

Evaluación externa de impacto del Programa Oportunidades 2006




Mauricio Hernández Ávila
Bernardo Hernández Prado
José E. Urquieta Salomón
EDITORES

Tomo III

Documento metodológico



Instituto Nacional
de Salud Pública



**Evaluación externa de impacto del
Programa Oportunidades 2006**

Autores*

Jere Behrman,[†] Petra Todd,[‡]
Bernardo Hernández,[§] José Urquieta,[§]
Orazio Attanasio,[#] Manuela Angelucci,[&]
Mauricio Hernández[§]

[†] Universidad de Pennsylvania,

[‡] Instituto Nacional de Salud Pública,

[#] University College London, Institute for Fiscal Studies,
National Bureau of Economic Research y BREAD,

[&] Universidad de Arizona e Institute for the Study of Labor.

Nota: los puntos de vista expresados por los autores de este documento no reflejan la postura institucional del Instituto Nacional de Salud Pública. Agradecemos la colaboración del grupo evaluador de *Oportunidades* para la recolección de información y definición de estrategias de análisis. Los autores agradecen a Lynnette Neufeld y a Armando García la revisión de las secciones sobre el diseño y levantamiento de la información biológica.

**Evaluación externa de impacto del
Programa Oportunidades 2006. Tomo III • Documento metodológico**

Primera edición, 2006

D.R. ©

Instituto Nacional de Salud Pública

Av. Universidad 655

Sta. Ma. Ahuacatlán

62508 Cuernavaca, Morelos, México

Impreso y hecho en México

Printed and made in Mexico

ISBN 970-9874-12-8

Citación sugerida: Behrman JR, Todd PE, Hernández-Prado B, Urquieta-Salomón JE, Attanasio O, Angelucci M, Hernández-Ávila M. Metodología de evaluación de *Oportunidades* en zonas urbanas 2002-2004 y rurales 1997-2003. En: Hernández-Prado B, Hernández-Ávila M, Urquieta-Salomón JE, eds. Evaluación externa de impacto del Programa Oportunidades 2006. Tomo III, Documento metodológico. Cuernavaca, México: Instituto Nacional de Salud Pública, 2006.

Evaluación externa de impacto del Programa Oportunidades 2006



Mauricio Hernández Ávila
Bernardo Hernández Prado
José E. Urquieta Salomón

EDITORES



Tomo III

Documento metodológico



Contenido

Metodología de evaluación de *Oportunidades* en zonas urbanas 2002-2004 y rurales 1997-2003

Presentación	9
I. Resumen ejecutivo	11
II. Introducción	15
III. Diseño y recolección de datos en zonas urbanas	17
IV. Diseño y recolección de datos en zonas rurales	31
V. Métodos de evaluación	37
VI. Metodología utilizada en el diseño de evaluación de <i>Oportunidades</i>	51
VII. Conclusiones	53
Anexos	55
Referencias	59



Presentación

El Programa de Desarrollo Humano Oportunidades, antes Progresá, se ha constituido en la iniciativa más importante dentro de la estrategia de combate a la pobreza del Ejecutivo Federal. El principal objetivo de *Oportunidades* es contribuir a que las familias que viven en extrema pobreza superen esta condición y se rompa en el largo plazo el círculo intergeneracional de la pobreza mediante la inversión en capital humano en educación, salud, alimentación, mejora patrimonial y protección social. Además de estos propósitos, el Programa busca contribuir a la vinculación de la población beneficiaria con nuevos programas de desarrollo que propicien la mejora de la calidad de vida de las familias beneficiarias. *Oportunidades* cuenta, en 2006, con una cobertura de 5 millones de familias, lo que representa alrededor de 25 millones de beneficiarios, y maneja un presupuesto de 35 mil millones de pesos.

Como pocos programas en el país, un aspecto distintivo de este Programa ha sido que desde el inicio de su operación consideró la necesidad de contar con un componente de evaluación que permitiera tanto mejorar su diseño e implementación como documentar sus efectos en la población beneficiaria. Esta fortaleza ha permitido mantener la operación de *Oportunidades* por nueve años y convertirse en una estrategia modelo de atención a la pobreza en varios países del mundo.

De acuerdo con lo que establecen el Presupuesto de Egresos de la Federación y las Reglas de Operación del propio Programa, *Oportunidades* debe ser evaluado de manera externa por instituciones académicas y de investigación con reconocimiento y experiencia en la materia. Si bien la evaluación se ha llevado a cabo desde una perspectiva integral que involucra estudios basados en metodologías cuantitativas y cualitativas, la presente serie de volúmenes temáticos integra únicamente los documentos de evaluación de impacto con enfoque cuantitativo elaborados en 2006, y cuya coordinación fue encomendada al Instituto Nacional de Salud Pública (INSP).

Por parte del INSP, Bernardo Hernández, Lynnette Neufeld, Juan Rivera, Stefano Bertozzi, Gustavo Olaiz, Juan Pablo Gutiérrez, Fernando Meneses, José Urquieta, Juan Eugenio Hernández, Javier Carlos, Carlos Oropeza y Manuel Castro, con sus respectivos equipos de trabajo, han realizado un importante trabajo en dicha evaluación. Asimismo, este esfuerzo no sería posible sin la participación de renombrados expertos, como Susan Parker y Luis Rubalcava, ambos del Centro de Investigación y Docencia Económicas (CIDE); Graciela Teruel, de la Universidad Iberoamericana; Paul Gertler, de la Universidad de California en Berkeley; Petra Todd y Jere Behrman, de la Universidad de Pennsylvania; y Orazio Attanasio, del University College de Londres. Todos ellos han contribuido con su reconocida experiencia en las diferentes áreas en que el Programa ha sido evaluado.

De igual modo, se reconoce el invaluable apoyo del personal de *Oportunidades* para la conducción de las actividades de evaluación en sus distintas etapas, en especial el de Concepción Steta, Mónica Orozco e Iliana Yashine, así como el de sus equipos de trabajo. La colaboración de los distintos sectores involucrados en la operación de *Oportunidades*, especialmente de la Secretaría de Salud, el Instituto Mexicano del Seguro Social y la Secretaría de Educación Pública, ha sido fundamental en la conducción de la presente evaluación, la cual también se ha visto beneficiada de la interacción con el grupo de trabajo de Carola Álvarez, del Banco Interamericano de Desarrollo.

La presente serie de volúmenes de evaluación 2006 forma parte de un conjunto de trabajos que conforman la medición de impacto de *Oportunidades* en diferentes indicadores relacionados con aspectos de interés, como educación, salud, alimentación y nutrición de la población beneficiaria.

El volumen I documenta los impactos del Programa en el uso de servicios médicos y las condiciones de salud de los beneficiarios urbanos, así como el impacto del Programa en los logros educativos de niños y





jóvenes, y su relación con algunas variables de tipo laboral. Adicionalmente, en el mismo volumen se analiza cómo las características de los centros educativos en áreas rurales afecta el impacto del Programa.

El volumen II integra los resultados de la evaluación de impacto en el estado nutricional de niños y madres residentes de áreas urbanas, así como el patrón de uso de los complementos alimenticios. Finalmente, el volumen III sistematiza la metodología de la evaluación de impacto de programas, contextualizada en el marco de los diseños de los modelos de evaluación rural y urbana de *Oportunidades*.

Los documentos incluidos en esta serie de volúmenes utilizan como insumo información derivada de distintas encuestas panel diseñadas específicamente para la evaluación del Programa, tanto en áreas rurales como en urbanas, y en las que participaron más de 50 mil familias, a quienes damos nuestro más sincero agradecimiento por habernos permitido visitar sus hogares y por la manera desinteresada en que respondieron nuestros cuestionarios.

Los investigadores participantes expresan que los resultados plasmados en los documentos de evaluación son producto de su trabajo profesional y académico, y que no reflejan una postura institucional.

Estos documentos contribuyen al conocimiento sobre los efectos de *Oportunidades* hasta el momento, y proporcionan información relevante para orientar las acciones del Programa en el futuro. Igualmente, sugieren futuras líneas de investigación que ayudarán a determinar si *Oportunidades* cumple con su misión de contribuir a romper el ciclo intergeneracional de la pobreza en nuestro país.

Dr. Mauricio Hernández Ávila
Director General del INSP

El Programa de Desarrollo Humano Oportunidades (antes Progresá) es un programa del Poder Ejecutivo del Gobierno Federal que atiende a familias en condiciones de extrema pobreza en México.^a Este Programa tiene como objetivo contribuir a romper el ciclo de transmisión intergeneracional de la pobreza al incrementar el capital humano de los miembros de las familias mediante acciones en las áreas de salud, nutrición y educación.

El objetivo de este trabajo es documentar la metodología de recolección de datos utilizada tanto en la evaluación rural como en la urbana. Adicionalmente, el documento presenta una revisión de las principales estrategias de análisis utilizadas y otras que pueden utilizarse para evaluar los efectos de *Oportunidades* en zonas urbanas en el periodo 2002 a 2004, y los efectos a mediano plazo en zonas rurales con información de 1997 a 2003.

El componente cuantitativo de la evaluación de impacto de *Oportunidades* incluye una serie de encuestas que comenzaron en 1997 en zonas rurales, y en 2002 en zonas urbanas. Para las zonas urbanas se adoptó un diseño de evaluación no experimental que aprovechó el proceso de incorporación de los hogares al Programa. En las zonas donde el Programa comenzaría operaciones en 2002 (zonas de intervención) se obtuvo una muestra probabilística de manzanas, a cuyos hogares residentes se les aplicó un cuestionario para evaluar su estatus de elegibilidad. La muestra de evaluación incluyó a todos los hogares que afirmaron haber sido seleccionados. Debido a que el número de hogares que manifestaron estar incorporados a *Oportunidades* resultó menor al esperado, se decidió incluir en la muestra a hogares identificados como incorporados al Programa que vivían en manzanas adyacentes (manzanas de barrido) a las seleccionadas inicialmente.

Con el fin de obtener un grupo de comparación, se utilizaron técnicas de pareamiento para conformar un grupo externo de control mediante la selección de hogares en zonas donde el Programa comenzaría a operar en 2004 (zonas de no intervención). También se incluyó en la muestra a familias ubicadas justo por encima del punto de corte de elegibilidad (llamado grupo cuasi elegible) y a un grupo de familias no elegibles, para evaluar la focalización de *Oportunidades*. Con esta información se pueden construir los siguientes grupos de comparación para la evaluación urbana:

- I. Grupo de intervención (grupo de tratamiento): muestra de hogares elegibles en zonas de intervención e incorporados en 2002
- II. Grupo de control interno: muestra de hogares elegibles en zonas de intervención pero que no solicitaron su incorporación
- III. Grupo de control externo: muestra de hogares en zonas de no intervención que fueron clasificados como elegibles
- IV. Controles cuasi elegibles en zonas de intervención: muestra de hogares en zonas de intervención y con un puntaje de pobreza justo por arriba del punto de corte que determina la elegibilidad al Programa

Asimismo, se cuenta con un grupo de control conformado por familias en zonas de intervención clasificadas como no elegibles por tener un puntaje de pobreza ampliamente por arriba del punto de

I. Resumen ejecutivo

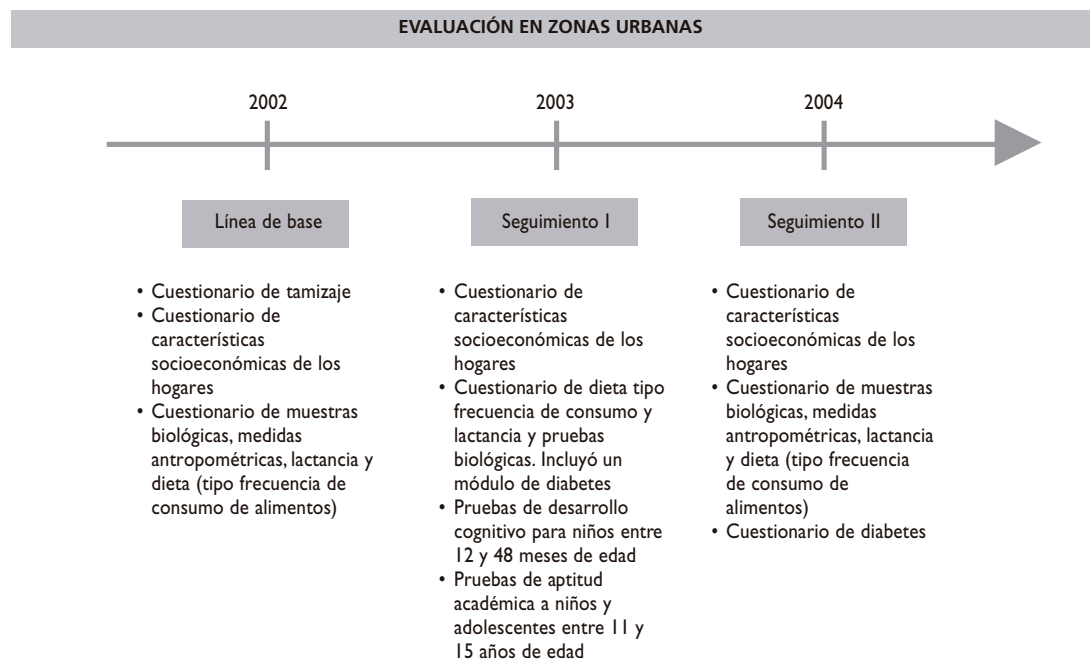


^a Las Reglas de Operación de *Oportunidades* establecen que el Programa está dirigido a la población en extrema pobreza. Si bien el término extrema pobreza tiene un significado muy específico en la literatura académica (no contar con suficientes ingresos para consumir una canasta básica de alimentos), en este documento el término extrema pobreza está entendido en términos de los criterios de elegibilidad de *Oportunidades*.

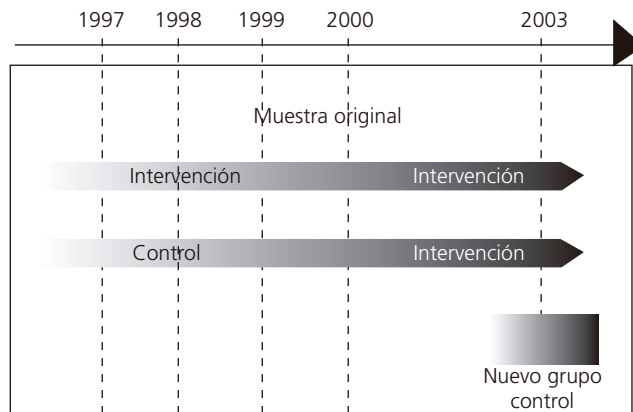


corte que determina la elegibilidad al Programa (es decir, por encima de las familias cuasi elegibles). Este grupo puede utilizarse para evaluar la focalización de *Oportunidades*.

La recolección de los datos en zonas urbanas inició con una encuesta basal llamada Encuesta de Evaluación de los Hogares Urbanos (Encelurb). Esta encuesta se aplicó por primera vez en 2002, antes de que las familias elegidas para participar en el Programa comenzaran a recibir los beneficios. Le siguieron dos encuestas de seguimiento llevadas a cabo en 2003 (Encelurb 2003) y 2004 (Encelurb 2004). Las encuestas recopilaron datos relevantes para los intereses de la evaluación, la cual incluyó información sobre educación, salud, nutrición, consumo, ingreso, actividad laboral, condiciones de vida y de la vivienda. Los instrumentos utilizados en cada encuesta se presentan en el siguiente cuadro:



En un principio, la evaluación rural se basó en un diseño experimental en el cual se conformaron los grupos de intervención y control a partir de la asignación aleatoria de 506 localidades rurales. Se recolectó información en encuestas realizadas semestralmente en 1998, 1999 y 2000. Durante la recolección de información para la evaluación rural a mediano plazo (2003) se conformó, además, un grupo de comparación no experimental, debido a que los hogares del grupo control original conformado de manera aleatoria ya se habían integrado al Programa. El nuevo grupo control se conformó mediante métodos estadísticos de pareamiento sobre la base de características observables de los hogares y las comunidades. De esta forma, la evaluación rural de mediano plazo permite establecer comparaciones entre los grupos incorporados al inicio de las operaciones de *Oportunidades* en 1998 (grupo de intervención temprana), en 2000 (grupo de intervención tardía) y el grupo de control conformado en 2003, como se muestra en la siguiente figura:



La recolección de datos en zonas rurales abarcó variables similares a las utilizadas para la evaluación urbana. Todos los procedimientos de recolección de la información biológica fueron aprobados por las comisiones de ética, bioseguridad e investigación del Instituto Nacional de Salud Pública (INSP). Toda la información fue recolectada con la obtención previa del consentimiento informado de los participantes en la encuesta.

Debido a las características especiales de los diseños de evaluación en zonas rurales y urbanas, es necesario utilizar técnicas específicas de evaluación con el fin de estimar correctamente los impactos del Programa. En este documento se discute el uso de métodos experimentales y no experimentales para analizar el impacto promedio de *Oportunidades* con un alto nivel de precisión, tales como: los métodos de diferencias en diferencias, técnicas de pareamiento, análisis de regresión discontinua y modelos estructurales, así como el uso de variables instrumentales para crear controles cuasi experimentales. Estos métodos son analizados tomando en cuenta sus supuestos y su interpretación. Por último, el documento incluye un resumen de todos los documentos de evaluación publicados hasta ahora, identificando las técnicas de análisis usadas en cada uno de ellos.

Este documento no sólo contribuye a ofrecer una descripción detallada de los diseños para la recolección de datos en zonas rurales y urbanas en la fase más reciente de la evaluación de *Oportunidades*, también analiza el uso de diferentes opciones de análisis. De esta forma se espera contribuir al mejor uso de la información recolectada por el Programa y dar una provechosa retroalimentación al mismo.



El Programa de Desarrollo Humano Oportunidades (antes Progresá) es un programa del Gobierno Federal que atiende a familias en condiciones de extrema pobreza en México.^b Su objetivo es contribuir a romper el ciclo de transmisión intergeneracional de la pobreza mediante transferencias en efectivo condicionadas a una serie de conductas de los individuos y de las familias, que buscan incrementar su educación y mejorar su salud y nutrición. El Programa inició operaciones en 1997 en comunidades rurales altamente marginadas con poblaciones menores a los 2 500 habitantes. El Programa ha expandido sus operaciones a zonas semiurbanas en 2001 y urbanas en 2002. Desde entonces, su cobertura se ha incrementado en forma sostenida hasta beneficiar en 2004 a 5 millones de familias en el país.

Desde su comienzo, el proceso de evaluación ha sido considerado parte esencial del Programa. Una evaluación confiable de los efectos de *Oportunidades* en la población beneficiaria sirve para dos propósitos: por un lado, ofrece información respecto de si el Programa ha alcanzado sus metas y objetivos; por otro, documenta aspectos susceptibles de ser mejorados.

El esquema de evaluación que analiza los efectos a corto y mediano plazo del Programa abarca metodologías tanto cualitativas como cuantitativas. La evaluación de impactos de *Oportunidades* ha sido coordinada por instituciones de investigación como el Instituto Internacional de Investigación sobre Políticas Alimentarias (IFPRI, por sus siglas en inglés), el Instituto Nacional de Salud Pública y el Centro de Investigación y Estudios Superiores en Antropología Social de Occidente (CIESAS-Occidente). Estas instituciones han coordinado los esfuerzos de un grupo de evaluadores integrado por acreditados académicos nacionales e internacionales de diversas instituciones. Con el fin de definir los lineamientos para el diseño e implementación de la evaluación de impacto, se creó el Grupo Asesor de Evaluación (GAE) conformado por renombrados expertos nacionales e internacionales.

