

## Capítulo 4

### Cómo aprovechar las evaluaciones de impacto de “Práctica Óptima”\*

En los capítulos anteriores se presentaron los métodos, problemas y dificultades clave que pueden surgir al evaluar el impacto de un proyecto. Al revisar los estudios de casos indicados en el Cuadro 4.1, surgen muchos ejemplos ilustrativos de interesantes enfoques frente al diseño, uso de datos, elección y aplicación de los métodos analíticos usados y la creación de capacidades en el país. A continuación se destacan estos ejemplos, además de un análisis de los costos de las evaluaciones y los problemas de economía política que podrían surgir en la implementación.

Los 15 estudios de casos incluidos en la revisión se seleccionaron de una variedad de evaluaciones realizadas por el Banco Mundial, otras agencias de donantes, instituciones de investigación y empresas consultoras privadas. Se seleccionaron como ejemplos de una “práctica óptima” por su rigor metodológico, intentando reflejar una gama de ejemplos de diferentes sectores y regiones. Aunque cada evaluación de impacto tiene sus fortalezas y debilidades, las lecciones aprendidas de estas experiencias deberían servir de ayuda a los administradores de proyectos o a analistas políticos que intentan diseñar e implementar un futuro trabajo.

#### **Planificación anticipada y cuidadosa del diseño de la evaluación**

Una preparación cuidadosa durante las etapas iniciales de la identificación del proyecto asegurará que se recopile la información correcta y que se puedan emplear los resultados para hacer ajustes a los componentes del proyecto a mitad del proceso. Mediante una planificación anticipada y cuidadosa, es posible incorporar todos los elementos que contribuirán a una evaluación de impacto rigurosa, como una encuesta básica (de referencia) con un grupo de control aleatorio y datos cualitativos sobre los procesos que podrían influir en los efectos.

**Proyecto de Nutrición y Desarrollo Preescolar de Uganda.** Esta evaluación, aunque aún no se implementa, representa un excelente ejemplo de una planificación anticipada y cuidadosa (véase Banco Mundial, 1998a; García, Alderman y Rudqvist, 1999). El proyecto mismo se centra en el fortalecimiento de la capacidad de los padres y comunidades para cuidar a los niños, proporcionándoles conocimientos sobre mejores prácticas de cuidado infantil y mejorando las oportunidades de aumentar el ingreso. Se basa en la comunidad y lo aplica una red de organizaciones no gubernamentales (ONG). El componente de evaluación, que se integró al ciclo del proyecto desde el primer día, se acerca al ideal en

---

\* Este capítulo se basa en gran medida en los estudios de casos de práctica óptima del Anexo I y en publicaciones de información general preparados por Gillette Hall y Julia Lane, y Subbarao y otros (1999).

términos de su diseño. En primer lugar, se generan datos de encuestas básicas y de seguimiento, junto con un grupo de control aleatorio, de modo que se puede evaluar en forma rigurosa el efecto del programa a nivel de los beneficiarios. En segundo lugar, mejora el componente cuantitativo con un proceso participativo (cualitativo) de supervisión y evaluación.

En el aspecto cuantitativo, el proyecto se diseñó con miras a un diseño de estudio experimental, en que las subdivisiones se asignarán en forma aleatoria a los grupos de tratamiento y control. Luego se usarán tarjetas de salud para registrar la información sobre el peso de los niños en las subdivisiones de tratamiento y control. Además, la encuesta básica domiciliaria se realizará antes de entregar los servicios a las comunidades, al igual que una encuesta de seguimiento de los mismos hogares dos años después. Se espera que una rápida revisión de estos datos sirva de guía para la decisión de ampliar algunos componentes de la intervención durante la revisión a mitad del proyecto. Un análisis más profundo de los datos al término del proyecto orientará el diseño de la segunda fase del proyecto.

