID.
- CIASES. (Febrero, 2014). *Informe Final. Levantamiento de una Línea de Base para evaluar los resultados del Programa Urbano de Bienestar para la niñez en Extrema Pobreza*. Managua.
- CIASES. (Septiembre 2015). *Manual de Aplicadores: Medición Proyecto Urbano 2015*. Managua.
- Cunha, F., & Heckman, J. (2010). Estimating the technology of cognitive and noncognitive skill formation . *Econometrica*, 883-931.

- Cunha, F., Heckman, J., Lochner, L., & Masterov, D. (2006). Interpreting the evidence on life cycle skill formation. En E. Hanushek, & F. Welch, *Handbook of the Economics of Education* (págs. 697-812). Amsterdam: North-Holland.
- Duflo, E., Glennerster, R., & Kremer, M. (2006). *Using Randomization in Development Economics Research: A Toolkit*.
- Gertler, P., Martínez, S., Premand, P., Rawlings, L., & Veermeersch, C. (2011). *La Evaluación de Impacto en la Práctica*. Washington DC: Banco Internacional de Reconstrucción y Fomento/Banco Mundial.
- Gertler, Paul; Galiani, S; Sánchez, E & Vivo, S. (Octubre 2009). *Preparación de la Operación NI0155. Programa Urbano de Bienestar para la Niñez en Extrema Pobreza. Evaluación de Impacto. Nota Metodológica*. Managua: Banco Interamericano de Desarrollo.
- Heckman, J. (2000). Policies to foster Human Capital . *Research in Economics*, 3-56.
- Heckman, J. (2008). Schools, Skill and synapses. *Economic Inquiry*.
- Heckman, J., & Masterov, D. (2007). The Productivity Argument for Investing in Young Children. *Allied Social Sciences Association annual meeting*. Chicago.
- Heinrich, C., Maffioli, A., & Vázquez, G. (2010). *A Primer for Applying Propensity-Score Matching*. Banco Interamericano de Desarrollo.
- INIDE. (2011). *Encuesta de Hogares sobre Medición del Nivel de Vida 2009. Principales Resultados: Pobreza, Consumo, Ingreso*. Managua.
- Instituto de Estudios Estratégicos y Políticas Públicas. (2011). *Evolución de Políticas y Prácticas en la Protección Social a la niñez de 0-3 años*. Managua.
- Khandker, S., Koolwal, G., & Samad, H. (2010). *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington DC: The International Bank for Reconstruction and Development / The World Bank.
- Magnuson, K., & Duncan, G. (2014). Can Early Childhood Interventions Decrease Inequality of Economic Opportunity? *Inequality of Economic Opportunity in the United States*. Boston, Massachusetts.
- Martínez, S. (2015). *Midiendo el Impacto: Métodos de Evaluación de Impacto*. Banco Interamericano de Desarrollo.
- National Scientific Council. (s.f.). *La Ciencia del Desarrollo Infantil Temprano*. Center on the Developing Child at Harvard University.

- Novoa-Hidalgo, G., & Urzúa, S. (2012). The Effects of Participation in Public Child Care Centers: Evidence from Chile. *Journal of Human Capital*, 2012, vol. 6, no. 1.
- Rosenbaum, P., & Rubin, D. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*. Volume 79. Number 387, Applications Section .
- Rosero, J., & Oosterbeek, H. (2011). *Trade-Offs between Different Early Childhood Interventions: Evidence from Ecuador*. Tinbergen Institute Discussion Paper 11-102/3.
- Shady, N., Behrman, J., Araujo, M., Azuero, R., Bernal, R., Bravo, D., . . . Vakis, R. (2014). *Gradientes de Riqueza en el Desarrollo Cognitivo durante la Primera Infancia en cinco países de América Latina*. BID.
- Shokoff, J., Ritcher, L., Van Der Gaag, J., & Bhutta, Z. (2012). An Integrated Scientific Framework for Child Survival and Early Childhood Development. *Pediatrics*. Volume 120. .
- UNICEF. (2008). *El Cuidado Infantil en los países industrializados: transición y cambio, Report Card Innocenti No.8*. Florencia: Fondo de las Naciones Unidas para la Infancia.
- Verdisco, A., Cueto, S., Thompson, J., & Neuschmidt, O. (2015). *PRIDI. Regional Project on Child Development Indicators. Urgencia y Posibilidad*. BID.