La evaluación de impacto está conformada por dos componentes: el cuantitativo (coordinado en los últimos años por el INSP) y el cualitativo (coordinado por el CIESAS-Occidente). El componente cuantitativo de la evaluación de impacto de *Oportunidades* incluye una serie de encuestas tipo panel levantadas desde 1997 en zonas rurales, y desde el 2002 en zonas urbanas. La evaluación rural inicial se basó en un diseño experimental en el cual se conformaron grupos de tratamiento y control mediante la asignación aleatoria de 506 localidades rurales. La metodología de evaluación en zonas rurales para el periodo de 1997 a 2000 ha sido descrita en detalle por Orozco, Parker y Hernández.¹ En el ámbito urbano no fue posible utilizar un diseño aleatorio para la evaluación, por lo que en su lugar se recurrió a un diseño de evaluación no experimental. De manera similar, para la evaluación rural a mediano plazo en 2003 se utilizó también un grupo de comparación no experimental, ya que, debido al avance en el proceso de incorporación al Programa, los hogares del grupo control original creado aleatoriamente ya habían sido incorporados al Programa desde el año 2000. Estos diseños no experimentales demandan la selección de grupos de comparación que no habían sido expuestos al Programa, mediante métodos estadísticos de pareamiento basados en características observables de las familias y de las comunidades.

El objetivo de este documento es exponer la metodología de recolección de datos utilizada para la evaluación del Programa en zonas rurales y urbanas. Asimismo, el documento ofrece una revisión, y

II. Introducción



^b Las Reglas de Operación de Oportunidades establecen que el Programa está dirigido a la población en extrema pobreza. Si bien el término extrema pobreza tiene un significado muy específico en la literatura académica (no contar con suficientes ingresos para consumir una canasta básica de alimentos), en este documento el término extrema pobreza está entendido en términos de los criterios de elegibilidad de *Oportunidades*.



algunos lineamientos, sobre cómo han sido utilizados y cómo pueden utilizarse estos datos para evaluar los efectos de corto plazo de *Oportunidades* en zonas urbanas (2002-2004) y de mediano plazo en zonas rurales con la información de 2003. Las siguientes dos secciones del documento describen los datos disponibles para la evaluación en zonas urbanas y rurales, así como el diseño y los métodos usados para su recolección. En las últimas secciones se describen, y discuten, diferentes estrategias de análisis que han sido utilizadas en la evaluación del Programa hasta 2004.

El principal reto de la evaluación de impactos de *Oportunidades* consistió en generar información que permitiera identificar cambios en la población a lo largo del tiempo y que pudieran ser atribuibles al Programa. Para esto, habría sido ideal contar con dos escenarios: uno que permitiera la observación de cambios producidos en la población de estudio al recibir los beneficios de *Oportunidades*, y otro de comparación, o contrafactual, que permitiera observar lo que habría sucedido en la misma población, durante el mismo periodo de tiempo, en un ambiente por lo demás idéntico, pero sin haber participado en el Programa.²⁻⁶ Sin embargo, esta comparación no es posible de realizar, ya que no podemos observar simultáneamente a la misma población con y sin *Oportunidades*.

El problema anterior ha dado lugar a diferentes enfoques analíticos. Una opción es comparar los cambios observados en el grupo que recibió el Programa con los cambios en otro grupo que no lo haya recibido. Ésta es una simulación de un escenario contrafactual que puede ser válida si se satisface el principio de intercambiabilidad.⁴⁻⁷ En otras palabras, si se intercambia la condición de tratamiento o no tratamiento entre los grupos de evaluación, los resultados esperados deberían ser exactamente los mismos. En este caso, las diferencias observables serían atribuibles solamente al efecto del Programa.

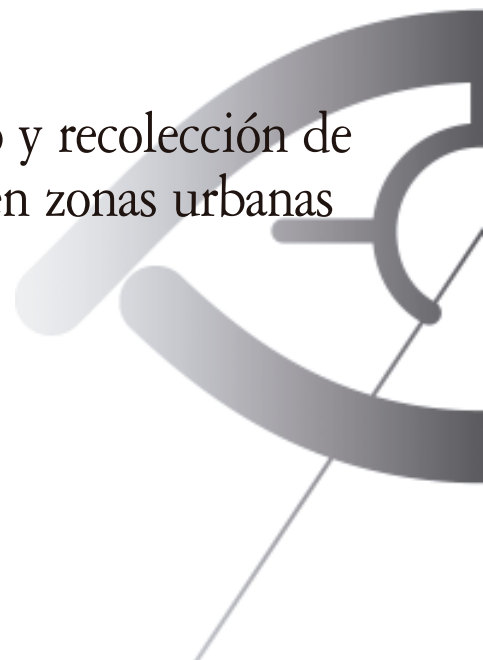
Los diseños experimentales satisfacen el principio de intercambiabilidad y son considerados los diseños más confiables y poderosos para la evaluación de programas. Los diseños experimentales asignan de manera aleatoria a los individuos, u otra unidad de muestreo, a los grupos de intervención y control, así que los individuos tienen una probabilidad conocida de ser asignados a uno u otro grupo.⁷ La asignación aleatoria asegura, en una muestra grande, que los grupos de comparación serán similares en cuanto a sus características observables y no observables, y si existe alguna diferencia entre ambos, ésta será debida a la casualidad y no a un sesgo.

En contraste con el diseño experimental utilizado para la evaluación de *Oportunidades* en zonas rurales,^{1,8} para zonas urbanas el GAE decidió no implementar un diseño experimental por razones logísticas. En su lugar se eligió un diseño no experimental utilizando métodos de pareamiento. Los métodos de pareamiento generan grupos de comparación lo más similares posibles con base en un conjunto de características observables. Sin embargo, únicamente se puede medir el nivel de similitud entre los grupos de intervención y no intervención con respecto a estas variables.^{9,10}

Antes de definirse el diseño de evaluación en zonas urbanas, los administradores del Programa elaboraron un plan de incorporación progresiva de las familias al mismo, tomando en cuenta la disponibilidad de recursos y aspectos logísticos. En este plan de incorporación los administradores del Programa tuvieron que identificar el nivel de pobreza en cada zona urbana del país (zonas de 15 000 a 1 millón de habitantes). Esta identificación se realizó con información de la Encuesta Nacional de Ingreso y Gasto de los Hogares 2000 (ENIGH 2000), con la cual se llevó a cabo un análisis multivariado basado en información demográfica y socioeconómica de las familias. Así fue posible estimar un índice de pobreza para clasificar a las familias en elegibles y no elegibles (de acuerdo con un punto de corte que determinó su condición de elegibilidad) para obtener los beneficios del Programa. Los administradores de *Oportunidades* ajustaron el modelo discriminante, y los estimadores y el punto de corte de elegibilidad fueron proporcionados al Instituto Nacional de Estadística, Geografía e Informática (INEGI), en donde se aplicó el modelo discriminante y se estimó un puntaje de pobreza para todas las familias en el país con información del Censo General de Población y Vivienda 2000. Una vez que el INEGI clasificó a todas las familias del Censo en elegibles y no elegibles, se agrupó la información a

III.

Diseño y recolección de datos en zonas urbanas





nivel de manzanas y se devolvió a los administradores de *Oportunidades* para la planeación del proceso de incorporación. La expansión del Programa en zonas urbanas no incluyó las ciudades de México, Puebla, Guadalajara, Monterrey ni otras con más de un millón de habitantes; tampoco localidades rurales y semiurbanas previamente incorporadas.

Con la información de pobreza a nivel de manzana, los administradores del Programa conformaron conglomerados o *clusters* a nivel localidad y zonas de influencia. Los conglomerados con 500 o más familias elegibles fueron considerados para su incorporación a partir de 2002; el resto sería incorporado tiempo después de acuerdo con la disposición de recursos. A diferencia de las zonas rurales, donde las familias fueron incorporadas a *Oportunidades* por medio de visitas domiciliarias, la incorporación en zonas urbanas se apoyó en módulos de afiliación.

En las zonas a ser incorporadas en 2002, se instalaron módulos en los cuales la gente podía solicitar su incorporación. Se llevó a cabo una campaña de promoción para informar a la población acerca de la disponibilidad del Programa e invitarla a acudir a los módulos de incorporación. Cuando las personas acudían al módulo se les aplicaba un cuestionario que clasificaba a su hogar como elegible o no elegible. A las familias que resultaron elegibles se les realizó una visita domiciliaria para verificar la información proporcionada. El dictamen final de la incorporación fue entregado posteriormente a las familias. Este proceso de incorporación se llevó a cabo entre julio y agosto de 2002.

Como se mencionó antes, el diseño no experimental utilizado en las zonas urbanas se construyó sobre el plan de incorporación definido por *Oportunidades*. En términos de la evaluación, este plan de incorporación permitió la identificación de zonas de intervención inmediata (programadas para incorporarse en 2002) y zonas de no intervención (programadas para incorporarse en 2004).

3.1. Esquema muestral de evaluación

Ya identificadas las zonas de intervención y no intervención, una de las principales preocupaciones del GAE era obtener una muestra lo suficientemente grande que permitiera probar las principales hipótesis de evaluación. Con base en análisis previos, se estimó que 100 manzanas con al menos tres familias elegibles en cada una de ellas sería posible estimar diferencias estadísticamente significativas de 0.2 desviaciones estándar de variables clave en las áreas de educación, salud y condiciones de vida. Por lo tanto, se propuso que el tamaño de muestra en zonas de intervención fuera de 6 000 familias elegibles e incorporadas; 2 000 familias elegibles no incorporadas; 4 000 familias cuasi elegibles (con puntajes de pobreza justo por arriba del punto de corte de elegibilidad); y 2 000 familias no elegibles (para evaluar la focalización del Programa).

3.1.1. Zonas de intervención

El tamaño de muestra se definió buscando un equilibrio entre la necesidad de una muestra más grande y las limitaciones de los recursos disponibles para el estudio. Una vez definido el tamaño de muestra y utilizando información del Censo de 2000, se seleccionaron 149 manzanas mediante un muestreo probabilístico en una sola etapa, estratificado y por conglomerados. Para obtener los tamaños necesarios de las muestras de hogares elegibles, cuasi elegibles y no elegibles se seleccionaron las manzanas con una probabilidad proporcional al número de hogares elegibles residentes en ellas. Se conformaron seis estratos de manzanas en función del número de hogares elegibles a tener los beneficios del Programa mediante el procedimiento de Tore Dalenius.¹¹ De este modo, se controló el tamaño y la dispersión de la muestra. El número de hogares elegibles por manzana varió de 1 a 2 para el estrato 1, a más de 50 en el estrato 6, como se muestra en el cuadro 1. Todas las manzanas con más de 50 hogares elegibles se incluyeron en la muestra con probabilidad igual a 1 (se incluyeron todas), y las manzanas sin hogares



elegibles fueron excluidas de la muestra. Con este procedimiento, las manzanas seleccionadas en zonas de intervención se ubicaron en 62 municipios y 71 localidades en 17 entidades federativas: Campeche, Colima, Chiapas, Guanajuato, Guerrero, Hidalgo, Estado de México, Michoacán, Morelos, Puebla, San Luis Potosí, Sinaloa, Sonora, Tabasco, Tamaulipas, Tlaxcala y Veracruz.

3.1.2. Zonas de no intervención

Una vez definida la muestra para las zonas de intervención, el GAE propuso conformar una muestra de control en las zonas de no intervención mediante un procedimiento de pareamiento. Se clasificó cada manzana de acuerdo con su región geográfica (se usaron 14 regiones geográficas) y al número de hogares elegibles residentes en cada una de ellas. Las manzanas se estratificaron, de acuerdo con su tamaño, en manzanas con 10 o más hogares elegibles y manzanas con menos de 10 hogares elegibles. Esta clasificación por región y tamaño de la manzana dio lugar a 28 grupos (14 regiones geográficas y dos tamaños de manzana según la cantidad de hogares elegibles en las mismas). Esta clasificación permitió evitar el problema de parear manzanas con un alto número de hogares elegibles en zonas de intervención con manzanas con pocos hogares elegibles en zonas de no intervención, y viceversa. Se ajustaron modelos de regresión logística para estimar la correlación entre las manzanas de intervención y las de no intervención para cada uno de los 28 grupos, utilizando covariables a nivel de manzana (cuadro 2). Posteriormente, se utilizó el método de pareamiento por vecino más cercano^c para identificar las manzanas en las zonas de no intervención que se parearían con cada una de las manzanas en las zonas de intervención. Como el número de hogares elegibles en las manzanas de intervención resultó mayor que el de las manzanas en zonas de no intervención, el pareamiento se hizo con reemplazo. Por tanto, una manzana en zona de no intervención puede estar pareada con varias manzanas en zonas de intervención. Un total de 10 manzanas de no intervención se seleccionaron para cada manzana de intervención con 12 o más hogares elegibles en términos de su proximidad, la cual fue calculada por la diferencia absoluta de las probabilidades estimadas de tener el Programa. Para cada manzana de intervención con menos de 12 hogares elegibles se escogieron tres manzanas de no intervención para la muestra. Asimismo, debido a consideraciones de costos, todas las manzanas con menos de 10% de hogares elegibles fueron excluidas de la muestra. Finalmente, se seleccionaron 387 manzanas en zonas de no intervención, distribuidas en

ESTRATO	NÚMERO DE FAMILIAS ELEGIBLES POR MANZANA	NÚMERO DE FAMILIAS ELEGIBLES POR ESTRATO	NÚMERO DE MANZANAS SELECCIONADAS POR ESTRATO
1	Manzanas con 1 y 2 familias elegibles	122 885	5
2	Manzanas con 3 a 5 familias elegibles	156 025	5
3	Manzanas con 6 a 12 familias elegibles	159 256	5
4	Manzanas con 13 a 28 familias elegibles	72 872	7
5	Manzanas con 29 a 49 familias elegibles	16 595	28
6	Manzanas con 50 o más familias elegibles	6 761	99

CUADRO 1

Tamaño de la muestra de las manzanas por estrato

Fuente: Censo Nacional de Población 2000 México

^c Los métodos de pareamiento permiten construir conjuntos de observaciones que son similares de acuerdo con un grupo dado de características. Éstos serán explicados con detalle más adelante en el documento.



76 municipios y 108 localidades en 14 estados (Colima, Chiapas, Guanajuato, Guerrero, Hidalgo, Estado de México, Michoacán, Puebla, Sinaloa, Sonora, Tabasco, Tamaulipas, Tlaxcala y Veracruz).

3.1.3. Identificación de las familias (primera etapa)

Con el fin de actualizar la información del estatus de elegibilidad de los hogares residentes en las manzanas seleccionadas para la muestra, todos los hogares en las manzanas de intervención y no intervención fueron visitados (primera etapa) y se les aplicó un cuestionario de tamizaje. La figura 1 muestra el número de hogares incluidos en la muestra en esta etapa. La clasificación de los hogares como elegibles o no elegibles se utilizó exclusivamente para la selección de la muestra de evaluación y no influyó en el proceso de incorporación al Programa.

CUADRO 2

Variables incluidas en el modelo logístico para el pareamiento de manzanas en zonas de intervención y de no intervención

Número de escuelas primarias por AGEB*
Número de escuelas secundarias por AGEB *
Número de escuelas preparatorias (educación media superior) por AGEB*
Centros de salud por AGEB *
Proporción de familias sin cuarto de baño
Proporción de familias sin agua entubada
Proporción de familias sin agua en el predio
Proporción de familias sin agua en el hogar
Proporción de familias con piso de tierra
Proporción de familias sin electricidad
Proporción de familias sin teléfono
Proporción de familias sin refrigerador
Proporción de familias sin lavadora
Proporción de familias sin automóvil
Proporción de familias sin estufa de gas
Media de personas por familia
Media de personas que trabajan
Media de personas que reciben un ingreso
Media de niños entre 12 y 15 años que trabajan
Media de niños entre 5 y 15 que no asisten a la escuela
Proporción de familias con jefe de familia mujer
Edad media del jefe de familia
Media de años de educación del jefe de familia
Proporción de familias con jefe de familia analfabeta
Índice de hacinamiento
Proporción de jefes de familia asalariados
Proporción de jefes de familia campesinos o agricultores
Proporción de jefes de familia auto empleados
Proporción de jefes de familia empleados en un negocio familiar
Índice de dependencia (personas entre 1 y 15 años de edad y mayores a 65, sobre personas entre 16 y 64 años de edad)
Proporción de familias elegibles (utilizando una función discriminante que se correlaciona 0.99 con la utilizada por Oportunidades para definir la elegibilidad)
Media del puntaje pobreza (obtenido por análisis discriminante) para definir la elegibilidad al Programa
Entidad de residencia (como variables indicadoras)
Número de familias elegibles para Oportunidades
Cuadrado del número de familias elegibles para Oportunidades

* Áreas geoestadísticas básicas. Son unidades básicas utilizadas para agrupar datos en la información del Censo.



3.1.4. Conformación del panel (segunda etapa de selección)

Todas las familias que indicaron estar incorporadas a *Oportunidades* durante el cuestionario de tamizaje fueron incluidas en la muestra de evaluación. Sin embargo, sólo 3 645 hogares elegibles reportaron haber estado incorporados al Programa, muchos menos de los 6 000 esperados. A través de los registros administrativos del Programa se pudieron identificar 656 hogares más en las manzanas de intervención, que respondieron que no sabían o no estaban incorporados al Programa, cuando efectivamente sí lo estaban. Incluyendo a estos hogares, el número de hogares incorporados en la muestra de evaluación se incrementó a 4 301. Sin embargo, este número era todavía muy inferior al esperado. Con el fin de obtener una muestra cercana a la esperada (6 000 hogares elegibles incorporados), se decidió agregar familias incorporadas a *Oportunidades* localizadas en las manzanas adyacentes a las manzanas previamente seleccionadas para la muestra (llamadas manzanas de barrido). El estatus de incorporación fue provisto por los implementadores del Programa según los registros administrativos. Se identificaron todas las manzanas adyacentes y todos los hogares incorporados al Programa fueron incluidos en la muestra. Mediante este procedimiento se agregaron 1 178 hogares incorporados, por lo que la muestra de evaluación quedó conformada por 5 479 hogares incorporados (nótese que estos números se refieren al estatus de incorporación, mientras que los números en el cuadro 1 y la figura 1 se refieren a la condición de elegibilidad). Los hogares en las manzanas de barrido no formaron parte de la muestra probabilística

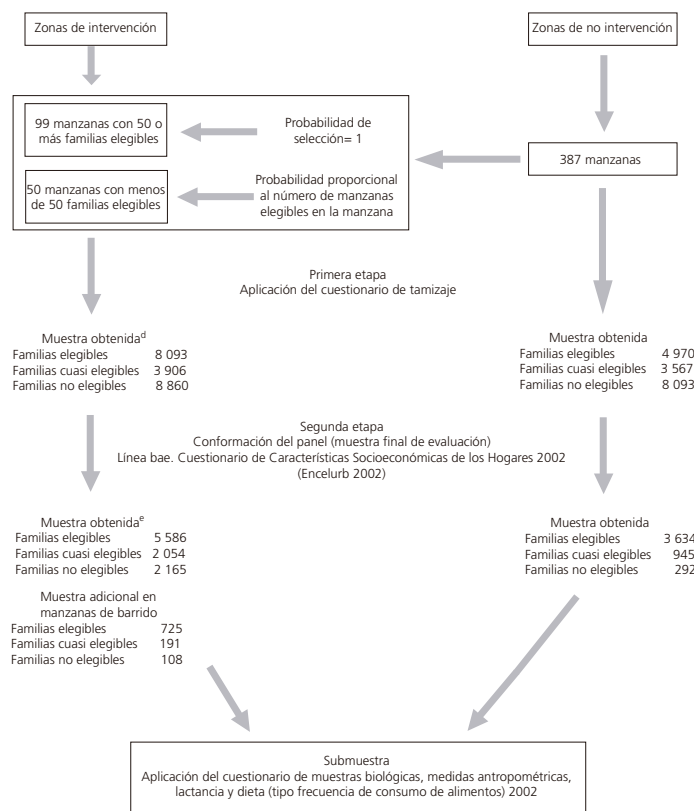


FIGURA 1
Esquema general de conformación de la muestra urbana

^d La clasificación de elegibilidad fue estimada con la información del Cuestionario de tamizaje

^e La clasificación de elegibilidad en esta etapa fue estimada con la información del Cuestionario de características socioeconómicas de los hogares 2002.



original en las zonas de intervención, y por lo tanto no respondieron al Cuestionario de tamizaje. Como consecuencia, no es posible estimar su estatus de elegibilidad con base en ese cuestionario.