**Proyecto de crédito con educación de Ghana.** La evaluación de este proyecto fue muy compleja y tuvo varios pasos intermedios. En el proyecto se combinan los elementos de un plan de préstamo grupal con la educación sobre nociones básicas en salud, nutrición, planificación y espaciamiento de los nacimientos y capacitación en pequeños negocios. En general, la evaluación se centra en analizar el estado nutricional de los niños, la capacidad económica de las mujeres para invertir en alimentos y atención de salud, conocimientos y adopción de la lactancia materna y destete por parte de las mujeres. Comienza con un marco conceptual muy claro, el que se ilustra más adelante. Este esquema detalla claramente los aportes, beneficios intermedios y resultados a largo plazo, en una forma que facilita la elaboración de diversos modelos y su interpretación. Al planificar en forma cuidadosa la evaluación y trabajar con un esquema desde una etapa inicial, fue posible aclarar muchos aspectos de un diseño relativamente complejo (véase el Anexo 1.6).

### **Enfoques para la evaluación cuando no hay datos básicos**

En la práctica, muchas evaluaciones no cuentan con los datos adecuados. Se agregan evaluaciones después que es posible realizar una encuesta básica o en ausencia de grupos de comparación. Algunos ejemplos de este aspecto son las evaluaciones de los proyectos Alimentos por Educación de Bangladesh, PROBECAT de México, Programas de Trabajo Activo de la República Checa y TRABAJAR de Argentina. Sin datos básicos, se deben crear controles usando los métodos de pareo analizados en los capítulos anteriores. Sin embargo, esto puede ser bastante complicado. Un buen ejemplo lo constituye la técnica de correspondencia de puntuación de la propensión usada en el proyecto TRABAJAR de Argentina para crear un grupo de control con datos transversales de los participantes y no participantes en el programa.

**Cuadro 4.1 Resumen de evaluaciones de impacto de “práctica óptima”**

<i>Proyecto:</i>	<i>País</i>	<i>Tipo de base de datos</i>	<i>Unidad de análisis</i>	<i>Medidas de resultados</i>	<i>Método econométrico</i>				<i>Cualitativo</i>	<i>Fortalezas</i>	
					<i>Aleatorización</i>	<i>Pareo</i>	<i>Comparaciones reflexivas</i>	<i>Doble diferencia</i>			<i>VARIABLES instrumentales</i>
<b>Educación</b>											
Radio Nicaragua	Nicaragua	Encuesta básica y posterior a la intervención	Estudiantes y clases	Calificaciones en las pruebas	Sí	No	Sí	No	No	No	Diseño de cuestionario
Reforma de autonomía escolar	Nicaragua	Evaluaciones cualitativas y panel de encuestas	Estudiantes, padres, maestros, directores	Calificaciones en las pruebas, grado de toma de decisiones local	No	Sí	Sí	No	No	Sí	Combinación cualitativa y cuantitativa
Libros de texto	Kenya	Encuesta básica y posterior a la intervención	Estudiantes, clases, maestros	Calificaciones en las pruebas	Sí	No	Sí	Sí	Sí	No	Análisis de factores distractivos
Abandono	Filipinas	Encuesta básica y posterior a la intervención	Estudiantes, clases, maestros	Calificaciones en las pruebas y nivel de deserción	Sí	No	Sí	Sí	Sí	No	Análisis de costos/beneficios; creación de capacidades
<b>Programas laborales</b>											
TRABAJAR	Argentina	Encuesta domiciliaria, censo, registros administrativos, evaluaciones sociales	Trabajadores, hogares	Ingreso, focalización, costos	No	Sí	No	No	Sí	Sí	Uso razonable de las fuentes de datos existentes, técnicas analíticas innovadoras
PROBECAT	México	Encuestas retrospectivas y de la fuerza laboral	Trabajadores	Resultados en el ingreso y el empleo	No	Sí	No	No	No	No	Técnica de correspondencia