10 Anexos

Anexo 1. Ejemplo de cuestionario CLAP implementado en el Programa para evaluar los resultados

FORMATO 1. ESCALA CLAP				EDAD: 22-23 MESES
CODIGO CP DEL NIÑO/A:	FECHA DE NACIMIENTO __/__/__	EDAD EN MESES : __ __	FECHA DE APLICACION DE PRUEBAS : __/__/__	
Instrucción: 1. Encierre en un círculo el logro que esté presente en cada área. 2. Poner en la columna Ev. A, el resultado de la evaluación de área (A, EP, N). 3. Realizar la evaluación global				Criterios para evaluar las áreas: → Alerta si el niño/a no realizó uno o varios logros resaltados en letra negrita(A) . → En Proceso: Si el niño/a únicamente realizó los logros resaltados en letra negrita (EP) . → Normal: si el niño/a cumplió con los logros que están resaltados en letra negrita, además con uno o más de los logros que NO lo están en letra negrita(N) .
1. ÁREA SOCIOAFECTIVA		3. ÁREA DE COORDINACIÓN		
Logro	Ev. A	Logro	Ev. A	
a) Se viste con ayuda		a) Introduce objetos pequeños en frascos o botellas.		
b) Juega en paralelo, o juega cerca de otros niños pero sin compartir los juguetes		b) Construye torres de 2 cubos o tacos de madera.		
c) Se alimenta con cuchara derramando un poco del contenido.		c) Construye torre de 3 cubos o tacos de madera.		
d) Ayuda en tareas simples de la casa		d) Construye torre de 4 cubos o tacos de madera.		
e) Se pone alguna ropa.		e) Construye torre de 5 cubos o tacos de madera.		
f) Se lava y seca las manos.				
2. ÁREA DE LENGUAJE		4. ÁREA MOTORA		
Logro	Ev. A	Logro	Ev. A	
a) Usa palabras como: dadá, mamá, pan, agua, oso.		a) Sube escaleras con ayuda		
b) Utiliza más palabras.		b) Sube escaleras (gradas) tomado del pasamano.		
c) Señala algunas partes del cuerpo		c) Patea la pelota.		
d) Cumple órdenes simples.		d) Salta en el lugar y lanza la pelota.		
		e) Pedalea en triciclo.		
Criterios para la Evaluación Global: <ul style="list-style-type: none"> ▪ Alerta: Dos o más alertas en cualquier área. ▪ En Proceso: Un alerta en cualquier área o dos o más áreas en Proceso. ▪ Normal: Un área En Proceso y tres áreas Normales o todas las áreas Normales. ▪ Otro resultado: que no corresponda a los anteriores, su evaluación será En Proceso. 			Resultados de la Evaluación Global: A ____ EP ____ N ____	

Anexo 2. Características de los niños y las madres: Estadística descriptiva y test de medias según UE tratadas y UE de control (2013)

VARIABLE	Grupo de tratamiento (2040 niños)	Grupo de control (718 niños)	Diferencia de medias	Total Muestra (2758 niños)
Porcentaje de niños con problemas en su desarrollo global	11.82	9.6	2.22	11.19
Edad promedio de los niños (en meses)	21.92	21.62	0.3	21.83
Edad de la madre (en años)	26.7	26.8	-0.1	26.7
Edad promedio de la madre biológica al dar a luz (en años)	19.09	19.56	-0.47*	19.22
Número de controles prenatales de la madre durante el embarazo	7.12	7.21	-0.09	7.15
Número de hijos vivos en la actualidad	2.33	2.10	0.23**	2.27
Si la madre tiene educación técnica o universitaria (1=sí)	13.10	17.88	-4.78*	14.44
Si la madre dio lactancia exclusiva al niño (1=sí)	82.37%	81.56%	0.81	82.14%
Promedio de horas de juego con su hijo al día	1.84	1.97	-0.13	1.87
Le lee a su hijo alguna vez (1=sí)	86.02	86.74	-0.72	86.22%
Observaciones (Número de UE)	56	22		78