La muestra de hogares cuasi elegibles (hogares localizados justo por encima del punto de corte que determina la elegibilidad al Programa) en zonas de intervención fue seleccionada a través de un muestreo sistemático con igual probabilidad de selección. Asimismo, se seleccionaron 2 000 hogares no elegibles de manera aleatoria simple en zonas de intervención.

En las zonas de no intervención, la muestra de hogares se seleccionó por medio de un muestreo sistemático de todos los hogares potencialmente elegibles para recibir los beneficios de *Oportunidades* y de hogares cuasi elegibles, todos residentes en las manzanas seleccionadas por el procedimiento de pareamiento.

Una vez identificada la muestra de evaluación, la segunda etapa consistió en la aplicación del Cuestionario de características socioeconómicas de los hogares 2002, y del Cuestionario de muestras biológicas, medidas antropométricas, lactancia y dieta (tipo frecuencia de consumo de alimentos), los cuales constituyeron la línea basal de la evaluación urbana. Con el fin de aprovechar la información recopilada en las manzanas de barrido, se volvieron a clasificar todos los hogares en elegibles, cuasi elegibles y no elegibles usando la información del Cuestionario de características socioeconómicas de los hogares 2002. En la figura 1 se observa la muestra obtenida en esta etapa bajo la nueva clasificación de elegibilidad. La composición de la muestra final se presenta en el cuadro 3.

Además de la muestra general de evaluación, se construyeron diversas submuestras para estudiar aspectos específicos de la evaluación. La selección de esas muestras se describe más adelante. Debido a los procedimientos empleados en la construcción del marco muestral, la inclusión de hogares de las manzanas de barrido en las zonas de intervención, y el pareamiento de manzanas en las zonas de no intervención, no es posible estimar la probabilidad de selección de todos los hogares en la muestra. Por tanto, la muestra final de evaluación y las submuestras derivadas no tienen representatividad nacional o regional específica, y no es posible estimar los ponderadores muestrales que permitirían hacer inferencias poblacionales.

CUADRO 3

Distribución de hogares por zona y estatus de elegibilidad* en la línea basal

ZONA	ELEGIBLES	CUASI ELEGIBLES	NO ELEGIBLES	CLASIFICACIÓN	
				FALTANTE**	TOTAL
Intervención	6 311	2 245	2 273	734	11 563
No intervención	3 634	945	292	767	5 638
Total de hogares	9 945	3 190	2 565	1 501	17 201
Número total de personas					76 002

* El estatus de elegibilidad se estimó a partir de la información del Cuestionario de características socioeconómicas de los hogares 2002

** Familias no clasificadas debido a falta de información en algunas variables utilizadas para estimar el índice de pobreza. Todos esos hogares reportaron encuestas incompletas



3.2 Grupos de comparación

El diseño de evaluación y la información sobre el estatus de elegibilidad de los hogares permiten identificar a los siguientes grupos de comparación para evaluar los impactos de *Oportunidades* en zonas urbanas (véase también cuadro 4):

- V. Grupo de intervención (grupo de tratamiento): muestra de hogares elegibles en zonas de intervención e incorporados en 2002 a *Oportunidades*
- VI. Grupo de control interno: muestra de hogares elegibles en zonas de intervención pero que no estaban incorporados al Programa
- VII. Grupo de control externo: muestra de hogares en zonas de no intervención que fueron clasificados como elegibles para recibir los beneficios del Programa
- VIII. Controles cuasi elegibles en zonas de intervención: muestra de hogares en zonas de intervención clasificados como cuasi elegibles debido a que su puntaje de pobreza se encontró justo por arriba del punto de corte que determina la elegibilidad al Programa

Adicionalmente, existe un grupo de comparación conformado por hogares no elegibles, con un puntaje de pobreza ampliamente por arriba del punto de corte que determina la elegibilidad a *Oportunidades* (es decir, por encima de las familias cuasi elegibles), residente en zonas de intervención que puede utilizarse para evaluar la focalización del Programa.

La sección 4 detalla los principales efectos de *Oportunidades* que pueden ser estimados con el uso de estos grupos de comparación.

3.3. Recolección de los datos

Los datos recolectados para la evaluación de impacto de *Oportunidades* en zonas urbanas incluyen una encuesta de línea de base (Encelurb 2002) y dos encuestas de seguimiento realizadas en 2003 y 2004 (Encelurb 2003 y 2004) (figura 2). En todos los casos se realizaron pruebas piloto de los instrumentos, y éstos fueron aplicados por personal altamente capacitado con experiencia previa en la aplicación de encuestas poblacionales. Toda la información se recolectó directamente en los hogares después de haber obtenido cartas firmadas de consentimiento informado por parte de las personas que respondieron. El protocolo, métodos e instrumentos de evaluación fueron aprobados por las comisiones de investigación,

GRUPO DE INTERVENCIÓN

Hogares elegibles en zonas de intervención que solicitaron su incorporación y fueron incorporados a *Oportunidades*

GRUPOS DE COMPARACIÓN

Control interno. Hogares elegibles no incorporados en zonas de intervención

Control externo. Hogares elegibles en zonas de no intervención

Cuasi elegibles. Hogares cuasi elegibles en zonas de intervención

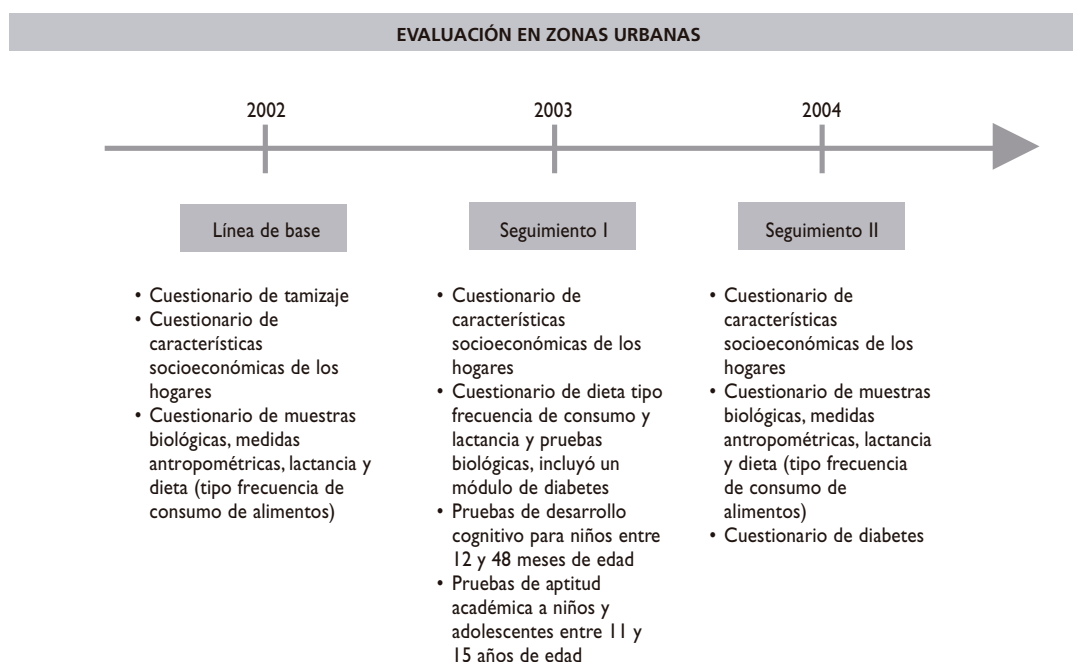
Hogares no elegibles en zonas de intervención (para análisis de la focalización del Programa)

CUADRO 4

Algunos grupos de comparación para la evaluación de impacto en zonas urbanas



FIGURA 2.
Diseño de evaluación
para zonas urbanas



ética y bioseguridad del INSP. Más detalles sobre la estructura y manejo de las bases de datos urbanas se encuentra en la Nota técnica general urbana.¹²

Los instrumentos utilizados para la recolección de los datos fueron:

3.3.1. Encelurb 2002 (línea de base)

La encuesta de línea de base (Encelurb 2002) se compuso de tres instrumentos:

- Cuestionario de tamizaje. Este cuestionario recolectó información socioeconómica de los hogares con el fin de calcular su puntaje de pobreza y determinar su estatus de elegibilidad al Programa. Se aplicó a todos los hogares en la muestra de evaluación, con excepción de los hogares residentes en las llamadas manzanas de barrido en zonas de intervención.
- Cuestionario de características socioeconómicas de los hogares. Este cuestionario recolectó información de los hogares y de sus miembros sobre variables sociodemográficas, de salud, educación, ocupación, consumo, gasto, ingreso y aspectos de género. Se aplicó a la muestra de evaluación seleccionada en la segunda etapa (conformación del panel) a partir de los resultados del Cuestionario de tamizaje. Este cuestionario fue respondido por personas mayores de 15 años que contaban con información sobre la familia y sus miembros. Algunas secciones del cuestionario fueron respondidas por residentes de entre 10 y 21 años de edad (para las secciones sobre el consumo de tabaco y alcohol) y por mujeres entre 15 y 49 años (para la sección sobre salud reproductiva).
- Cuestionario de muestras biológicas, medidas antropométricas, lactancia y dieta (tipo frecuencia de consumo de alimentos). Este cuestionario recolectó información sobre niños menores de cuatro años de edad y sus madres, y sobre mujeres embarazadas. Se recolectó información biológica sobre niños de 24 a 47 meses de edad, madres de niños de 0 a 47 meses de edad y mujeres embarazadas. La concentración de hemoglobina se midió utilizando un fotorreflectómetro portátil (HemoCue,



Angelholm, Suecia). Se obtuvieron medidas antropométricas (peso y talla) de los niños menores de cuatro años de edad y sus madres. Las medidas de peso en niños de 24 a 47 meses de edad fueron obtenidas utilizando básculas electrónicas Tanita con un margen de error de +/- 20 g. Para adultos (mujeres embarazadas y madres de niños menores de 47 meses de edad) se usaron básculas portátiles Solares. La talla de los niños (24 a 47 meses de edad) se midió utilizando infantómetros de PVC, y la de adultos con estadímetros portátiles. Se obtuvo información sobre lactancia materna y alimentación complementaria en niños menores a 35 meses de edad. El cuestionario de frecuencia de consumo de alimentos se aplicó a niños entre 6 y 35 meses de edad y a sus madres. El cuadro 5 resume las medidas y poblaciones incluidas en este cuestionario.

La recolección de datos de la Encelurb 2002 se llevó a cabo en dos etapas. En la primera etapa se recolectó información en las zonas de intervención entre septiembre 2 y noviembre 27 de 2002. En la segunda etapa se recolectó información en las zonas de no intervención de octubre 22 a diciembre 17 de 2002. Cada una de estas fases incluyó a su vez tres fases: 1) actualización geográfica y localización de todas las familias en el marco muestral, y la aplicación del Cuestionario de tamizaje; 2) aplicación del Cuestionario de características socioeconómicas de los hogares; y 3) aplicación del Cuestionario de muestras biológicas, medidas antropométricas, lactancia y dieta (tipo frecuencia de consumo de alimentos).

3.3.2. Encelurb 2003 (primer panel)

La encuesta Encelurb 2003 fue la primera encuesta de seguimiento aplicada a la muestra de evaluación seleccionada en la Encelurb 2002. Los instrumentos utilizados en la Encelurb 2003 fueron: 1) Cuestionario de características socioeconómicas de los hogares; 2) Cuestionario de dieta tipo frecuencia de consumo, lactancia y pruebas biológicas, con un módulo especial sobre diabetes; 3) pruebas de aptitud académica aplicadas a niños de entre 11 y 15 años de edad; y 4) pruebas de desarrollo cognitivo a niños de 12 a 48 meses de edad. En la Encelurb 2003 se utilizaron, prácticamente, la misma metodología y los mismos instrumentos que en la Encelurb 2002, salvo algunos cambios menores tras una revisión a cargo del GAE. Esta encuesta no recolectó información antropométrica ni muestras biológicas de madres, pero en su lugar incluyó un módulo especial sobre la calidad de la atención y el control glicémico en diabéticos mayores de 40 años de edad.

Con el fin de evitar una disminución excesiva de la muestra de hogares en la Encelurb 2003 (que por alguna causa ya no fuera posible levantarles información) con la consecuente pérdida de poder en los análisis estadísticos, se adicionaron a la muestra alrededor de 1 500 hogares elegibles y casi elegibles, tanto de las zonas de intervención como de no intervención. Estos hogares fueron seleccionados sistemáticamente de entre la muestra de hogares que completaron el cuestionario de tamizaje en 2002,

VARIABLE	POBLACIÓN
Hogares	4 770
Hemoglobina y/o antropometría	5 762 niños de 24-47 meses de edad
	4 787 mujeres embarazadas o madres de niños menores a 47 meses de edad
Lactancia materna y alimentación complementaria	3 704 niños de 0-35 meses de edad
Cuestionario de frecuencia de consumo de alimentos	3 351 niños de 6-35 meses de edad
	3 113 mujeres embarazadas o madres de niños menores a 36 meses de edad

CUADRO 5

Poblaciones incluidas en la recolección de muestras biológicas, medidas antropométricas, lactancia y dieta. Encelurb 2002



y que no se incluyeron en la segunda etapa de selección. La muestra final con información socioeconómica en la Encelurb 2003 quedó compuesta por 18 041 hogares y 77 764 personas.

La muestra para la aplicación del Cuestionario de dieta tipo frecuencia de consumo, lactancia y pruebas biológicas se presenta en el cuadro 6. Por razones técnicas,⁹ fue necesario redefinir la muestra de estudio, por lo que no todos los niños y sus madres incluidos en esta muestra formaron parte del componente biológico de la Encelurb 2002. De este modo, se recolectaron muestras de sangre capilar para determinar concentraciones de hemoglobina en niños de entre 12 y 35 meses de edad, que tenían de 0 a 23 meses de edad al momento de la aplicación de la Encelurb 2002. Esto se hizo con la intención de incluir a niños que tendrían entre 24 y 47 meses de edad en la encuesta de seguimiento en 2004. La selección de estos grupos de edad permite realizar análisis longitudinales y de corte transversal aprovechando el proceso de crecimiento de los niños. Se recolectó información sobre lactancia materna y dieta complementaria en niños menores de tres años de edad.

Para estudiar la calidad en la atención y el control glicémico en adultos diabéticos, se planeó una muestra de 500 individuos diabéticos mayores de 40 años de edad para hacer determinaciones de hemoglobina glicosilada. Con datos de la Encelurb 2002 se identificaron los hogares en los que se reportó la presencia de por lo menos una persona mayor de 40 años de edad con diagnóstico de diabetes mellitus. Se visitaron un total de 919 hogares, y 677 adultos completaron el cuestionario, de los cuales 502 (74%) tuvieron una determinación válida de hemoglobina glicosilada. El otro 26% (n=175) no pudo ser localizado o se rehusó a participar en el estudio.

La Encelurb 2003 incluyó la aplicación de una prueba de aptitud académica en niños y adolescentes de 11 a 15 años edad, en un intento por medir el impacto de su participación en el sistema escolar. La prueba seleccionada, llamada DAT-2, ha sido aplicada a adolescentes mexicanos y ofrece mediciones de razonamiento matemático y verbal. La aplicación de esta prueba asume que las aptitudes académicas pueden ser mejoradas mediante la asistencia a la escuela y el trabajo escolar, por lo cual puede ser un indicador de cómo el Programa está influyendo en el desarrollo de los adolescentes.

CUADRO 6

Poblaciones de estudio para la aplicación del Cuestionario de dieta tipo frecuencia de consumo, lactancia y pruebas biológicas. Encelurb 2003

VARIABLE	POBLACIÓN DE ESTUDIO
Hemoglobina	3 862 niños de 12-35 meses de edad que en la Encelurb 2002 tenían 0-23 meses de edad
Lactancia materna y alimentación complementaria	2 817 niños de 12-35 meses de edad que en la Encelurb 2002 tenían 0-23 meses de edad y niños de mujeres embarazadas detectadas en la 2002.
Dieta tipo frecuencia de consumo	3 857 niños de 18-47 meses de edad que tenían 6-35 meses de edad en la Encelurb 2002, niños de 12-18 meses de edad que tenían 0-6 meses de edad en la Encelurb 2002.
Se incluyeron también niños de más de 6 meses de edad, nacidos de mujeres embarazadas identificadas en la Encelurb 2002	2 975 madres de niños de 0-47 meses de edad en 2003.
Se visitó un total de 4 533 hogares	

⁹ No se trata de un panel de niños, ya que los que fueron detectados con anemia severa en la Encelurb 2002 (sean de zonas de intervención como de no intervención), fueron sujetos a tratamiento nutricional por parte del INSP, según lo dictan las normas éticas.



La prueba DAT-2 se aplicó a una submuestra de 3 556 niños de entre 11 y 15 años de edad que vivían en hogares de la muestra de evaluación urbana, y residentes en los estados del país donde se encontraba la mayor concentración de hogares con niños de 11 a 15 años de edad. La muestra final quedó compuesta por 2 103 niños en zonas de intervención y 1 453 en zonas de no intervención.

Por último, la Encelurb 2003 incluyó la aplicación de una batería de pruebas de desarrollo cognitivo para niños de 12 a 48 meses de edad, con el fin de analizar si el Programa tiene un impacto en las habilidades cognitivas y motoras de los niños. La batería incluyó las siguientes pruebas: la prueba de vocabulario Peabody, las escalas McCarthy de habilidades infantiles para medir habilidades físicas y motoras, la prueba de desarrollo cognitivo Woodcock-Johnson-Muñoz III, los hitos de desarrollo cognitivo de MacArthur (llamados CDIs) y el cuestionario MacArthur de salud y conducta para medir el desarrollo conductual. Las pruebas se aplicaron a un total de 1 961 niños de 12 a 48 meses de edad, de los cuales 1 305 vivían en zonas de intervención y 656 en zonas de no intervención.

La recolección de datos de la Encelurb 2003 se llevó a cabo del 1 de septiembre al 14 de noviembre de 2003, en las zonas de intervención y no intervención simultáneamente.

Encelurb 2004 (segundo panel)

La Encelurb 2004 constituye el segundo panel de la encuesta, y es un seguimiento a los hogares incluidos en la muestra de las Encelurb 2002 y 2003. La Encelurb 2004 utilizó como instrumentos: 1) el Cuestionario de características socioeconómicas de los hogares; 2) Cuestionario de dieta, pruebas biológicas, mediciones antropométricas (peso y talla), y toma de sangre venosa en niños; 3) Cuestionario de diabetes y toma de muestra de sangre capilar para determinaciones de hemoglobina glicosilada.