Programas de Trabajo Activo	República Checa	Encuestas retrospectivas por correo	Trabajadores	Resultados en el ingreso y el empleo	No	Sí	No	No	No	No	Técnica de correspondencia
<b>Finanzas</b>											
Microfinanzas	Bangladesh	Encuesta posterior a la intervención y registros administrativos	Hogares	Consumo y educación	Sí	Sí	No	Sí	No	No	Análisis de factores distractivos
Crédito con educación	Ghana	Encuesta básica y posterior a la intervención	Pares madre e hijo	Ingreso, salud y habilitación	Sí	Sí	Sí	No	No	Sí	Uso de información cualitativa y cuantitativa
Financiamiento de la salud	Níger	Encuesta básica y posterior a la intervención, más registros administrativos	Hogares y centros de salud	Recuperación de costos y acceso	No	Sí (en distritos)	Sí	No	No	No	Uso de datos administrativos
<b>Alimentos y nutrición</b>											
Alimentos para la educación	Bangladesh	Encuesta sobre gastos familiares	Hogares y comunidades	Asistencia escolar	No	No	No	No	Sí	No	Uso imaginativo de instrumentos para abordar el problema de selección con datos estándar
Salud, educación y nutrición	México	Encuestas básicas y posteriores a la intervención	Hogares	Resultados en salud, educación y nutrición	Sí	Sí	Sí	Se ignora	Se ignora	No	Conceptualización clara; análisis de factores distractivos

<b>Infraestructura</b>											
Fondo de Inversión Social	Bolivia	Encuestas básicas y de seguimiento	Hogares, proyectos	Indicadores de educación y salud	Sí	Sí	Sí	Sí	Sí	No	Aplicación de una variedad de metodologías de evaluación
Caminos rurales	Vietnam	Encuestas básicas y de seguimiento	Hogares, comunidades	Indicadores de bienestar a nivel de hogar y comunidad	No	Sí	Sí	Sí	Sí	No	Mide resultados en el bienestar
<b>Agricultura</b>											
Proyecto de extensión nacional	Kenya	Evaluaciones de beneficiarios, datos de panel	Hogares, granjas	Productividad y eficiencia de granjas	No	No	Sí	No	No	No	Aplicabilidad de los resultados en la política

---

**El proyecto TRABAJAR II en Argentina.** Este proyecto se centró en proporcionar empleos de bajo salario en pequeños subproyectos de infraestructura social y económica seleccionados por agrupaciones locales. La evaluación de impacto del programa se diseñó para evaluar si los ingresos de los participantes en el programa eran más altos de lo que hubieran sido en ausencia del programa. Los métodos que se usan habitualmente para estimar el ingreso de los hogares sin la intervención no eran factibles de aplicar en el caso del programa TRABAJAR: no se había realizado aleatorización para crear un grupo de control que se pudiera usar para comparar el ingreso de los beneficiarios del proyecto y no se contaba con una encuesta básica, lo que eliminaba la posibilidad de realizar una evaluación antes y después.

En su lugar, en la evaluación de TRABAJAR se utilizaron los datos existentes para crear un grupo de control a través de la correspondencia entre participantes del programa y no participantes de la población nacional para un conjunto de variables socioeconómicas, como escolaridad, género, vivienda, percepciones subjetivas del bienestar y afiliación a partidos políticos y asociaciones de vecinos, usando una técnica llamada puntuación de la propensión. El estudio es un ejemplo del uso imaginativo de los datos existentes de una encuesta domiciliaria a nivel nacional —la Encuesta de Desarrollo Social (EDS)— para generar el grupo de comparación, junto con una encuesta más reducida de participantes en TRABAJAR realizada específicamente para fines de la evaluación. La encuesta más reducida se diseñó cuidadosamente, de modo de utilizar el mismo cuestionario que la EDS y los mismos equipos de entrevistadores y se llevó a cabo aproximadamente al mismo tiempo, para poder realizar el ejercicio de correspondencia o pareo. Esta técnica fue posible en el caso de TRABAJAR, porque se estaba realizando una encuesta domiciliaria a nivel nacional y los evaluadores podían aprovechar esta encuesta para sobremuestrear los participantes en TRABAJAR. Se usaron los mismos equipos de entrevistadores para la encuesta nacional y la encuesta del proyecto, lo que arrojó ganancias en eficiencia en la recopilación de datos (véase el Anexo 1.1).