***Diferencias significativas $p < 0.01$, ** Diferencias significativas $p < 0.05$, * Diferencias significativas $p < 0.10$

Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013

Anexo 3. Características socioeconómicas de los hogares: Estadística descriptiva y test de medias según UE tratadas y UE de control (2013)

VARIABLE	Características socioeconómicas de los hogares			
	Grupo de tratamiento (2040 niños)	Grupo de control (718 niños)	Diferencia de media	Total Muestra (2758 niños)
Número de miembros en el hogar	5.41	5.60	-0.19	5.46
Número de cuartos de los que dispone el hogar	2.25	2.58	-0.33**	2.34
Porcentaje de hogares con piso de tierra	50.19	38.24	11.95**	46.82
Porcentaje de hogares que tienen letrina	64.10	47.94	16.16**	59.55
Porcentaje de hogares con paredes de concreto	42.39%	48.98%	-6.59	44.25%
Porcentaje de hogares que usan leña como combustible	42.39%	49.08%	-6.69	44.27%
Observaciones (Número de UE)	56	22		78

***Diferencias significativas $p < 0.01$, ** Diferencias significativas $p < 0.05$, * Diferencias significativas $p < 0.10$

Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013

**Anexo 4. Regresion para la probabilidad de participar en el programa: efectos marginales
(variable dependiente: 1= participa, 0= no participa)**

VARIABLES	regresión (1)
mal_global_13	-0.0439 (0.0529)
mama_sin_dormir	0.0295 (0.0331)
hijos_vivos	0.0272** (0.0115)
madre_jefaHogar	0.1171*** (0.0426)
agua_tuberia	0.0224 (0.0379)
cuartos_hogar	-0.0253* (0.0139)
techo_zinc	0.3327*** (0.0675)
cocina_leña	0.0224 (0.0379)
paredes	0.1225*** (0.0387)
letrina	0.2093*** (0.0357)
Pobre	0.0055 (0.0223)
edad_luz_madre	-0.0019 (0.0048)
lactancia_exclusiva	0.0530 (0.0426)
nivel_educativo_mom	-0.0085 (0.0107)
casa_propia	0.0117 (0.0459)

VARIABLES	regresión (1)
luz	0.0006 (0.1655)
piso_ceramica	-0.0823** (0.0419)
R2 ajustado	0.1105
Observations	1,053
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Donde:

Variable dependiente: si el niño participa del programa=1, 0 si no siendo que su barrio es intervenido

mal_global_13: 1 si el niño tiene problemas en su desarrollo y 0 si no

Mama_sin_dormir: variable dicotómica donde 1 si la mamá no durmió bien en la última semana y 0 si no

Hijos_vivos: variable discreta del número de hijos vivos que tiene la madre

Madre_jefaHogar: variable dicotómica donde 1 si la mamá es jefa del Hogar y 0 si no

Agua_tubería: 1 si el hogar recibe agua potable por medio de tubería, 0 si no

Cuartos_hogar: variable continua que da cuenta del número de cuartos de los que dispone el hogar

Techo_zinc: variable dicotómica donde 1 si el hogar tiene techo de zinc y 0 si no

Cocina_leña: variable dicotómica donde 1 si el hogar cocina con leña y 0 si no

paredes: variable dicotómica donde 1 si el hogar tiene paredes de concreto y 0 si no

letrina: 1 si el hogar tiene letrina como servicio higiénico, 0 si no

Pobre: 1 si el gasto en la semana anterior a la encuesta fue menor a C\$1000, 0 si el gasto en la semana anterior fue entre C\$1001 y C\$1500

Edad_luz_madre: variable continua que da cuenta de la edad a la que la madre dio a luz (en años)

Lactancia_exclusiva: 1 si la mamá le dio lactancia exclusiva al niño los primeros seis meses, 0 si no

Nivel_educ_mom: 0 si ninguno, 2 si preescolar, 3 educación de adultos, 4 primaria, 5 secundaria, 6 técnico básico, 7 técnico medio, 8 técnico superior, 9 universitario.