En la Encelurb 2004, la información socioeconómica y de muestras biológicas en niños y adultos se recolectó en dos fases: la primera incluyó solamente a los hogares en zonas de no intervención (comprendió del 21 de junio al 31 de agosto de 2004), y la segunda incluyó a los hogares en zonas de intervención (del 1 de septiembre al 12 de noviembre de 2004). Fue necesaria esta estrategia de dos fases debido a que estaba por iniciar un intensivo proceso de incorporación en las zonas de no intervención, por lo que era necesario obtener la información de estas zonas antes de que éste comenzara. La muestra final con información socioeconómica en 2004 quedó compuesta por 17 023 hogares y 72 421 personas.

La Encelurb 2004 utilizó los mismos instrumentos y metodología que las Encelurb 2002 y 2003, con algunas modificaciones de contenido después de una revisión por parte del GAE. Fue especialmente necesario modificar el módulo de educación para diferenciar entre asistencia escolar actual y asistencia escolar futura de los miembros del hogar. Se aplicó, además, junto con el cuestionario de características socioeconómicas de los hogares, un cuestionario de diabetes. El cuestionario de diabetes fue muy similar al módulo de diabetes utilizado en la Encelurb 2003, e igualmente pretendía obtener datos sobre la calidad en la atención y el control glicémico de la diabetes en personas mayores de 40 años de edad. Los procedimientos para la recolección de muestras para determinar hemoglobina glicosilada fueron idénticos a los utilizados en 2003. Es importante señalar que la muestra para este estudio de diabetes no necesariamente se conforma de las mismas personas participantes de la encuesta de diabetes de la Encelurb 2003, ya que por diversas circunstancias no todos volvieron a participar en el estudio, por lo que hubo necesidad de completar la muestra con residentes de hogares que mencionaron tener diabetes en la encuesta de características socioeconómicas de 2003. La muestra final quedó conformada por 598 personas con información de diabetes y determinaciones de hemoglobina glicosilada.

En cuanto a la información biológica, la Encelurb 2004 incluyó niños de 24 a 47 meses de edad y a sus madres, con información socioeconómica de la Encelurb 2002 y que fueron parte de la muestra de medidas biológicas en la Encelurb 2003. La recolección de datos se basó en el Cuestionario de dieta tipo frecuencia de consumo, lactancia y pruebas biológicas de la Encelurb 2003, a la cual se le añadió



un módulo de antropometría (talla y peso) similar al utilizado en la Encelurb 2002. El cuadro 7 muestra la conformación de la muestra para este levantamiento.

En la Encelurb 2004 el peso fue medido con un solo modelo de báscula electrónica con capacidad máxima de 14 kg y con una precisión de 20 g para niños de 0 a 23 meses, y una capacidad de 135 kg para mayores de dos años (modelo 1582 Mommy and Family, Tanita Corp., Arlington Heights, IL, EU). Las básculas fueron calibradas diariamente antes de comenzar el trabajo de campo. La talla y longitud se midió con infantómetros/estadímetros portátiles de madera con capacidad para 198 cm y precisión de 1 mm (Industrias Schorr, Glen Burney, MD, EU).

Asimismo, se recolectaron muestras de sangre venosa en una submuestra de niños de 6 a 47 meses de edad y mujeres que formaron parte de la fase final de un estudio especial sobre el estado nutricional y consumo de complementos alimenticios de mujeres y niños. La toma de muestra de sangre venosa se recolectó con el fin de obtener mediciones de las concentraciones séricas de ferritina, receptor soluble de transferrina y proteína C reactiva. La recolección de muestras biológicas se llevó a cabo del 23 de julio al 31 de agosto de 2004 en zonas de no intervención y del 1 de septiembre al 15 de octubre de 2004 en zonas de intervención. Todas las muestras fueron recolectadas por personal altamente capacitado y mediante procedimientos estándar.

Recolección de datos adicionales en 2005

En 2005 no se llevó a cabo una encuesta de seguimiento a la muestra general de evaluación urbana. No obstante, se recolectó información de seguimiento de dos estudios importantes para evaluar el impacto del Programa. El primero fue la aplicación, por segunda vez, de la prueba de aptitudes académicas (DAT-2) a los adolescentes de entre 12 y 17 años de edad que respondieron la prueba en 2003, a un grupo adicional de adolescentes de la muestra de evaluación, y a un grupo de adolescentes (grupo control) residentes en zonas donde el Programa aún no estaba disponible en 2005. Con el fin de compensar pérdidas potenciales en el seguimiento, se incluyeron en la muestra adolescentes que no habían respondido a la prueba DAT-2 en 2003. La conformación de la muestra final para este levantamiento se presenta en el cuadro 8.

CUADRO 7

Población de estudio para la aplicación del Cuestionario de muestras biológicas, medidas antropométricas, lactancia y dieta (tipo frecuencia de consumo). Encelurb 2004

VARIABLE	POBLACIÓN DE ESTUDIO
Muestra total	7,976 niños de 24 a 47 meses y sus madres que tuvieron medición basal en la Encelurb 2002 y medición en la encuesta de seguimiento Encelurb 2003
Hemoglobina, lactancia y antropometría	3 315 niños de 24-47 meses, se les midió antropometría, hemoglobina y se les preguntó condición de lactancia. 1 041 niños de 0-23 meses se les midió únicamente antropometría
Dieta tipo frecuencia de consumo	3 315 niños de 24-47 meses de edad
Prácticas de lactancia y alimentación complementaria	1 041 niños de 0-23 meses de edad
Dieta tipo frecuencia de consumo	3 385 madres de niños de 24-47 meses
Hemoglobina y antropometría	3 385 madres de niños de 24-47 meses



	NO INTERVENCIÓN	INTERVENCIÓN	TOTAL
Adolescentes de 12 -17 años de edad con pruebas DAT-2 en 2003 y 2005	732	1 049	1 781
Adolescentes de 12 -17 años de edad con pruebas DAT-2 en 2005, pero no en 2003	1 742	2 458	4 200
Total	2 474	3 507	5 981

CUADRO 8

Adolescentes que respondieron a la prueba de aptitudes académicas (DAT-2) en 2003 y 2005 en zonas de intervención y de no intervención

En forma similar, en 2005 se obtuvieron datos de seguimiento para las pruebas de desarrollo cognitivo aplicadas a niños de 12 a 48 meses de edad en 2003. Se aplicó la prueba Woodcock-Johnson-Muñoz III, la prueba Stroop para medir la función ejecutiva, la prueba de vocabulario Peabody y la prueba MacCarthy para medir el desarrollo físico. Asimismo, se midió la talla de los niños y se aplicó la prueba Pegboard, con el fin de contar con una medida de las habilidades motoras finas de los niños, y la prueba Achenbach para medir el desarrollo conductual de los niños. Se utilizaron computadoras personales para recolectar la información en campo. Se recolectó información de 5 717 niños, 4 034 en zonas de intervención y 1 683 en las zonas de no intervención.



La evaluación de impacto del Programa en zonas rurales inició en 1997. Parte del diseño de esta evaluación ha sido descrito detalladamente en otros documentos.^{1,27} La evaluación rural utilizó un diseño experimental en el cual las localidades fueron asignadas aleatoriamente a los grupos de intervención y control. El proceso para la asignación aleatoria de las localidades a los grupos de intervención y control fue como sigue: I) se seleccionó un conjunto de localidades rurales que satisficieran los requerimientos de inclusión al Programa; II) cada localidad fue aleatoriamente asignada al grupo de intervención o al grupo control; y III) mediante un criterio de pobreza basado en la Encuesta de Características Socioeconómicas de los Hogares Rurales (Encaseh 1997) se determinaron las familias elegibles de recibir los beneficios del Programa en cada localidad previamente asignada al grupo de intervención (localidades asignadas como beneficiarias). Así se conformó el grupo de localidades que participarían inicialmente en la evaluación de *Oportunidades*, también llamada la muestra original. Este diseño sirvió para la evaluación a corto plazo del impacto del Programa en zonas rurales; se tomó como línea de base la información de la Encaseh 1997 y la primera ronda de la Encuesta de Evaluación de los Hogares Rurales (Encel 1998), y se aplicaron encuestas panel subsecuentes cada seis meses de 1998 a 2000. Para la evaluación de mediano plazo, en 2003 se regresó a la localidades y se les aplicó la Encuesta de Evaluación de los Hogares Rurales 2003 (Encel 2003).

El objetivo de las Encel fue recolectar datos de los hogares y sus miembros sobre distintos temas de interés para el Programa. Las Encel se conforman de diferentes cuestionarios que abordan temas específicos. Si bien las preguntas pueden variar de una encuesta a otra, dependiendo del periodo de tiempo pertinente para observar resultados atribuibles a *Oportunidades* en ciertos indicadores, siempre se buscó mantener la comparabilidad de las preguntas a lo largo del tiempo. El principal instrumento de recolección de datos en las Encel 1998-2003 fue el Cuestionario de características socioeconómicas de los hogares. Este cuestionario, al igual que el aplicado en las zonas urbanas, incluye secciones sobre características sociodemográficas del hogar, educación, salud, ocupación, consumo e ingreso, entre otros temas.

El tamaño de la muestra original para la evaluación en zonas rurales incluyó 330 localidades de intervención, de un universo de 4 546 localidades; y 191 comunidades de control, de un universo de 1 850 localidades. Después de los ajustes finales, la muestra final quedó definida en 320 localidades de intervención, donde el Programa operó inicialmente; y 186 comunidades control, donde *Oportunidades* inició operaciones posteriormente.

Para evaluar los impactos a mediano plazo se aplicó la Encel 2003. Esta encuesta recolectó información en las 506 localidades que conformaron la muestra original de evaluación. Para ese momento, todas las localidades que habían participado en las Encel anteriores habían sido incluidas en el Programa (320 se incorporaron en 1998 y 186 en el año 2000). Por esta razón, se decidió crear un nuevo grupo de comparación que no participara en el Programa, por lo que se agregaron a la muestra 151 localidades que no habían sido incluidas en él (figura 3). Más detalles sobre la estructura y manejo de las bases de datos rurales se encuentra en la Nota técnica general rural.¹³

El reto en esta nueva fase de la evaluación rural era conformar un nuevo grupo control con hogares que fueran lo más similares posible a aquellos que ya estaban recibiendo los beneficios del Programa, y poder cuantificar los efectos de *Oportunidades* después de seis años de intervención.

Para la construcción de este nuevo diseño, algunas localidades de un conjunto de 14 000 localidades potenciales donde aún no operaba el Programa fueron pareadas a cada localidad de la muestra original de evaluación (506). Para hacer el pareamiento de localidades, se obtuvo información del Censo 2000

IV.

Diseño y recolección de datos en zonas rurales

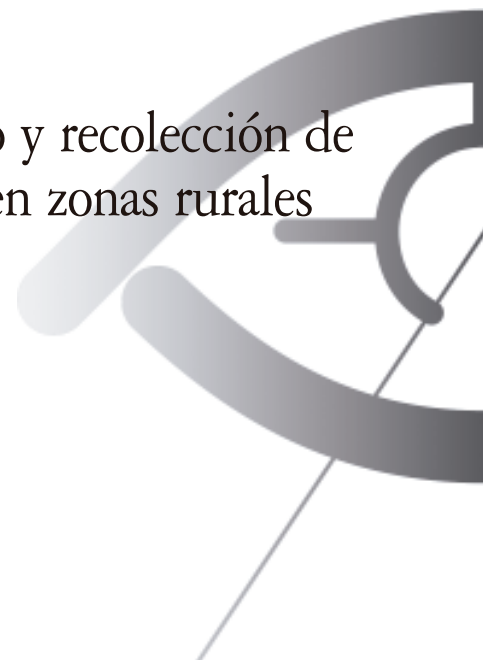
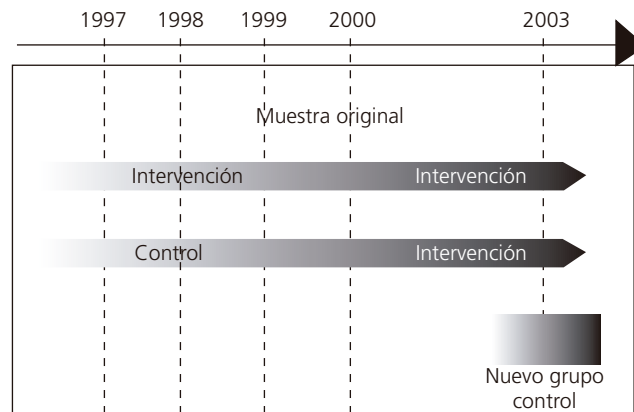




FIGURA 3
Encuestas levantadas para la evaluación de impactos de *Oportunidades* en zonas rurales



sobre diversas características de los hogares en esas localidades. Se incluyeron características sobre los miembros de los hogares, tales como edad, condición laboral y de escolaridad, etc.; así como características de la vivienda: número de cuartos, tipo de piso, si tenían electricidad, etc. Adicionalmente, se incluyeron variables demográficas de la localidad. Las comunidades control pareadas fueron seleccionadas sobre la base de mayor similitud en términos de esas características observables. La estrategia de pareamiento fue:

- I. Las localidades pareadas debían pertenecer a la misma entidad federativa que las localidades de intervención. En el caso de Querétaro, donde hubo pocas localidades control, se eligieron localidades de los estados vecinos (San Luis Potosí o Hidalgo).
- II. Para maximizar la comparabilidad entre las localidades en términos de los servicios de salud y educación disponibles, se construyeron, para cada localidad de intervención, las distancias a la escuela primaria, secundaria, y clínica de salud más cercanas. Para cada estado se obtuvo el máximo de estas distancias; con esta información se eliminaron las localidades del conjunto inicial de 14 000 que tenían distancias a los servicios de educación y salud más cercanos, mayores a la distancia máxima estimada. Aquellas localidades con distancias menores o iguales a la distancia máxima fueron consideradas potencialmente como localidades control.
- III. También se excluyeron aquellas localidades potencialmente de control con 25% o más de familias incorporadas al Programa.
- IV. Una vez aplicadas las restricciones de selección, el grupo de localidades potenciales a ser pareadas (5 190) se les implementó la estrategia de pareamiento a través de un puntaje de propensión, que refleja la probabilidad de la localidad de tener el Programa dada una serie de características observadas. El puntaje de propensión se estimó por un modelo logístico y se hizo para cada entidad federativa. Con la probabilidad estimada de participación, se parearon localidades de intervención con localidades de control considerando el valor absoluto de la diferencia entre sus probabilidades estimadas de tener el Programa. El pareamiento se hizo por el método del vecino más cercano y con reemplazo (con un máximo de 3 o 4 reemplazos). Otros documentos ofrecen discusión sobre esta metodología, una descripción más detallada de las variables incluidas en el puntaje de propensión y la selección de las localidades incluidas en el nuevo grupo control.^{14,15}

La integración de un nuevo grupo control permite la conformación de tres grupos de comparación para la evaluación a mediano plazo de *Oportunidades* en áreas rurales:



- I. Grupo de intervención temprana: hogares elegibles residentes en las localidades de intervención temprana (1998) y participantes en el Programa.
- II. Grupo de intervención tardía: muestra de hogares elegibles y participantes en el Programa residente en las localidades que fueron incorporadas en el año 2000.
- III. Grupo de control externo: Muestra de hogares elegibles residentes en las localidades del nuevo grupo control donde aún no se ofertaba el Programa en 2003.

La Encel 2003 constituye el séptimo panel de la encuesta de seguimiento en áreas rurales, y utilizó como principal instrumento para la recolección de datos el Cuestionario de características socioeconómicas de los hogares, que incluyó un cuestionario de fecundidad que fue aplicado a mujeres de 15 a 49 años de edad. El objetivo del Cuestionario de características socioeconómicas de los hogares fue recabar información sobre diversos temas vinculados con el hogar en su conjunto y sus integrantes, tales como aspectos demográficos del hogar, condiciones de salud, educación, actividad laboral, características de la vivienda, consumo, transferencias, y un pequeño módulo de migración. La Encel 2003 fue estructurada de acuerdo con las temáticas de las anteriores Encel, procurando conservar la comparabilidad de las preguntas a lo largo del tiempo.

Para el levantamiento de la información de fecundidad se hizo una selección de los tres tipos de localidades: las que recibían el Programa desde 1998, las que lo hacían desde 2000 y las del nuevo grupo control. Los tres grupos se dividieron en siete estratos, que corresponden a cada uno de los estados de la República en donde se levantó la encuesta. Posteriormente se hizo una selección de localidades y de hogares considerando la distribución porcentual de mujeres en edad fértil por estado y por localidad. El procedimiento de selección terminó a nivel de hogar encuestándose a todas las mujeres en edad fértil de cada hogar seleccionado. La selección de la muestra de mujeres fue realizada por el Consejo Nacional de Población (Conapo).¹⁶ La Encel 2003, con excepción del componente biológico, fue levantada por personal de encuestas de la Coordinación Nacional del Programa Oportunidades.

Como parte de la recolección de la Encel 2003, el INSP se encargó de planear y llevar a cabo el componente biológico de esta encuesta. La recolección de datos se llevó a cabo en todas las 506 localidades de intervención y las 151 localidades de no intervención. Consistió en la recolección de muestras biológicas, medidas antropométricas (talla y peso), pruebas de desarrollo cognitivo y cuestionarios a los centros de salud. La recolección de datos biológicos incluyó varias pruebas y cuestionarios. El cuadro 9 muestra las pruebas y las mediciones que se aplicaron a cada población de estudio.

Los datos adicionales incluidos en los cuestionarios fueron:

- Cuestionario para niños: secciones relacionadas con el estado de salud y desarrollo conductual de los niños
- Cuestionario para adolescentes: secciones relacionadas con comportamientos de riesgo, estrés, depresión, estado de salud, patrones alimenticios, e historial académico y laboral
- Cuestionario para adultos: secciones relacionadas con patrones alimenticios, conductas de riesgo asociadas a enfermedades crónicas, estado de salud, estrés, y morbilidad

Se consideró importante recolectar datos no sólo a nivel de hogar, sino también datos que pudieran describir las características económicas, sociales, demográficas y de infraestructura que pudieran afectar los resultados de los indicadores de interés. Por tanto, se aplicó también un cuestionario dirigido a la localidad como parte de la Encel 2003. Este cuestionario se levantó en todas las localidades de la muestra de evaluación, uno por localidad. La entrevista se aplicó a autoridades locales, líderes o funcionarios públicos que pudieran proporcionar información relevante. El objetivo de este cuestionario era recolectar información sobre la infraestructura local, la disponibilidad de servicios, principales actividades económicas, y las organizaciones, agrupaciones o asociaciones que se reúnen para realizar



CUADRO 9
Pruebas biológicas
aplicadas a cada
población de estudio

PRUEBAS	NIÑOS DE 2 A 6	ADOLESCENTES DE 15	MADRES DE LOS	ADULTOS MAYORES
	AÑOS DE EDAD	A 21 AÑOS DE EDAD	NIÑOS EN LA MUESTRA	A 30 AÑOS DE EDAD
Saliva	•		•	
Anemia	• (H)	• (P/A)	• (H)	
Talla y peso	•	•	•	•
Glucosa				•
Embarazo		•		
General de orina		•	•	•
Pulso	•		•	•
Presión arterial			•	•
Prueba de esfuerzo				•
Muestras		Orina en papel filtro	Sangre venosa	Sangre venosa

* Niños de 24 a 36 meses de edad; H: HemoCue; P/A: Presencia/Ausencia

actividades comunitarias. Se aplicó además, un cuestionario de precios para obtener información sobre los precios de los principales productos alimenticios, agrícolas y farmacéuticos en cada localidad de la muestra.