**Evaluación de programas del mercado laboral en la República Checa.** En esta evaluación se intentó cubrir cinco programas de trabajo activo para (a) determinar si los participantes en los distintos programas tenían más éxito para reincorporarse al mercado laboral que los que no participaban en el programa y si esto variaba a través de los subgrupos y con las condiciones del mercado laboral y (b) determinar la eficacia en función de los costos de cada programa y sugerir mejoramientos. En esta evaluación se utilizó una técnica de correspondencia, porque no se recopilaron datos básicos (de referencia). Los evaluadores entrevistaron a los participantes y luego seleccionaron una muestra aleatoria de no participantes. Puesto que los no participantes eran sistemáticamente de mayor edad y con un menor nivel de instrucción, los evaluadores tuvieron que crear un grupo de comparación razonable para cada programa. Para ello, se tomó a cada participante de a uno y se le comparó con cada individuo del grupo de no participante sobre la base de siete características: edad, género, educación, número de meses empleado antes del registro, tamaño del pueblo, estado civil y tipo del último empleo. La correspondencia más cercana se introdujo en el grupo de comparación. Aunque este enfoque es directo, existe la

posibilidad de sesgo en la selección, esto es, que el grupo de no participantes sea sistemáticamente diferente al grupo de participantes sobre la base de elementos no observables (Anexo 1.5).

### **El manejo de las restricciones al elaborar controles adecuados**

A veces es posible hacer una selección aleatoria o contar con controles experimentales, pero no es factible en términos de política. En este caso, se puede efectuar la aleatorización aprovechando cualquier plan para probar experimentalmente el proyecto en ciertas áreas restringidas. Inicialmente se pueden seleccionar aleatoriamente las áreas en las cuales se aplicará el proyecto piloto, fijando posibles áreas futuras del proyecto como controles. En el tiempo se pueden incluir aleatoriamente comunidades adicionales en el proyecto. Tres ejemplos ilustran la forma en que se puede manejar una situación en la cual la aleatorización no era factible en términos de política o debido a otros aspectos. En Vietnam se evaluará un proyecto de transporte rural con información limitada y sin aleatorización. El Fondo Hondureño de Inversión Social es un buen ejemplo de la creación de un grupo de control en proyectos guiados por la demanda, usando una comparación pareada ex post basada en un corte transversal único de datos. Evaluar proyectos guiados por la demanda puede ser especialmente difícil, dado que no se conocen anticipadamente los proyectos o comunidades que participarán en el proyecto. Y tercero, la evaluación del Fondo de Inversión Social de Bolivia en la Región del Chaco proporciona un buen ejemplo de cómo incorporar la aleatorización en proyectos guiados por la demanda de forma que permita su orientación hacia grupos destinatarios específicos.

**El Proyecto de caminos rurales de Vietnam.** Este proyecto está destinado a reducir la pobreza en las áreas rurales mejorando el acceso a las comunidades rurales y conectándolas con las redes viales del distrito y la provincia. El diseño de la evaluación del impacto se centra en los datos de encuestas básicas y de seguimiento recopilados para una muestra de comunidades del proyecto y del grupo de comparación identificado mediante técnicas de comparación pareada. De una muestra aleatoria de miembros de la comunidad dentro del proyecto (tratamiento) y fuera de éste (comparación), se recopilará información básica y posterior a la intervención sobre indicadores como rendimientos de la producción agrícola a nivel de la comunidad, diversificación de la fuente de ingresos, oportunidades de empleo, disponibilidad de bienes, servicios e instalaciones y patrimonio y distribución de activos. Estos datos se emplearán para comparar los cambios en los resultados antes y después de la intervención entre comunidades del proyecto y fuera de éste, usando la técnica de “doble diferencia”.

Idealmente, las comunidades de tratamiento y de comparación deberían ser equivalentes en todas sus características observadas y no observadas, siendo la única diferencia entre ellas que las comunidades de tratamiento se benefician con la intervención, mientras que las comunidades de comparación no. Puesto que no se había realizado una asignación aleatoria a grupos de tratamiento y de comparación y no se disponía de los datos necesarios para tomar decisiones fundamentadas sobre los controles adecuados en el momento de la selección de la muestra, se tomaron muestras aleatorias de comunidades dentro del proyecto

y fuera de éste. Específicamente, se seleccionaron las comunidades del proyecto de una lista de todas las comunidades con proyectos propuestos en cada provincia. Luego, se seleccionaron comunidades de comparación de una lista de todas las comunidades restantes sin proyectos propuestos, pero en los mismos distritos que las comunidades de tratamiento. Con la información recopilada para la evaluación, se utilizarán técnicas de correspondencia de puntuación de la propensión para garantizar que las comunidades fuera del proyecto que fueron seleccionadas sean grupos de comparación adecuados. Se eliminará de la muestra cualquier control con atributos inusuales en relación con las comunidades de tratamiento (Anexo 1.15).