Casa_propia: 1 si la casa es propia, 0 si no

Luz: 1 si la casa tiene luz, 0 si no

Piso_cerámica: 1 si el hogar tiene piso de cerámica, 0 si no

Anexo 5. Diferencias de Medias en variables descriptivas según grupo de tratamiento y control generados por método de emparejamiento

VARIABLE	VMC (1)	VMC (3)	VMC (5)	Radius (0.1)	Radius (0.5)	Radius (0.01)
Porcentaje de niños con problemas en su desarrollo global_2013	0.90	1.25	1.48	0.90	-0.17	0.73
Edad promedio de los niños (en meses)	-0.84	-0.73	-0.34	-0.84	-0.73	-0.95
Edad de la madre (en años)	-0.06	0.37	0.35	-0.06	-0.15	-0.77
Edad promedio de la madre biológica al dar a luz (en años)	-0.31	-0.39*	-0.51**	-0.31	-0.26	-0.14
Número de controles prenatales de la madre durante el embarazo	-0.02	-0.09	-0.11	0.02	0.02	0.06
Número de hijos vivos en la actualidad	0.29**	0.35**	0.37***	0.29**	0.21*	0.03
Si la madre tiene educación técnica o universitaria (1=sí)	-0.88	-3.78*	-5.56**	-0.88	-0.34	1.45
Si la madre no ha dormido bien en la última semana (1=sí)	0.70	4.21	3.63	0.70	-2.94	-1.53
Si la madre tiene menos de 27 años (1=sí)	-0.45	-3.41	-3.22	-0.45	-1.13	3.08
Si el niño colorea en casa (1=sí)	5.18	2.02	2.73	5.18	5.72	4.18
Si la madre dio lactancia exclusiva al niño (1=sí)	0.58	2.23	2.64	0.58	0.12	-0.69
Promedio de horas de juego con su hijo al día	-0.16	0.20**	0.21**	0.16	-0.13	-0.09
Le lee a su hijo alguna vez (1=sí)	1.58	0.69	-0.09	1.58	2.67	2.44

VARIABLE	VMC (1)	VMC (3)	VMC (5)	Radius (0.1)	Radius (0.5)	Radius (0.01)
Número de miembros en el hogar	0.01	-0.01	-0.04	-0.01	0.10	-0.18
Número de cuartos de los que dispone el hogar	-0.15	-0.24**	-0.30**	-0.15**	-0.12	-0.08
Porcentaje de hogares con piso de tierra	7.24*	10.71**	12.78***	7.24**	5.91	3.19
Porcentaje de hogares que tienen letrina	10.30***	16.25***	19.34***	10.30**	8.72**	2.21
Porcentaje de hogares con paredes de concreto	-2.4	7.01**	-8.02**	2.48	-0.46	2.63
Porcentaje de hogares que usan leña como combustible	2.07	8.32**	9.79**	2.07	1.08	2.78
Porcentaje de hogares con tubería	-11.48***	-12.48***	-12.95***	-11.48***	7.17**	2.24
Porcentaje de hogares con piso de cerámica	-4.06	-6.87**	-10.02***	-4.06	-2.98	1.34

***Diferencias significativas $p < 0.01$, ** Diferencias significativas $p < 0.05$, * Diferencias significativas $p < 0.10$

Fuente: Elaboración propia basada en encuestas realizadas en noviembre del año 2013