El Programa tiene la intención de evaluar el impacto de *Oportunidades* en indicadores de demanda de servicios educativos, tales como asistencia escolar, tasas de deserción, reprobación, reinscripción, etc. Asimismo, se busca también analizar las condiciones de calidad predominantes en la provisión de educación primaria, secundaria y de nivel medio superior. Con el fin de obtener información para evaluar estos aspectos, la Encel 2003 incluyó un cuestionario para las escuelas que fue respondido por el director y dos maestros en cada una. Esta información pretende documentar las características de los servicios educativos que reciben los miembros de las familias. Para seleccionar la muestra de escuelas, se dio prioridad a las escuelas donde asistían la mayoría de los beneficiarios (o potenciales beneficiarios en caso de las localidades control) del Programa. Se seleccionaron las escuelas a las que dichos sujetos asistían y que se encontraban en un radio de 10 km de cualquiera de las 506 localidades. La ponderación muestral se realizó tomando en cuenta el número de beneficiarios por escuela, de modo que cada beneficiario tuviera la misma probabilidad de que su escuela fuera seleccionada. En el caso de las localidades control no fue posible tomar en cuenta la asistencia escolar de los beneficiarios, ya que no había familias incorporadas, por lo que el único criterio de selección fue la distancia. Las escuelas más cercanas a estas localidades tuvieron una mayor probabilidad de ser elegidas. La muestra final en las zonas de intervención quedó conformada por 190 escuelas primarias, 337 escuelas secundarias y 177 escuelas de nivel medio superior. La muestra para las nuevas localidades de control consistió de 74 escuelas primarias, 120 escuelas secundarias y 68 escuelas de nivel medio superior.

La recolección de datos de la Encel 2003 se llevó a cabo entre el 22 de septiembre y el 4 de diciembre de 2003. La recolección de datos biológicos se llevó a cabo de julio a octubre de 2003.

Como parte de la encuesta de evaluación rural de mediano plazo, en 2003 se levantó información relacionada al desarrollo cognitivo en niños y desempeño académico en adolescentes, así como pruebas biológicas en adultos, con el fin de obtener datos que documenten si los niños y jóvenes que han ingresado al Programa en diferentes etapas de su vida pueden tener diferentes respuestas a los beneficios del mismo. El levantamiento se llevó a cabo en 676 localidades, que incluyó a las 506 localidades seleccionadas inicialmente en 1997 (zonas de intervención) y 170 localidades adicionales de comparación (nuevas zonas control).



Al total de hogares identificados en esas zonas (41 833) se les aplicó un cuestionario de tamizaje; se seleccionaron aquellos hogares que cumplieran con algunos de los siguientes criterios: 1) viviendas con niños de 2 a 6 años de edad y 2) viviendas con adolescentes entre los 15 y 21 años. El tamaño de muestra planeado fue de 9 000 niños en el grupo de 2 a 6 años de edad (3 000 en localidades que se incorporaron en 1998, 3 000 en comunidades que se incorporaron en 2000, y 3 000 de las comunidades del nuevo grupo control); la misma distribución fue aplicada para la muestra de adolescentes entre 15 y 21 años de edad. Adicionalmente, se incluyó un componente biológico a 6 000 adultos mayores de 30 años (3 000 en cada una de las áreas de intervención).

En los hogares con niños de 2 a 6 años se recolectó información de muestras biológicas para identificar marcadores biológicos de estrés en binomios madre-hijo. La prueba de estrés constó de tres muestras de saliva para la determinación de cortisol, una de las hormonas esenciales del eje hipófisis-hipotálamo-adrenal. El cortisol de la saliva es un indicador de situaciones de estrés a corto plazo, está directamente asociado con la función del sistema inmune y tiene implicaciones críticas para la salud física y mental a largo plazo. A todos los niños se les aplicaron las pruebas de desarrollo cognitivo de McCarthy y Woodcock-Johnson, la prueba de McArthur fue aplicada en los niños de 2 a 3 años de edad, la cual fue reemplazada por la prueba de Peabody en los niños de 3 a 6 años. La concentración de hemoglobina fue medida por Hemocue, asimismo se obtuvieron medidas antropométricas de peso y talla, y se les midió el pulso.

En el caso de las madres se les aplicó la prueba de Peabody, se hicieron determinaciones de hemoglobina para detectar anemia por Hemocue, de la misma manera que se realizó en el niño. Se obtuvieron medidas antropométricas de peso y talla, medición del pulso, un examen general de orina,^h medición de la tensión arterial y se tomaron muestras de sangre venosa. Cuando la mujer se encontró en ayunas se tomaron dos muestras (una para suero y otra de sangre total), y en los casos en los que no se encontró en ayunas se tomó sangre total para la determinación de hemoglobina glicosilada. En los casos de las madres adolescentes, se les aplicaron pruebas de embarazo con tiras reactivas.

Al resto de los adultos mayores de 30 años en el hogar, residentes de las localidades de intervención temprana o tardía, se le realizaron las siguientes determinaciones biológicas y antropométricas: peso, talla, pulso, presión arterial con prueba de esfuerzo, examen general de orina, muestras de sangre venosa de acuerdo con los criterios señalados para las madres, y se determinó glucosa en sangre capilar. En la prueba de esfuerzo se le solicitó al individuo que realizará el mayor número de sentadillas posibles, siempre y cuando la persona pudiera hacerlas. Antes y después de la prueba se les tomó el pulso y la tensión arterial para determinar los cambios registrados.

En los hogares con adolescentes de 15 a 21 años de edad, se les realizó una entrevista los adolescentes y a los adultos mayores de 30 años de las comunidades de intervención 1998 y 2000. Las pruebas para adolescentes incluyeron la de Woodcock-Johnson en tres esferas y el cuestionario de sexualidad. La información biológica y antropométrica recolectada incluyó determinaciones de anemia, mediciones de peso y talla, y toma de muestras de orina y sangre capilar en papel filtro. A las adolescentes se les realizó una prueba de embarazo. Al igual que en los hogares con niños, los adultos mayores de 30 años residentes de las áreas de intervención temprana o tardía fueron incluidos al estudio, y se les recolectó la misma información biológica que en esos hogares. El cuadro 10 muestra el resumen de las mediciones en las distintas poblaciones objetivo.

La muestra final quedó compuesta por 8 845 niños de 2 a 6 años de edad (6 411 en zonas de intervención y 2 434 en nuevas zonas control); 7 525 jóvenes de 15 a 21 años de edad (5 412 en zonas de intervención y 2 113 en zonas control); y 12 932 adultos mayores de 30 años (9 496 en zonas de intervención y 3 436 en zonas de control).

^h El examen de orina fue medido con las limitaciones que señala el manual correspondiente.



CUADRO 10
Pruebas biológicas
aplicadas a cada
población de estudio

PRUEBAS	NIÑOS 2-6	NIÑOS 2-3	NIÑOS 3-6	ADOLESCENTES	MADRES	ADULTOS
McCarthy	•					
Peabody			•		•	
Woodcock Johnson	•			•		
McArthur		•				
Saliva	•				•	
Anemia	• (H)			• (P/A)	• (H)	
Peso y talla	•			•	•	•
Glucosa						•
Embarazo				• (M)		
Examen general de orina				•	•	•
Pulso	•				•	•
Tensión arterial					•	•
Prueba de esfuerzo						•
Muestras				Orina en papel filtro	Sangre venosa	Sangre venosa

* H: HemoCue; P/A: Presencia/Ausencia; M: mujeres

En esta sección se discuten los métodos estadísticos y econométricos utilizados en el marco de los diseños urbano y rural. Los retos metodológicos que surgieron en ambos diseños fueron diferentes, debido a que la evaluación rural inicial se fundamentó en la asignación aleatoria del Programa entre las localidades (aunque la muestra de control adicional de 2003 no fue asignada así), y por otra, las diferencias en el proceso de incorporación de los beneficiarios. Por lo tanto, se abordará primero la evaluación rural y posteriormente la urbana. Antes de entrar en detalles de las dos evaluaciones, se hará un breve resumen de algunas de las técnicas relevantes en los dos contextos a revisar. Adicionalmente a los métodos que se han aplicado en el trabajo de evaluación de *Oportunidades* realizado hasta la fecha, se abordarán tres técnicas de evaluación que aún no han sido utilizadas (o se han utilizado fuera de este contexto), pero que podrían ser aplicadas: análisis de regresión discontinua, método de variables instrumentales y modelos estructurales.

V. Métodos de evaluación



5.1. El problema de evaluación

Para establecer una notación unificada, consideremos un resultado de interés Y , donde Y_1 sea el resultado que un individuo alcanzaría si recibe los beneficios del Programa y Y_0 el resultado alcanzado sin el Programa. Por supuesto Y_1 y Y_0 no pueden ser observadas para un mismo individuo simultáneamente. Por consecuencia, será imposible establecer el impacto del Programa definido como $Y_1 - Y_0$ para un individuo en particular. Por este motivo, la mayoría de la literatura se enfoca en la estimación de los impactos promedio.

Cuando un programa se asigna aleatoriamente, los impactos promedio pueden estimarse fácilmente a partir de la estimación de resultados en participantes y no participantes. Por diseño, en una asignación aleatoria, la participación en el Programa no se relaciona con el resultado de interés y las dos poblaciones son estadísticamente idénticas. El problema de evaluación surge fundamentalmente en presencia de una asignación no aleatoria, donde las diferencias en los resultados promedio entre beneficiarios y no beneficiarios pueden reflejar diferencias entre ambas poblaciones más allá de la presencia del programa.

5.2. Parámetros de interés

Se comenzó el análisis con la definición de los parámetros a estimar. Se definió el efecto del programa para un individuo en particular como

$$\Delta = Y_1 - Y_0$$

donde Y_1 es el resultado potencial en presencia del tratamiento y Y_0 el resultado potencial sin tratamiento. Como sólo se observó un resultado potencial para cada individuo, se deben estimar algunos momentos de la distribución de los efectos del Programa.

Varios parámetros pueden ser de interés para el evaluador de políticas. Este documento se enfoca hacia dos parámetros: el Efecto Promedio del Tratamiento sobre los Tratados (EPTT) y el Efecto Promedio de la Intención al Tratamiento (EPIT). El EPTT es el efecto promedio del Programa para los individuos tratados, mientras que el EPIT es el efecto promedio del Programa en la población elegible. Ambos



parámetros son importantes para el diseño de políticas: el primero indica la efectividad del Programa para aquellos que lo reciben, mientras que el segundo mide el efecto del Programa en los beneficiarios posibles, lo cual es sumamente importante para el análisis de políticas, tanto en situaciones en que los diseñadores de políticas pueden influir en la participación (por ejemplo, con una campaña informativa) como en situaciones donde esto es más difícil (simplemente porque algunos individuos pueden no estar interesados en el Programa).

Si definimos un indicador $T=1$ si el individuo vive en una zona de intervención y $P=1$ si el individuo es participante, el EPTT y el EPIT se definen como:

$$\begin{aligned} \text{EPIT}(X) &= E[\Delta | T=1, X] = E\{Y_1 | T=1, X\} - E\{Y_0 | T=1, X\} \\ \text{EPTT}(X) &= E[\Delta | T=1, P=1, X] = E\{Y_1 | T=1, P=1, X\} - E\{Y_0 | T=1, P=1, X\} \end{aligned}$$

donde la dependencia en el vector de variables observables X pone énfasis en que estos parámetros pueden ser heterogéneos en dimensiones observables. Todas las esperanzas en las expresiones de arriba son condicionales al vector de variables X . Por lo tanto, los parámetros son estimados para individuos con ciertas características (X). Promediando para los diferentes valores de X obtenemos el EPIT y el EPTT para el grupo de interés.ⁱ

5.3. Método experimental. Identificación y estimación de los parámetros de interés con asignación aleatoria y tasa de participación de 100%

Cuando todos los individuos a los que se les ofrece la participación deciden participar en el programa, el EPIT y el EPTT coinciden (es decir, todos los individuos con $T=1$ también tienen $P=1$ y todos los sujetos con $T=0$ también tienen $P=0$). Si además, la asignación de las zonas (intervención y control) al Programa se realiza de manera aleatoria, los parámetros de interés pueden ser identificados asumiendo que el EPT en ausencia del Programa es igual al resultado promedio observado en las zonas de control, es decir, $E\{Y_0 | T=1, X\} = E\{Y | T=0, X\}$. Esto ocurre si no se presentan efectos indirectos en individuos dentro de las zonas de control. Entonces,

$$\text{EPIT} = \text{EPTT} = E\{Y_1 | T=1, X\} - E\{Y_0 | T=1, X\} = E\{Y | T=1, X\} - E\{Y | T=0, X\}.$$

Para estimar este parámetro sólo es necesario calcular la diferencia en los resultados promedio observados entre los individuos de zonas de intervención y en zonas de control.

En el caso de *Oportunidades*, éste fue el método de evaluación utilizado para determinar el efecto a corto plazo del Programa en zonas rurales, donde las localidades de tratamiento y control fueron asignadas aleatoriamente, y la tasa de participación entre los hogares elegibles fue prácticamente universal. Como se ha mencionado anteriormente, los datos para la evaluación rural a mediano plazo (que incluye la

ⁱ Otros parámetros que con frecuencia se mencionan en la literatura son el Efecto Promedio del Tratamiento (EPT) y el Efecto Promedio del Tratamiento en los no Tratados (EPTNT). El primero es el promedio del parámetro Δ sobre toda la población de interés.

$$\text{EPT} = E[\Delta] = E\{Y_1\} - E\{Y_0\}$$

Resulta útil definir estos parámetros condicionados a un conjunto X de variables observables. De modo que el EPT condicional es:

$$\text{EPT}(X) = E[\Delta | X] = E\{Y_1 | X\} - E\{Y_0 | X\}$$

el EPT puede entonces obtenerse a partir del EPT(X), calculando el promedio para los diferentes valores de X . El EPTNT es el efecto que el programa tendría en aquellas personas que no han sido tratadas.

$$\text{EPTNT}(X) = E[\Delta | T=0, X] = E\{Y_1 | T=0, X\} - E\{Y_0 | T=0, X\}$$



Encel 2003) no son aleatorios, por lo que se requieren métodos de evaluación no experimentales descritos más adelante.

El caso del diseño urbano de evaluación presenta, además, el problema de que sólo alrededor de 50% de los hogares elegibles decidieron participar en el Programa. Los métodos de identificación y estimación que a continuación se presentan atienden la problemática de la falta de aleatoriedad y las bajas tasas de participación.

5.4. Métodos no experimentales para identificar y estimar EPIT

Cuando la asignación al Programa se realiza en forma no aleatoria, debemos cuidar que las comparaciones entre participantes y no participantes sean originadas por la participación en el Programa y no por diferencias sistemáticas y preexistentes entre ambos grupos. Tales diferencias pueden surgir, por ejemplo, si las personas que viven en zonas donde el Programa está disponible tienen características diferentes, algunas de ellas relacionadas con el programa. En el caso de *Oportunidades*, las zonas con mayor densidad de pobreza fueron las primeras en ser incorporadas al Programa.

Una de las ventajas de centrarse en el parámetro EPIT es que no se requiere tomar en cuenta la autoselección de las personas al programa. Sin embargo, aún es necesario considerar la no aleatoriedad en la asignación del Programa en la medida en que existan diferencias sistemáticas entre las zonas donde se ofreció el Programa y las zonas donde no se ofreció.

A continuación se exponen los métodos que pueden abordar el problema de las diferencias entre los grupos de participantes y no participantes en el caso en que las diferencias sistemáticas se deban totalmente a características observables o a características no observables fijas en el tiempo. Posteriormente, se discutirá cómo se combinan ambos casos.

5.4.1. Diferencias en diferencias

El estimador de diferencias en diferencias (DID, por sus siglas en inglés) es comúnmente utilizado en el trabajo de evaluación de programas. Mide el impacto de un programa a través del cambio en los resultados de individuos en zonas de tratamiento, expuestos al programa, con una medición previa y otra posterior a la intervención, y se compara con el cambio de los resultados en individuos que viven en zonas de control. De este modo, el estimador elimina las fuentes, que no varían en el tiempo, que originan las diferencias entre los grupos de tratamiento y control.

El estimador de DID requiere datos previos y posteriores al Programa (datos t y t') de individuos en zonas de tratamiento y de control. Dichos datos están disponibles tanto para las muestras rurales como las urbanas. Asumimos que la relación entre los resultados de interés y las características observables y no observables se puede expresar como:

$$\begin{aligned} Y_{1t} &= X\beta_1 + U_{1t} \\ Y_{0t} &= X\beta_0 + U_{0t} \end{aligned}$$

donde β_1 y β_0 son las matrices de coeficientes que se refieren al efecto de los elementos de X en los respectivos resultados (sea Y_{1t} ó Y_{0t}). Sea $T_t=1$ si a un individuo se le ha ofrecido el tratamiento por un tiempo t . Sea para los individuos con Programa en las zonas de tratamiento $T_t=1$ y $T_{t'}=0$ para un periodo t' previo al Programa. Para aquellos en zonas de control, $T_t=0$ y $T_{t'}=0$. El modelo para la diferencia en los resultados de interés entre el tiempo t y el tiempo t' puede expresarse como:



$$Y_t - Y_{t'} = T_t (X\beta_1 - X\beta_0 + E(U_1 - U_0 | T_t = 1, X)) + U_{0t} - U_{0t'} + T_t \{U_{1t} - U_{0t} - T_t (E(U_{1t} - U_{0t} | T_t = 1, X))\} \\ = T_t T T(X) + \{U_{0t} - U_{0t'} + T_t \{U_{1t} - U_{0t} - T_t (E(U_{1t} - U_{0t} | T_t = 1, X))\}\},$$

donde el término en corchetes es el residual. El estimador puede ser justificado bajo una estructura de error de efectos fijos, tal que:

$$U_{0t} = f + v_{0t} \\ U_{1t} = f + v_{1t}$$

donde f es un efecto fijo (que no varía con el tiempo para una persona determinada y que mide características tales como habilidad innata, genética o salud) y v_t es un *iid* término de error con media cero.

Es posible obtener el estimador DID para el EPIT a partir de mínimos cuadrados ordinarios del cambio en los resultados en una variable que indique si pertenece a la zona de intervención o de control, interactuada con las variables X .

5.4.2. Pareamiento

El pareamiento es un método de evaluación no experimental que se utiliza ampliamente en estadística y economía para evaluar los efectos de las intervenciones de un programa. Los métodos de pareamiento han demostrado ser efectivos en contextos donde hay datos abundantes para modelar el proceso de participación en el programa (ver Heckman, Ichimura y Todd, 1997).¹⁵ El método compara los resultados de individuos o familias que participan en el Programa con los de individuos o familias similares (en términos de características observables) “pareados” que no participan. Las parejas se eligen sobre la base de un conjunto de características observables en las familias y los individuos. Por ejemplo, un niño de 12 años de una familia participante puede ser comparado con un niño de la misma edad, similar en términos de educación de los padres, bienes, ingreso, ubicación geográfica y composición familiar, pero cuya familia no participa en el programa.

En la evaluación de impactos de *Oportunidades* se han utilizado estimadores de pareamiento por puntaje de propensión. Estos estimadores fueron desarrollados en Rosenbaum y Rubin (1983).¹⁷ La idea detrás del uso de los estimadores por puntaje de propensión es que reducen a una sola dimensión todas las dimensiones en las que se basa el pareamiento de individuos. Ésta es una importante ventaja, dado que en el momento del análisis sólo se requiere ajustar a una dimensión en lugar de un conjunto de covariables.

Sea Y_1 el resultado de interés para individuos que reciben el tratamiento y Y_0 el resultado sin tratamiento. Como se ha dicho anteriormente, $T=1$ si la persona vive en zonas de tratamiento y $T=0$ en caso contrario. Asimismo, sea X el conjunto de características del individuo utilizadas como variables condicionantes. Se asume que la presencia o no del tratamiento no afecta la distribución de X . Sea $P(X) = \Pr(T=1|X)$ la probabilidad condicional de una familia para recibir el Programa. $P(X)$ es el puntaje de propensión al Programa.

Otros estimadores de pareamiento difieren en términos de los supuestos necesarios para justificar su aplicación y de los métodos utilizados para el pareamiento de individuos u hogares. En general, pueden clasificarse en dos tipos de estimadores: I) los estimadores de pareamiento de corte transversal (CT) comparan los resultados de participantes y no participantes, donde los resultados son medidos en un tiempo posterior al Programa, II) los estimadores de pareamiento por DID comparan el cambio en los resultados de los participantes en el Programa con el cambio en los resultados de los miembros de un grupo control, donde el cambio es medido en relación con los resultados en un tiempo previo al Programa.



5.4.3. Combinación del método de DID y pareamiento

La ventaja de utilizar un estimador DID en lugar de un estimador de corte transversal es que el primero permite que los resultados sean afectados por diferencias no observables invariantes en el tiempo entre los resultados de participantes y no participantes, que puedan surgir, por ejemplo, por diferencias en los mercados laborales regionales. Sin embargo, para implementar esta estrategia de evaluación DID se requiere de datos preliminares al Programa o de línea de base. Si se cuenta con datos longitudinales, como es el caso de *Oportunidades*, es posible combinar una estrategia DID con pareamiento. Intuitivamente, esta estrategia requiere de supuestos más débiles que cualquiera de las otras dos estrategias expuestas anteriormente. En particular, el supuesto de independencia condicional que típicamente se usa en el pareamiento debe mantenerse frente a los cambios en el tiempo en las variables de interés, y por lo tanto permite diferencias sistemáticas no observables (aunque invariables en el tiempo) entre los grupos de tratamiento y control.

5.4.4. Identificación de supuestos de los diferentes estimadores

(a) Pareamiento de corte transversal

El estimador de corte transversal asume que:

$$(CS.1) \quad E(Y_{0t}|P(X),T=1)=E(Y_{0t}|P(X),T=0)$$

$$(CS.2) \quad 0 < \Pr(T=1|X) < 1$$

en algún tiempo t posterior al Programa.

Bajo estas condiciones, el parámetro EPIT puede ser estimado mediante:

$$IT(P(X)) = (1/n_1) \sum_i Y_{1i}(P(X_{1i})) - E(Y_{0i}|P(X_i),T=0),$$

donde la suma es sobre n_1 , el número de individuos expuestos con X valores que satisfacen el CS.2. El término $E(Y_{0i}|P(X_i),T=0)$ representa el resultado pareado para cada individuo tratado. El resultado pareado puede estimarse en forma no paramétrica por el método de “vecino más cercano”, regresión lineal local, o *kernel*, como se explica más adelante.

(b) Pareamiento por diferencias en diferencias (DID)

Un estimador de pareamiento por diferencias en diferencias requiere datos transversales repetidos (o datos longitudinales) de los individuos en los grupos de tratamiento y control. Sean t y t' dos periodos de tiempo, uno anterior al inicio del Programa y otro posterior. Y_{0t} es el resultado observado en el tiempo t . Los supuestos necesarios para justificar la aplicación del estimador son:

$$(DID.1) \quad E(Y_{0t} - Y_{0t'}|P(X),T=1) = E(Y_{0t} - Y_{0t'}|P(X),T=0)$$

$$(DID.2) \quad 0 < \Pr(T=1|X) < 1$$

donde t es un periodo de tiempo posterior al Programa y t' uno anterior al Programa. El estimador de pareamiento DID de la intención al tratamiento está dado por



$$IT(P(X)) = \left\{ \frac{1}{n_{1t}} \sum_i Y_{1it}(P(X_i)) - E(Y_{0t} | P(X_i), T=0) \right\} - \left\{ \frac{1}{n_{1c}} \sum_i Y_{0ic}(P(X_i)) - E(Y_{0c} | P(X_i), T=0) \right\}$$

donde n_{1t} y n_{1c} son el número de observaciones de los individuos tratados en los dos periodos de tiempo. Nótese cómo en el periodo de tiempo inicial se observan Y_{0it} (resultados sin tratamiento) tanto para el grupo $T=1$ como $T=0$, ya que en el periodo de línea de base ninguno de los grupos ha recibido el tratamiento.

(c) *Diferencias en diferencias con pareamiento*

Cuando el estimador de diferencias en diferencias se combina con el pareamiento, se asume que los supuestos relevantes para el pareamiento se mantienen para los cambios en la variable de interés.

$$(CSDID.1) \quad E(\Delta Y_{0c} | P(X), T=1) = E(\Delta Y_{0c} | P(X), T=0)$$

$$(CSDID.2) \quad 0 < \Pr(T=1 | X) < 1$$

5.4.5. Implementación de los estimadores de pareamiento

Los estimadores de pareamiento por puntaje de propensión generalmente se implementan en dos pasos.

Paso 1. Estimar un modelo de participación en el Programa

La probabilidad condicional de participar en el Programa (llamada también puntaje de propensión) juega un papel importante en el pareamiento y en otros estimadores para evaluación. Como lo muestran Rosenbaum y Rubin,¹⁷ cuando la probabilidad de participación en el Programa puede ser estimada mediante un procedimiento paramétrico (como logit o probit), el problema del pareamiento puede reducirse a un pareamiento basado en una variable aleatoria unidimensional. Es decir, el problema se reduce a estimar $E(Y_{0it} | T=0, P(X))$ –en lugar de un problema de k dimensiones– tal como estimar $E(Y_{0it} | T=0, X)$.

Para estimar el puntaje de propensión $P(X)$ es necesario elegir primero un conjunto X de variables condicionantes. Es importante restringir la elección de variables X a aquellas que no sean influidas por la presencia del programa. Por esta razón, las variables X generalmente se establecen como características individuales o familiares medidas antes del inicio del Programa (información de línea de base). Con frecuencia el conjunto de variables X es elegido para maximizar la proporción de observaciones clasificadas correctamente dentro del modelo, aunque no existe una justificación teórica para este procedimiento. Una vez que se ha elegido el conjunto de variables X , se pueden realizar pruebas de balance de muestras para identificar si ciertas variables X deben interactuarse en el modelo (véase Smith y Todd, para una variedad de pruebas de balance que se utilizan comúnmente en la literatura).¹⁸

Paso 1. Construir los resultados pareados

Para construir los resultados pareados es necesario estimar $E(Y_{0t} | P(X), T=0)$ para el estimador de pareamiento de corte transversal y $E(Y_{0c} | P(X), T=0)$ y $E(Y_{0c} | P(X), T=0)$ para el estimador DID. Existen diversos estimadores no paramétricos que se han utilizado para estimar estas medias condicionales. El método más sencillo es el estimador del “vecino más cercano”. El estimador de vecino más cercano k puede implementarse como sigue:

- (I) Calcular $|P(X_i) - P(X_j)|$ para la observación de tratamiento i y para todas las observaciones j del grupo de comparación.



- (II) Ordenar las observaciones j en términos de $|P(X_i) - P(X_j)|$ de la más baja a la más alta.
- (III) Indexar en A_x las observaciones k en $T=0$ con los valores más bajos de $|P(X_i) - P(X_j)|$. Éstos son los llamados vecinos más cercanos.
- (IV) Construir el resultado pareado como el promedio simple de los resultados para los vecinos más cercanos

$$E(Y_{0i} | P(X_i), T_i=0) = (1/k) \sum_j Y_{0j}$$

Una preocupación al implementar los estimadores de pareamiento por el vecino más cercano es que los pares pueden estar lejos en el sentido de que $|P(X_i) - P(X_j)|$ puede ser muy grande, lo cual puede ocurrir cuando no se satisface el requerimiento del soporte común —es un conjunto de observaciones cuyos puntajes de propensión se traslapan— (las condiciones CS.2 y DID.2 antes mencionadas).^j

Una manera de determinar cuáles observaciones caen en la región de traslape de soporte consiste simplemente en diagramar el histograma de los valores $P(X_i)$ para los grupos de tratamiento y control, y posteriormente identificar visualmente los rangos de $P(X_i)$ donde no haya pares cercanos. Otra manera de abordar este problema es mediante pareamiento con *Caliper*, donde sólo se seleccionan pares si $|P(X_i) - P(X_j)| < \varepsilon$, para algún valor predeterminado ε , llamado *Caliper*. Una forma más sistemática de determinar la región de traslape consiste en calcular directamente la densidad $f(P(X_i) | T=0)$ para cada valor $P(X_i)$ observado en las observaciones con $T_i=1$, a través de un estimador no paramétrico de densidad *kernel*:

$$f(P(X_i) | T=0) = \sum_i K((P(X_i) - P(X_k)) / (h_n)),$$

donde K es una función kernel, h_n es un parámetro de ancho de banda, y la sumatoria es sobre todas las observaciones de no tratados con $T_j=0$. Una vez obtenidas las estimaciones de la densidad en cada punto, se ordenan las estimaciones de densidad y se encuentra el cuantil 1% de las densidades positivas estimadas. Todos los valores de $P(X_i)$ para los cuales la densidad estimada exceda este umbral se consideran que están dentro de la región de traslape de soporte. Los valores por debajo del umbral están fuera de la región, se considera que están fuera de la región de soporte común, por lo que se excluyen de la estimación.

A pesar de su fácil implementación, una dificultad al utilizar estimadores de pareamiento por vecino más cercano consiste en cómo analizar la variabilidad de los estimadores. Recientemente se ha mostrado que los métodos de remuestreo (*bootstrap*) son inconsistentes para estimadores de pareamiento por vecino más cercano.¹⁹ Una alternativa consiste en utilizar estimadores de pareamiento por kernel o regresión lineal local propuestos por Heckman, Ichimura y Todd.¹⁵ Estos métodos usan una regresión no paramétrica para construir los resultados pareados $E(Y_{0i} | P(X_i), T_i=0)$. Los métodos *bootstrap* pueden utilizarse para analizar la variabilidad de los métodos kernel o de regresión lineal local.

5.5. Métodos no experimentales para identificar y estimar el EPTT

En esta sección discutimos métodos para identificar y estimar el EPTT que consideran los problemas de falta de aleatoriedad tanto en la asignación del Programa como en la no participación en el mismo. Este último problema es particularmente importante en la evaluación de *Oportunidades* en el ámbito

^j Establecemos como soporte de $P(X)$ los valores para los cuales $f_x(P(X) | T=1) > 0$ y $f_x(P(X) | T=0) > 0$ (la región de traslape de soporte). Para implementar estimadores de pareamiento se requiere determinar, para cada valor de $P(X_i)$, si se encuentra o no dentro de la región de traslape del soporte. El impacto promedio del Programa sólo puede obtenerse para el grupo de personas en tratamiento dentro de la región de traslape.



urbano, ya que las tasas de participación en zonas urbanas son inferiores a 50% en 2003 y no se incrementaron significativamente en 2004. De este modo, las familias que participaron en el Programa son una submuestra selecta de hogares elegibles de participar.

Se consideraron dos diferentes enfoques. El primero consiste en una generalización de los métodos de pareamiento antes mencionados. El segundo, por el contrario, utiliza la disponibilidad del Programa (entre las áreas) como una variable instrumental para la participación. En ambos casos se toman en cuenta las diferencias no observables fijas en el tiempo mediante la explotación del carácter longitudinal de los datos y la consideración de los cambios, más que niveles, de las variables de interés.

5.5.1. Métodos de pareamiento

En el caso de *Oportunidades*, la presencia de hogares no participantes en las zonas de tratamiento ofrece la ocasión de considerar diferentes grupos control. En particular, es posible comparar a los beneficiarios del Programa, ya sea con individuos elegibles en zonas de tratamiento que no participan en el Programa o con individuos elegibles en zonas de control (zonas de no intervención). Sin embargo, en ambos casos hay que considerar el hecho de que los participantes no son una muestra aleatoria de individuos elegibles. Aquí se consideran dos posibles alternativas dentro de la estrategia del pareamiento: comparar participantes con no participantes dentro de las zonas de intervención, y comparar participantes con hogares elegibles en zonas de control.

El propósito del estimador de pareamiento es equilibrar la composición de las muestras de tratamiento y control en términos de las variables observables X . Una vez establecido el grupo control, uno puede fácilmente adaptar los métodos discutidos en la sección anterior para la identificación y estimación de EPTT adoptando la siguiente notación: considerar sólo individuos que viven en zonas de intervención ($T=1$), y dejar que Y_1 denote el resultado para las personas que reciben el tratamiento y Y_0 el resultado de los no tratados. Igual que antes, sea $P=1$ para los participantes en el Programa y $P=0$ para los no participantes. Así, $P(X)=\Pr(P=1|X)$ denota ahora la probabilidad condicional de ser tratado. Esta nueva notación aplicada a las secciones 5.4.1 y 5.4.2 define los supuestos de identificación. La implementación del estimador se hace también como se describió en esas secciones, con la única diferencia de que se comparan individuos con y sin tratamiento dentro de las zonas de tratamiento, en lugar de comparar individuos a los que se les ofrece o no el Programa.

Si uno compara participantes y no participantes en zonas de tratamiento, se debe asumir que el Programa no tiene efectos indirectos en los no participantes que viven en zonas de tratamiento. Una estrategia de evaluación alternativa dentro del marco del pareamiento, que sea robusta ante potenciales efectos indirectos, consiste en comparar a los participantes en el Programa con sujetos elegibles en zonas de control, es decir, individuos con características observables similares que vivan en zonas donde no se ofrece el Programa. Es posible estimar la probabilidad de participación en el Programa utilizando participantes y no participantes dentro de las zonas de intervención y posteriormente utilizar el modelo de estimación para asignar puntajes de propensión a los controles en zonas de no intervención. Una ventaja de esta selección de datos para el grupo de comparación es que evita suponer que el Programa no tiene efectos indirectos en los no participantes en las áreas de intervención. Simplemente se necesita suponer que el Programa no afecta el cambio en los resultados de sujetos que viven en diferentes grupos, lo cual es más plausible, dada la mayor distancia (económica y geográfica) entre estos dos grupos. Sin embargo, al mismo tiempo requerimos que la participación en el Programa sea aleatoria, condicional a las variables utilizadas en el puntaje de propensión y a cualquier característica no observable invariante en el tiempo.

Un argumento similar puede utilizarse para otra posibilidad, que es estimar un puntaje de propensión que prediga el grupo de residencia, y posteriormente comparar sujetos tratados con sujetos en el grupo control con puntaje de propensión igual o similar. En este caso se están equilibrando variables observables



entre los grupos, pero nada asegura que también se estén equilibrando las variables que determinan la participación en el Programa, y los resultados del mismo.

5.5.2. Utilización de la disponibilidad del Programa como variable instrumental

Una forma de evitar el fuerte supuesto de que se observan todas las variables que determinan la participación en el Programa y los resultados, es utilizar la disponibilidad del Programa como una variable instrumental. La idea que está detrás de este estimador es simple: el efecto promedio de la intención al tratamiento (EPIT) puede descomponerse en:

$$E(Y_1 - Y_0 | X, P=1) = \Pr(T=1 | P=1, X) E(Y_1 - Y_0 | X, P=1, T=1) + \Pr(T=0 | P=1, X) E(Y_1 - Y_0 | X, P=1, T=0),$$

donde $\Pr(T=1 | P=1, X)$ es la fracción de hogares que participan en el Programa dado que son elegibles. Si las familias que son elegibles pero que no participan en el Programa no experimentan ningún efecto del tratamiento, entonces se podría esperar que $E(Y_1 - Y_0 | X, P=1, T=0) = 0$. En tal caso, el parámetro del efecto promedio del tratamiento sobre los tratados (EPTT) puede expresarse como una función del parámetro EPIT:

$$E(Y_1 - Y_0 | X, P=1, T=1) = E(Y_1 - Y_0 | X, P=1) / \Pr(T=1 | P=1, X). \\ \text{EPTT}(X) = \text{EPIT}(X) / \Pr(T=1 | P=1, X).$$

Nótese que el supuesto de que el Programa no afecta a las personas que no participan en él descarta los efectos indirectos del tratamiento, como en algunos otros estimadores del EPTT arriba mencionados. Angelucci y Attanasio (2006)²⁰ exponen detalladamente el uso de este método para la identificación y estimación de EPTT.

Este EPTT puede estimarse de diversas maneras: por ejemplo, se pueden utilizar métodos de pareamiento por puntaje de propensión en diferencias en diferencias para estimar el EPIT, dividirlo entre la probabilidad de participación, y posteriormente integrarlo sobre la densidad de X para los participantes en el Programa.

5.6. Métodos y parámetros adicionales

Los métodos presentados hasta ahora han sido implementados en el contexto de evaluación de *Oportunidades*. Sin embargo, en principio pueden utilizarse otros estimadores. Algunos de ellos, sin embargo, identifican parámetros ligeramente diferentes a los hasta ahora analizados. Aquí se presentan tres métodos: variables instrumentales (VI), regresión discontinua y métodos estructurales.

5.6.1. Variables instrumentales (VI)

Otro método que puede ser utilizado para analizar el impacto de las intervenciones del Programa en presencia de autoselección es el método de variables instrumentales. Para ver cómo este método puede ser aplicado para estimar el parámetro EPTT, se escribe el modelo para resultados en los tratados como:

$$Y_1 = g(X_1) + U_1 \\ Y_0 = g(X_0) + U_0$$



o

$$Y = g(X_0) + D(g(X_1) - g(X_0) + E(U_1 - U_0)) + U_0 + D(U_1 - U_0 - E(U_1 - U_0) | T=1) \\ = g(X_0) + D(TT(X)) + \{U_0 + D(U_1 - U_0 - E(U_1 - U_0) | T=1)\}$$

Supóngase que existe una posible variable instrumental Z que está correlacionada con la decisión de participar en el Programa, pero que no está correlacionada con el término residual (entre llaves). Por ejemplo, en el caso de *Oportunidades*, la distancia física entre un hogar y el módulo de incorporación puede ser considerada determinante en la decisión de la familia para solicitar o no su participación en el Programa, pero puede ser aleatoria respecto de componentes idiosincráticos del efecto del tratamiento. La aplicación de variables instrumentales para estimar EPTT se justifica si:

$$E(U_0 | X, Z, T) = 0 \\ \text{y} \\ E(D(U_1 - U_0 - E(U_1 - U_0) | T=1) | X, Z, T) = 0.$$

Como ha sido señalado por Heckman (1997),²¹ esta última condición implica que este método genera un estimado del EPTT sólo bajo el supuesto de que los efectos son homogéneos o si las personas no seleccionan diferencialmente su participación en el Programa con base en las ganancias potenciales que podrían recibir de él. En caso de que esto no sea cierto, el uso de la técnica de VI produce un estimado del EPTT para individuos que son inducidos a participar por una variación observada en el instrumento Z . El parámetro es, por lo tanto, específico al instrumento.

5.6.2. Diseño de regresión discontinua (DRD)

Oportunidades ha sido ofrecido a individuos con un índice de marginalidad (o de pobreza) superior^k al valor de un punto de corte. La discontinuidad en la regla de asignación al Programa puede ser explotada mediante estimadores de regresión discontinua, como fue aplicado en Buddelmeyer y Skoufias²² para estudiar la implementación inicial del programa Progresá en el ámbito rural. Sea z el valor del índice de marginalidad y sea z_0 un punto de corte que determina si la familia es o no elegible a recibir el Programa. El estimador por regresión discontinua asume que:

$$(RD-1) E(Y_0 | z) \text{ es continua en } z_0$$

$$(RD-2) \Pr(T=1 | z) \text{ es discontinua en } z_0$$

donde

$$Y_1 = Y_0 + D(Y_1 - Y_0) = Y_0 + D\Delta,$$

y donde Δ es un efecto aleatorio que representa el efecto del tratamiento. Una comparación de las medias de los resultados de individuos cuyos valores z caen justo arriba y debajo de z_0 nos da

$$\lim_{\epsilon \rightarrow 0} E(Y | z = z_0 + \epsilon) - E(Y | z = z_0 - \epsilon) = \lim_{\epsilon \rightarrow 0} E(\Delta | z = z_0 + \epsilon) \{ \Pr(T=1 | z = z_0 + \epsilon) - \Pr(T=1 | z = z_0 - \epsilon) \}$$

^k En las localidades rurales del diseño original, de acuerdo con el algoritmo de clasificación allí implementado, el Programa fue ofrecido a aquellos hogares por debajo del punto de corte establecido por los implementadores de *Oportunidades*. En el caso urbano, el Programa fue ofrecido a los hogares que tuvieron un puntaje de pobreza por encima del punto de corte.



Así, un estimador para el efecto del tratamiento en personas o familias con índices de marginalidad cercanos al valor del punto de corte está dado por

$$E(\Delta|z=z_0+e) = \frac{\lim_{e \rightarrow 0} E(Y|z=z_0+e) - \lim_{e \rightarrow 0} E(Y|z=z_0-e)}{\lim_{e \rightarrow 0} \Pr(T=1|z=z_0+e) - \lim_{e \rightarrow 0} \Pr(T=1|z=z_0-e)}$$

Cada una de las expresiones límite puede ser estimada mediante una regresión no paramétrica, como ha sido expuesto por Hahn, Todd y Van der Klaauw.²³ Cuando el índice de marginalidad es el único determinante para que una persona reciba o no el tratamiento, entonces se dice que el diseño es estricto y el denominador tendrá $\lim_{e \rightarrow 0} \Pr(T=1|z=z_0+e)=1$ y $\lim_{e \rightarrow 0} \Pr(T=1|z=z_0-e)=0$. Sin embargo, cuando existe algún ruido en el mecanismo de asignación o cuando existen también otros determinantes en la asignación del Programa, entonces se dice que el diseño es difuso y el efecto del Programa cerca del punto de discontinuidad puede también ser estimado mediante una especificación análoga a la expresión anterior. El estimador por regresión discontinua generalmente descarta de forma implícita cualquier efecto indirecto, ya que asume que la gente sólo experimenta efectos del Programa cuando participa en él directamente. Por ejemplo, pueden ocurrir efectos indirectos si un aumento en las tasas de asistencia escolar inducido por el Programa cambia la calidad de las escuelas para niños que no participan directamente en el Programa.

Buddelmeyer y Skoufias²² utilizaron un enfoque de regresión discontinua para evaluar los efectos de Progresá con los datos experimentales del ámbito rural. La fuente de discontinuidad es el punto de corte utilizado para determinar la elegibilidad, y su estrategia de estimación consistió en comparar familias con valores de índice de marginalidad justo por arriba y debajo del valor del punto de corte. Adoptaron un diseño difuso, debido a la presencia de cierto ruido en el mecanismo de asignación con respecto del índice. Al comparar las estimaciones resultantes con las estimaciones experimentales, encontraron valores bastante cercanos y concluyeron que el estimador por regresión discontinua habría sido una estrategia de estimación confiable en caso de no haber contado con los datos experimentales.

Al igual que con las variables instrumentales, el parámetro identificado mediante DRD generalmente es diferente al EPTT. Es decir, en una situación donde el efecto del Programa es heterogéneo, el DRD identifica el EPTT para el grupo de familias que se encuentran cerca del punto de corte que determina su elegibilidad. Este parámetro puede ser útil en algunas situaciones (por ejemplo, si se piensa cambiar el punto de corte), pero no debe generalizarse si existen razones para pensar que los efectos son muy heterogéneos.

5.6.3. Modelos estructurales

Los métodos de estimación antes mencionados son útiles para analizar los efectos del programa existente, pero no han sido adaptados para evaluar los efectos de programas que difieren de algún modo del programa implementado. Por ejemplo, en el contexto de *Oportunidades*, puede haber interés en saber de qué forma serían diferentes los impactos del Programa si las familias recibieran apoyo económico sin condicionarlo a que sus hijos asistan a la escuela. Podría ser también interesante saber cómo cambiarían sus efectos si se cambiara el nivel de subsidios. Los métodos estructurales de estimación pueden utilizarse para analizar los efectos de variaciones del Programa que aún no han sido implementadas. Pueden también utilizarse para predecir el efecto del mismo antes de su implementación.

Para ilustrar su uso en una evaluación ex ante, describimos un modelo de escolarización de tres periodos presentado en Todd y Wolpin (2006),²⁴ el cual es una versión simplificada del modelo estimado en Todd y Wolpin (2006).²⁵ El modelo describe las decisiones de los padres acerca de enviar a sus hijos a la escuela. En los dos primeros periodos, los padres deciden entre enviar a sus hijos a la escuela o a trabajar. Si el hijo



o hija va a la escuela, entonces su logro escolar se incrementa en un año. Al finalizar el periodo, los padres reciben una utilidad por el aprovechamiento escolar de sus hijos. Véase Attanasio, Meghir, Santiago (2005)²⁶ para una formulación alternativa de un modelo estructural de escolarización que asume que las decisiones de escolarización son tomadas sobre la base de las ganancias futuras esperadas.

Sea $d(t)=1$ si un niño asiste a la escuela en un periodo t y $d(t)=0$ si el niño trabaja y no asiste a la escuela (se asume que ambas actividades son mutuamente excluyentes). Sea $S(t)$ el logro escolar del niño en el periodo t . En los primeros dos periodos de tiempo, los padres deciden si su hijo asiste a la escuela [es decir, deciden en $d(1)$ y $d(2)$]. En el tercer periodo de tiempo, el niño es ya un adulto y $d(3)=0$.

La utilidad de los padres en el tercer periodo está dada por $U(3)=S(3)$, donde

$$S(3) = d(1)+d(2)=S(2)+d(2).$$

Para los periodos 1 y 2, la función de la utilidad está dada por

$$U(t)= \alpha S(t)+(\beta_1-\beta_2k^s+\varepsilon^s(t))d_1(t)+C(t)$$

donde k^s denota la distancia a la escuela. El consumo $C(t)$ depende del ingreso de los padres y^p y el ingreso del hijo o hija (en caso de que trabaje) y^c :

$$C(t) = y^p + y^c (t)(1-d(t))$$

$$y^c(t) = \gamma_0 - \gamma_1 k^c + \varepsilon^c(t),$$

donde k^c denota la distancia a la ciudad, un determinante en el salario de los hijos.

Se asume que los errores tienen una distribución normal conjunta:

$$\begin{pmatrix} \varepsilon^c \\ \varepsilon^s \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_c^2 & \sigma_{cs} \\ \sigma_{cs} & \sigma_s^2 \end{pmatrix} \right)$$

A continuación se ilustra cómo pueden identificarse los parámetros del modelo mediante los datos disponibles. Primero, se considera la regla de decisión en el periodo 2, tomando como dada la elección del periodo uno $d(1)$. Las funciones valor asociadas con las elecciones $d(2)=1$ y $d(2)=0$ son:

$$V^1(2) = d(1)+ \beta_1 - \beta_2 k^s + y^p + \varepsilon^s(2) + \delta \alpha (d(1)+1)$$

$$V^0(2) = d(1)+y^p + \gamma_0 - \gamma_1 k^c + \varepsilon^c(2) + \delta \alpha d(1)$$

y los padres eligen

$$d(2) = 1 \text{ iff } V^1(2) \geq V^0(2)$$

$$\delta$$

$$\varepsilon^c(2) - \varepsilon^s(2) \leq (\beta_1 - \gamma_0) - \beta_2 k^s + \gamma_1 k^c + \delta \alpha$$

Si se observan salarios potenciales para todos los niños, entonces se pueden identificar los parámetros de la ecuación del salario infantil. La probabilidad de que los padres manden a sus hijos a la escuela en el segundo periodo está dada por

$$\Pr(d(2)=1) = \Phi \left(\frac{(\beta_1 - \gamma_0) - \beta_2 k^s + \gamma_1 k^c + \delta \alpha}{\sigma^*} \right)$$



donde $\sigma^* = \sqrt{\sigma_c^2 + \sigma_s^2 - \sigma_{cs}}$. Los parámetros $\beta_1 + \delta\alpha$ y β_2 pueden ser identificados a partir de un modelo probit sobre la decisión de escolarización en el segundo periodo.

Posteriormente, se examina cómo identificar el factor de descuento. Sea Γ los parámetros que podrían ser únicamente identificados por el argumento anterior. Las funciones valor en el periodo 1 están dadas por:

$$\begin{aligned} V^1(1) &= \beta_1 - \beta_2 k^s + y^p + \varepsilon^s(1) + \delta E \max(V^0(2), V^1(2)) | d(1)=1 \\ V^0(1) &= y^p + \gamma_0 - \gamma_1 k^c + \varepsilon^c(1) + \delta E \max(V^0(2), V^1(2)) | d(1)=0 \end{aligned}$$

$$E \max(V^0(2), V^1(2)) | d(1) = 1 = F1(\Gamma, \beta_1 + \delta\alpha)$$

$$E \max(V^0(2), V^1(2)) | d(1) = 0 = F0(\Gamma, \beta_1 + \delta\alpha)$$

Los padres eligen la escolarización en el primer periodo si

$$d(1)=1 \text{ if } V^1(1) > V^0(1),$$

o

$$PR(d(1)=1) = \Phi \left(\frac{(\beta_1 - \gamma_0) - \beta_2 k^s + \gamma_1 k^c + \delta [F_1(\Gamma, \beta_1 + \delta\alpha) - F_0(\Gamma, \beta_1 + \delta\alpha)]}{\sigma_c^2 - 2\sigma_{cs} + \sigma_s^2} \right)$$

Podemos identificar por separado δ , β_1 y α mediante supuestos de forma funcional.

Ahora consideremos cómo puede utilizarse el modelo para estimar los contrafactuales, tales como (i) predecir los efectos de un subsidio escolar o cambiar el esquema de subsidios, (ii) ofrecer un bono de graduación o (iii) evaluar los efectos de incrementar la oferta de escuelas. Para predecir los efectos de incrementar el subsidio, puede introducirse un subsidio b en la restricción presupuestal:

$$C(t) = y^p + y^c(t)(1-d(t)) + bd(t)$$

y el modelo puede utilizarse para simular elecciones con el subsidio. De modo similar, se puede introducir un bono de graduación para el tercer periodo como:


$$U(3) = S(3) + b I(S(3)=2),$$

donde $I(S(3)=2)$ es una función indicadora de si el niño o niña ha alcanzado dos grados escolares en el periodo tres. Un programa de construcción de escuelas podría verse como una reducción en el costo del transporte a la escuela debido a la distancia. Es posible analizar, por ejemplo, la toma de decisiones siendo k^s igual a 0 (por ejemplo, una escuela en cada población).



La sección anterior describe varias opciones para analizar los datos relacionados con la evaluación del Programa frente a diversos problemas de estimación. Los evaluadores han utilizado muchas de las estrategias de análisis arriba descritas, como puede apreciarse al revisar los documentos de evaluación.²⁷⁻⁵⁰ La complejidad del diseño de evaluación presenta un reto interesante para el análisis de datos. En esta sección y en la siguiente se enumeran las principales técnicas de análisis utilizadas en la evaluación rural y urbana de *Oportunidades*.

VI. Metodología utilizada en el diseño de evaluación de *Oportunidades*



6.1 Métodos de evaluación utilizados en zonas urbanas

Durante su primera fase de expansión a zonas urbanas, el diseño de evaluación aprovechó la incorporación progresiva al Programa, de modo que las zonas excluidas se utilizaron para construir la muestra control con la intención de comparar las familias elegibles en dichas zonas con las familias beneficiarias desde 2002.

Como ya se mencionó, la implementación de *Oportunidades* dentro de las zonas urbanas elegidas no fue aleatoria. Se decidió atender primero las zonas con mayor número de familias pobres. Otra característica importante del Programa es el método de incorporación. En las zonas urbanas, *Oportunidades* promocionó y dio oportunidad a las familias para acudir a registrarse en módulos centralizados que generalmente permanecieron abiertos unas cuantas semanas. Las tasas de registro en estas zonas fueron mucho más bajas que en zonas rurales. Una posible razón para esta baja participación es la falta de información sobre el Programa, asociada con el método de implementación del mismo. Otra posible explicación es que algunas familias no hayan percibido a *Oportunidades* (cuyos esquemas de pago se mantuvieron al mismo nivel que en zonas rurales) lo suficientemente atractivo. En consecuencia, los hogares que sí participaron son una submuestra selecta de los hogares elegibles. La evaluación en zonas urbanas debe tomar en cuenta la no aleatoriedad en la asignación al Programa en las diferentes zonas, así como el proceso de auto-asignación.

Casi todos los análisis que se han llevado a cabo dentro del diseño de evaluación urbana han aplicado el método no experimental descrito en las secciones 5.5.1 y 5.5.2. El cuadro A.1 en el apéndice muestra algunas estrategias utilizadas hasta ahora para el análisis de los datos de la evaluación urbana. Como se observa, las estrategias más utilizadas por los evaluadores para controlar las diferencias sistemáticas pretratamiento entre los grupos de intervención y de control, y que deben tomarse en consideración para el análisis, son los métodos por diferencias en diferencias con ajustes mediante puntaje de propensión y controlando por efectos fijos en el caso de información tipo panel, o el método simple mediante pareamiento por puntaje de propensión en el análisis de corte transversal. En la mayoría de los casos, diversos estimadores han sido calculados mediante puntaje de propensión utilizando técnicas no paramétricas como: *kernel*, pareamiento por vecino más cercano o regresión lineal local. Asimismo, los evaluadores han incluido en los modelos efectos fijos a nivel de localidad, y considerado la correlación de los errores a ese nivel con el fin de obtener errores estándar robustos.

La variedad de métodos utilizados por los investigadores pone en evidencia la complejidad de la evaluación en zonas urbanas, la cual dista mucho de tener una solución sencilla y directa. De cualquier forma, el uso coincidente de diversas metodologías es posiblemente el consejo más razonable para obtener hallazgos robustos. Sin embargo, como se describió en la sección 5.6, existen diversos enfoques alternativos que pueden utilizarse para estimar los principales efectos del Programa como el EPIT o el EPTT en el marco de un diseño no experimental.



6.2 Métodos utilizados en la evaluación rural a mediano plazo

Como ya se mencionó en la primera sección de este documento, la evaluación original de impacto a corto plazo de *Oportunidades* en el ámbito rural se llevó a cabo con un diseño experimental, en el cual se asignaron aleatoriamente 320 comunidades a un grupo de tratamiento y 186 a un grupo control. En el año 2000, aproximadamente 18 meses después de iniciado el experimento, las comunidades del grupo control se incorporaron también al Programa.¹ Los análisis de evaluación para este periodo aprovecharon la estructura panel de los datos, y aplicaron, principalmente, métodos de primeras y dobles diferencias con efectos fijos. Más detalles sobre los métodos de evaluación utilizados en la evaluación de *Oportunidades* bajo el diseño experimental, así como los principales resultados obtenidos, se pueden encontrar en Skoufias, Emmanuel (2005).²⁷ Las comparaciones entre los grupos experimentales de tratamiento y control son informativas respecto de los efectos de una exposición diferencial al Programa, pero el grupo control experimental no puede utilizarse para evaluar impactos posteriores a 18 meses a partir de un punto de referencia previo al inicio de *Oportunidades*. Para tal propósito se recolectó información adicional en 2003 para un nuevo grupo de comparación. El grupo se conformó por familias en 151 comunidades muy similares a las comunidades tratadas pero que no habían sido incorporadas al Programa en 2003. Sin embargo, el uso de esta muestra adicional de comunidades control se enfrenta a dos problemas. Primero, es claro que su exclusión del Programa no fue aleatoria. En efecto, existe evidencia de que si bien las comunidades adicionales se eligieron con la intención de maximizar su comparabilidad con la muestra original en términos de características observables, los nuevos controles resultaron menos pobres que la muestra original. Segundo, no se cuenta con información de línea de base correspondiente a 1997, el año anterior al inicio del Programa rural. Para obviar este último problema, la Encel 2003 incluyó varias preguntas retrospectivas a 1997 con el fin de estimar la situación prevalente de estos nuevos hogares en ese periodo pretratamiento.

Si bien la evaluación de *Oportunidades* en zonas rurales siguió originalmente un diseño experimental, la incorporación de un nuevo grupo control y los avances en el proceso de incorporación al Programa representan retos que debemos enfrentar para poder realizar estimaciones consistentes de los efectos del Programa. En general, los estudios de evaluación de impactos a mediano plazo utilizaron principalmente el pareamiento por puntaje de propensión como una estrategia no experimental para estimar el impacto del Programa. En todos los casos se utilizó la información recolectada del nuevo grupo control respecto de 1997 para estimar la probabilidad de que estas familias hubieran sido incorporadas al Programa si éste hubiera estado disponible, en un intento por maximizar la comparabilidad entre los grupos de tratamiento y control en el periodo de base previo a su inicio. Al igual que el análisis realizado en el ámbito urbano, se calcularon diferentes estimadores de los efectos por medio de pareamiento, como vecino más cercano, kernel, regresión lineal local o estimación por diferencias en diferencias (la cual ha sido detalladamente descrita anteriormente). Es importante mencionar que el estimador por DID puede computarse sólo para aquellos resultados en los cuales se recolectó información retrospectiva correspondiente a 1997.

Los estudios de evaluación también han utilizado información del diseño de evaluación aleatorio original para comparar las familias que comenzaron a recibir los beneficios del Programa en 1998 contra el grupo control de familias que comenzaron a recibir los beneficios a comienzos de 2000. Este análisis permite estimar los impactos del Programa ante diferentes intensidades de exposición al tratamiento por medio del análisis de DID.

El cuadro A.2 en el apéndice muestra algunas de las estrategias analíticas utilizadas en diferentes documentos relacionados con la evaluación de *Oportunidades* a mediano plazo en el ámbito rural.

¹ Sin embargo, durante el periodo de experimentación no se le informó al grupo control sobre la intención de ofrecerles posteriormente los beneficios del Programa.

Oportunidades ha recolectado una gran cantidad de información para su evaluación por medio de diseños experimentales y no experimentales en zonas rurales y urbanas. Las características de estos datos, que han sido recolectados considerando el desarrollo del Programa, presentan diversos retos para la estimación de los impactos del mismo. Este documento no sólo contribuye a ofrecer una descripción detallada de los diseños para la recolección de los datos en zonas rurales y urbanas en la fase más reciente de la evaluación, también expone el uso de diferentes estrategias que pueden ser utilizadas en el análisis. Se espera, por lo tanto, que este documento contribuya al mejor uso de la información recolectada por *Oportunidades*, y que permita dar una mejor retroalimentación al mismo.

VII. Conclusiones





Anexos

DOCUMENTO TÉCNICO DE EVALUACIÓN*	ESTIMADOR USADO PARA MEDIR IMPACTO	FUENTES DE DATOS	AUTORES	AÑO
Evaluación del impacto de <i>Oportunidades</i> sobre la inscripción, reprobación y abandono escolar ²⁸	Estimadores de dobles diferencias con efectos fijos	Base de datos SEP 1995-2003 Base becarios <i>Oportunidades</i> 1997-2003	Parker S	2003
Impacto de <i>Oportunidades</i> sobre la educación de niños y jóvenes de áreas urbanas después de un año de participación en el Programa ²⁹	Estimadores de dobles diferencias con efectos fijos ajustados por puntaje de propensión	Encelurb 2002, 2003	Todd P, Gallardo J, Behrman J et al	2004
Impacto de <i>Oportunidades</i> en los comportamientos de riesgo de los adolescentes y en sus consecuencias inmediatas. Resultados de corto plazo en zonas urbanas y de mediano plazo en zonas rurales ³⁰	Pareamiento por puntaje de propensión y estimación de efecto ponderado con el método de <i>Kernel</i>	Encelurb 2002 y 2003	Gutiérrez J, Gertler P, Hernández M et al	2004
Impacto de <i>Oportunidades</i> en la morbilidad y el estado de salud de la población beneficiaria y en la utilización de los servicios de salud. Resultados de corto plazo en zonas urbanas y de mediano plazo en zonas rurales ³¹	Pareamiento por puntaje de propensión y estimación de efecto ponderado con el método de <i>Kernel</i>	Encelurb 2002 y 2003	Gutiérrez J, Bautista S, Gertler P et al	2004
Impacto de <i>Oportunidades</i> en la salud reproductiva de la población beneficiaria ³²	Estimadores de dobles diferencias ajustado por puntaje de propensión	Encelurb 2002 y 2003	Hernández B, Urquieta J, Ramírez D et al	2004
Evaluación de la calidad de la atención médica para la población diabética usuaria de <i>Oportunidades</i> ³³	Estimadores de dobles diferencias	Encelurb 2002 y 2003	Hernández M, Meneses F, Téllez M et al	2004
Impacto de <i>Oportunidades</i> sobre los conocimientos y prácticas de madres beneficiarias y jóvenes becarios. Una evaluación de las sesiones educativas para la salud ³⁴	Estimación de efecto ajustando por puntaje de propensión urbano	Encelurb 2003	Duarte M, Morales S, Idrovo A, et al	2004
Estudio comparativo sobre el estado nutricional y la adquisición de lenguaje entre niños de localidades urbanas con y sin <i>Oportunidades</i> ³⁵	Modelo de regresión lineal múltiple ajustando por puntaje de propensión	Encelurb 2002 y 2003	Neufeld L, Sotres D, García R. et al	2004
Estudio sobre el consumo de los suplementos alimenticios Nutrisano y Nutrividá en niños y mujeres de zonas urbanas beneficiarios de <i>Oportunidades</i> ³⁶	Métodos no paramétricos Prueba de medidas repetidas con modelos mixtos	Subestudio específico	Neufeld L, Sotres D, Flores L. et al	2004
El efecto de <i>Oportunidades</i> sobre el nivel y la composición del consumo en áreas urbanas ³⁷	DID y selección con base en variables observables	Encelurb 2002 y 2003	Angelucci M, Attanassio O, Shaw J.	2004
Evaluación del mecanismo de focalización de <i>Oportunidades</i> en zonas urbanas ³⁸	Regresiones logísticas controlando por efectos fijos	Encelurb 2002 Registros administrativos <i>Oportunidades</i>	Coady D, Parker S	2004
Impacto del Programa <i>Oportunidades</i> en nutrición y alimentación en zonas urbanas de México ³⁹	Métodos de regresión múltiple ajustando por puntaje de propensidad urbano	Encelurb 2002 y 2004	Neufeld L, García A, Leroy J, et al	2005

CUADRO A.1
Metodología de evaluación de impacto usada en el contexto urbano

* Todos los documentos están disponibles en: <http://evaloportunidades.insp.mx>



CUADRO A.2

Metodología de evaluación de impacto usada en el contexto rural

* Todos los documentos están disponibles en: <http://evaloportunidades.insp.mx>

DOCUMENTO TÉCNICO DE EVALUACIÓN*	ESTIMADOR USADO PARA MEDIR IMPACTO	FUENTES DE DATOS	AUTORES	AÑO
Evaluación de efecto de <i>Oportunidades</i> sobre la utilización de servicios de salud en el medio rural ⁴⁰	Estimadores de dobles diferencias con efectos fijos	SISPA (IMSS) 1996-2002	Bautista S, Martínez S, Bertozzi S, et al.	2002
Evaluación del impacto de <i>Oportunidades</i> sobre la inscripción escolar: primaria, secundaria y media superior ⁴¹	Estimadores de dobles diferencias con efectos fijos	Base de datos SEP 1995-1996 Base de datos SEP 2001-2002 Base becarios <i>Oportunidades</i> 1997-2001	Parker S	2002
Evaluación del impacto de <i>Oportunidades</i> sobre la inscripción, reprobación y abandono escolar ²⁸	Estimadores de dobles diferencias con efectos fijos	Base de datos SEP 1995-2003 Base becarios <i>Oportunidades</i> 1997-2003	Parker S	2003
Evaluación del impacto de <i>Oportunidades</i> en la mortalidad materna e infantil ⁴²	Estimadores de dobles diferencias con efectos fijos y aleatorios	Bases de mortalidad SSA/INEGI 1995-2002	Hernández B, Ramírez D, Moreno H et al	2003
Impacto de <i>Oportunidades</i> en los comportamientos de riesgo de los adolescentes y en sus consecuencias inmediatas. Resultados de corto plazo en zonas urbanas y de mediano plazo en zonas rurales ³⁰	Pareamiento por puntaje de propensión y estimación de efecto ponderado con el método de <i>Kernel</i>	Encaseh 1997	Gutiérrez J, Gertler P, Hernández M et al	2004
Impacto de mediano plazo del Programa <i>Oportunidades</i> sobre la educación y el trabajo de jóvenes del medio rural que tenían de 9 a 15 años de edad en 1997 ⁴³	Estimadores de dobles diferencias Estimadores de dobles diferencias ajustados por puntaje de propensión	Encaseh 1997, Encel 1998-2000, Encel 2003	Parker S, Behrman J, Todd P.	2004
Impacto de mediano plazo del paquete de <i>Oportunidades</i> , incluyendo el aspecto nutricional, sobre la educación de niños rurales que tenían entre 0 y 8 años de edad en 1997 ⁴⁴	Estimadores con primeras diferencias Estimadores de dobles diferencias Estimadores de dobles diferencias con efectos fijos ajustados por puntaje de propensión	Encaseh 1997, Encel 1998-2000, Encel 2003	Behrman J, Parker S, Todd P.	2004
Impacto de <i>Oportunidades</i> en la morbilidad y el estado de salud de la población beneficiaria y en la utilización de los servicios de salud. Resultados de corto plazo en zonas urbanas y de mediano plazo en zonas rurales ³¹	Pareamiento por puntaje de propensión y estimación de efecto ponderado con el método de <i>Kernel</i>	Encaseh 1997, Encel 1998-2000, Encel 2003	Gutiérrez J, Bautista S, Gertler P et al	2004
Impacto de <i>Oportunidades</i> en la salud reproductiva de la población beneficiaria ³²	Estimadores de dobles diferencias Estimación de diferencias ajustando por puntaje de propensión	Encaseh 1997, Encel 1998-2000, Encel 2003	Hernández B, Urquieta J, Ramírez D et al	2004
Impacto de mediano plazo del Programa <i>Oportunidades</i> sobre la obesidad y las enfermedades crónicas en áreas rurales ⁴⁵	Pareamiento por puntaje de propensión y estimación de efecto ponderado con el método del vecino más cercano	Encel 2003	Fernald L, Gertler P, Olaiz G	2004
Impacto de <i>Oportunidades</i> sobre los conocimientos y prácticas de madres beneficiarias y jóvenes becarios. Una evaluación de las sesiones educativas para la salud ³⁴	Estimación de efecto ajustando por puntaje de propensión rural	Encel 2003	Duarte M, Morales S, Idrovo A, et al	2004
Impacto de <i>Oportunidades</i> en el crecimiento y estado nutricional de niños en zonas rurales ⁴⁶	Modelos de regresión GEE Estimación de efecto ajustando por puntaje de propensión rural	Encel 2003	Neufeld L, Sotres D, Gertler P et al	2004



Impacto de mediano plazo del Programa Oportunidades sobre el desarrollo infantil en áreas rurales ⁴⁷	Métodos de regresión multivariados ajustado por puntaje de propensión	Encel 2003	Gertler P, Fernald L	2004
Efectos de mediano plazo del Programa Oportunidades sobre el consumo en áreas rurales ⁴⁸	Estimación de efecto ajustando por puntaje de propensión usando el método del vecino más cercano y pareamiento lineal de interacción total	Encel 2003	Attanasio O, Di Maro V.	2004
El efecto de <i>Oportunidades</i> sobre el incremento en el consumo de los hogares a partir de inversiones productivas en microempresas y producción agrícola ⁴⁹	Variables instrumentales. Mínimos cuadrados de dos etapas	Encaseh 1997, Encel 1998-2003 Datos administrativos <i>Oportunidades</i>	Gertler P, Martínez S, Rubio M	2004
El efecto de <i>Oportunidades</i> sobre la dinámica demográfica de los hogares beneficiarios y las decisiones de migración de sus integrantes, en localidades rurales ⁵⁰	Estimadores de dobles diferencias ajustado por puntaje de propensión	Encaseh 1997, Encel 2003	Rubalcava L, Teruel G	2004



Referencias

1. Orozco M, Parker S, Hernández D. El modelo de evaluación de Progresá. En Secretaría de Desarrollo Social. Más oportunidades para las familias pobres: evaluación de resultados del programa de educación, salud y alimentación. Metodología de la evaluación de Progresá 2000 (primera reimpresión). México, 2003
2. Parascandola M, Weed DL. Causation in epidemiology. *J Epidemiol Community Health* 2001;55:905-912.
3. Pearl J. Causal inference in the health sciences: A conceptual introduction. *Health Serv. Outcomes Res. Method* 2001;2:189-220.
4. Heckman JJ, Vytlacil EJ. Structural equations, treatment effects and econometric policy evaluation. *Econometrica* 2005;73:669-738.
5. Winship C, Morgan SL. The estimation of causal effects from observational data. *Annu. Rev. Sociol.* 1999;25:659-706.
6. Heckman J, LaLonde R, Smith J. The Economics and Econometrics of Active Labor Market Programs, *Handbook of Labor Economics* 1999;3A:31:1-7
7. Schulz KF, Grimes DA. Allocation concealment in randomised trials: Defending against deciphering. *Lancet* 2002;359:614-618.
8. Behrman J, Todd P. Randomness in the Experimental Samples of Progresá (Education, Health, and Nutrition Program). International Food Policy Research Institute, Washington, DC, 1999a.
9. Constanza MC. Matching. *Prev Med* 1995;24:425-433.
10. Smith HL. Matching with multiple controls to estimate treatment effects in observational studies. *Sociol. Method.* 1997;27:325-353.
11. Dalenius T, Hodges Jr. JL. Minimum Variance Stratification. *Journal of the American Statistical Association* 1959;54(285):88-101
12. Nota técnica general urbana. Disponible en: <http://evaloportunidades.insp.mx>
13. Nota técnica general rural. Disponible en: <http://evaloportunidades.insp.mx>
14. Todd P. Technical note on using matching estimators to evaluate the Oportunidades program for six year follow-up evaluation of Oportunidades in rural areas. 2004. Disponible en: <http://evaloportunidades.insp.mx>
15. Heckman J, Ichimura H, Todd P. Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies* 1997;64(4):605-654.
16. Consejo Nacional de Población. Encuesta para medir el impacto en salud reproductiva del Programa Oportunidades 2003. Diseño muestral (Documento técnico). México DF: Conapo, 2004. Disponible en: <http://evaloportunidades.insp.mx>
17. Rosenbaum P, Rubin D. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 1983;70:41-55.
18. Smith J, Todd P. Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* 2005;125:305-353.
19. Abadie A, Imbens G. Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica, Econometric Society* 2006;74(1):235-267, 01.
20. Angelucci M, Attanasio O. Estimating ATT effects with non-experimental data and low compliance, 2006, mimeo.
21. Heckman J. Instrumental Variables. A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations, *Journal of Human Resources* 1997;32(3):441-462
22. Buddelmeyer H, Skoufias E. An Evaluation of the Performance of Regression Discontinuity Design on Progresá. *Policy Research Working Paper Series* 3386, Banco Mundial, 2004.
23. Hahn, J, Todd P, Van der Klaauw W. Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica* 2002;69(1):201-209.
24. Todd P, Wolpin K. Ex Ante Evaluation of Social Programs. Penn Institute for Economic Research, Working Paper 2006:06-022.
25. Todd P, Wolpin K. Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico, *American Economics Review*, 2006, forthcoming.
26. Attanasio O, Meghir C, Santiago A. Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresá, IFS working paper EWPO5/01, 2005.
27. Skoufias E. Progresá and its impacts on the welfare of rural households in Mexico. International Food Policy Research Institute, 2005. Disponible en: www.ifpri.org
28. Parker SW. Evaluación del impacto de Oportunidades sobre la inscripción, reprobación y abandono escolar. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2003. Tomo I. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
29. Todd P, Gallardo-García J, Behrman JR, Parker S. Impacto de Oportunidades sobre la educación de niños y jóvenes de áreas urbanas después de un año de participación en el Programa. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo I. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
30. Gutiérrez JP, Gertler P, Hernández M, Bertozzi S. Impacto de Oportunidades en los comportamientos de riesgo de los adolescentes y en sus consecuencias inmediatas. Resultados de corto plazo en zonas urbanas y de mediano plazo en zonas rurales. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo II, Salud. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
31. Gutiérrez JP, Bautista S, Gertler P, Hernández M, Bertozzi S. Impacto de Oportunidades en la morbilidad y el estado de salud de la población beneficiaria y en la utilización de los servicios de salud. Resultados de corto plazo en zonas urbanas y de mediano plazo en zonas rurales. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo II, Salud. Cuernavaca: Instituto Nacional de Salud Pública, 2005.



32. Hernández B, Urquieta JE, Ramírez MD, Figueroa JL. Impacto de Oportunidades en la salud reproductiva de la población beneficiaria. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo II, Salud. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
33. Hernández M, Meneses F, Téllez MM, Manzano A, López R. Evaluación de la calidad de la atención médica para la población diabética usuaria de Oportunidades México. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo II, Salud. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
34. Duarte MB, Morales S, Ochoa SC, Van der Wal SB, Idrovo AJ, Caballero M et al. Impacto de Oportunidades sobre los conocimientos y prácticas de madres beneficiarias y jóvenes becarios. Una evaluación de las sesiones educativas para la salud. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo II, Salud. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
35. Neufeld L, Sotres D, García R, García A, Tolentino L, Fernald L et al. Estudio comparativo sobre el estado nutricional y la adquisición de lenguaje entre niños de localidades urbanas con y sin Oportunidades. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo III, Alimentación. Cuernavaca: Instituto Nacional de Salud Pública, 2005a.
36. Neufeld L, Sotres D, Flores L, Tolentino L, Jiménez J, Rivera J. Estudio sobre el consumo de los suplementos alimenticios Nutrisano y Nutrivida en niños y mujeres de zonas urbanas beneficiarios de Oportunidades. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo III, Alimentación. Cuernavaca: Instituto Nacional de Salud Pública, 2005b.
37. Angelucci M, Attanasio OP, Shaw J. Evaluación del efecto de Oportunidades en el nivel y la composición del consumo en localidades urbanas. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo IV. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
38. Coady D, Parker S. Evaluación del mecanismo de focalización de Oportunidades en zonas urbanas. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo IV. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
39. Neufeld LM, García-Guerra, Leroy J, Flores ML, Fernández AC, Rivera-Dommarco J. Impacto del Programa Oportunidades en nutrición y alimentación en zonas urbanas de México. En Hernández B, Hernández M, eds. Evaluación Externa de impacto del Programa Oportunidades 2006. Tomo I, Salud. Cuernavaca: Instituto Nacional de Salud Pública. 2006b, mimeo.
40. Bautista S, Martínez S, Bertozzi S, Gertler PM. Evaluación del efecto de Oportunidades sobre la utilización de servicios de salud en el medio rural. Evaluación de resultados de impacto del Programa de Desarrollo Humano Oportunidades 2002. Cuernavaca: Instituto Nacional de Salud Pública, 2003.
41. Parker SW. Evaluación del impacto de Oportunidades sobre la inscripción escolar, primaria, secundaria y media superior, en Resultados de la Evaluación Externa del Programa de Desarrollo Humano Oportunidades 2002. México: Sedesol, 2003.
42. Hernández B, Ramírez D, Moreno H, Laird N. Evaluación del impacto de Oportunidades en la mortalidad materna e infantil. Evaluación externa de impacto del Programa Oportunidades 2003. Capítulo II. Cuernavaca: Instituto Nacional de Salud Pública, 2004.
43. Parker SW, Behrman JR, Todd P. Impacto de mediano plazo del Programa Oportunidades sobre la educación y el trabajo de jóvenes del medio rural que tenían de 9 a 15 años de edad en 1997. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo I. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
44. Behrman JR, Parker SW, Todd P. Impacto de mediano plazo del paquete de Oportunidades, incluyendo el aspecto nutricional, sobre la educación de niños rurales que tenían entre 0 y 8 años de edad en 1997. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo I. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
45. Fernald LC, Gertler PJ, Olaiz G. Evaluación del impacto de Oportunidades a mediano plazo en obesidad y enfermedades crónicas en zonas rurales. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo II, Salud. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
46. Neufeld L, Sotres D, Gertler P et al. Impacto de Oportunidades en el crecimiento y estado nutricional de niños en zonas rurales. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo II, Salud. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
47. Gertler PJ, Fernald LC. Impacto de mediano plazo del Programa Oportunidades sobre el desarrollo infantil en áreas rurales. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo III, Alimentación. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
48. Attanasio O, Di Maro V. Evaluación del efecto de Oportunidades a mediano y corto plazo en el consumo en áreas rurales. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo IV. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
49. Gertler PJ, Martínez S, Rubio M. Evaluación del efecto de Oportunidades en el incremento del consumo de los hogares mediante inversiones a la micro-empresa y a la producción agrícola. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo IV. Cuernavaca: Instituto Nacional de Salud Pública, 2005.
50. Rubalcava L, Teruel G. El Efecto del Programa Oportunidades sobre la Composición Demográfica de los Hogares y Decisiones de Migración. En Hernández B, Hernández M, eds. Evaluación externa de impacto del Programa Oportunidades 2004. Tomo IV. Cuernavaca, México, Instituto Nacional de Salud Pública, 2005.

Evaluación externa de impacto del Programa Oportunidades 2006.
Tomo III • Documento metodológico

Se terminó de imprimir en noviembre de 2006.
La edición consta de 1 000 ejemplares
y estuvo al cuidado de la
Subdirección de Comunicación Científica
y Publicaciones del Instituto Nacional de Salud Pública