**Fondo Hondureño de Inversión Social.** El Fondo Hondureño de Inversión Social (FHIS) (véase Banco Mundial, 1998b) tiene el objetivo de mejorar las condiciones de vida de grupos sociales marginales mediante el financiamiento de subproyectos de infraestructura social y económica de pequeño tamaño. El FHIS es una institución guiada por la demanda que responde a las iniciativas de municipalidades, ministerios de gobierno, ONG y grupos de la comunidad. Proporciona financiamiento para inversiones en infraestructura, equipos y capacitación. La evaluación de impacto del FHIS emplea técnicas de comparación pareada, obteniendo aleatoriamente la muestra del grupo de tratamiento a partir de una lista de comunidades en que los proyectos del FHIS han estado en funcionamiento al menos un año. El grupo de comparación, en cambio, se seleccionó de una lista de proyectos en tramitación: aquellos que han sido solicitados o aprobados, pero para los cuales aún no se ha efectuado la inversión del FHIS. En teoría, las comunidades de los grupos de comparación se parea automáticamente con las comunidades del proyecto de acuerdo a un proceso de autoselección y los criterios de aprobación de proyectos del FHIS. Luego se realizó una encuesta domiciliaria en las comunidades del grupo de tratamiento y de comparación, complementada con un componente de evaluación cualitativa (grupos representativos y entrevistas a informantes clave) aplicado a un subconjunto de comunidades de tratamiento. La evaluación inicial es un primer paso hacia el establecimiento de un sistema supervisión y evaluación permanente que finalmente quede integrado a las operaciones del FHIS. Específicamente, los datos recolectados de las comunidades con proyectos en tramitación pasarán a ser una línea base útil con la cual se puede realizar un seguimiento de futuros cambios en los indicadores de efecto después que se aplique la inversión del FHIS.

**Inversiones educacionales en la Región del Chaco de Bolivia.** Los proyectos de educación financiados por el Fondo de Inversión Social (FIS) de Bolivia tienen el objetivo de mejorar las instalaciones físicas y capacitar a los maestros de escuelas públicas rurales. Los retrasos en la implementación del proyecto en la Región del Chaco y la limitación de fondos para el mejoramiento de las escuelas proporcionó la oportunidad de utilizar un diseño de evaluación experimental, asegurando al mismo tiempo que las escuelas más necesitadas se beneficiaran del proyecto. Las escuelas de la Región del Chaco se clasificaron de acuerdo con un índice de calidad de escuelas basado en la suma de cinco indicadores de infraestructura y equipos escolares: luz eléctrica, alcantarillado, fuente de agua, pupitres por alumno y metros cuadrados de espacio por alumno. Sólo las escuelas bajo un determinado valor límite estaban calificadas para recibir una intervención del FIS. Entre los establecimientos calificados, se seleccionaron automáticamente las escuelas en

peores condiciones para que se beneficiaran de las inversiones financiadas por el FIS. El grupo de la siguiente prioridad más alta incluía 120 escuelas, pero sólo se contaba con fondos para mejorar menos de la mitad de ellas. Así, las escuelas calificadas de este grupo de segunda prioridad se asignaron aleatoriamente a los grupos de tratamiento o de comparación, proporcionando las condiciones para un diseño de evaluación experimental (Anexo 1.4).

### **Combinación de métodos**

Para la mayoría de las evaluaciones, se requiere más de una técnica para obtener resultados sólidos que respondan a varias preguntas de evaluación. Cada pregunta podría necesitar diferentes técnicas, incluso dentro de un único diseño de proyecto. Tres ejemplos ilustran la forma en que se combinaron diversas técnicas en una sola evaluación; el Fondo Social de Bolivia, la Evaluación de TRABAJAR de Argentina y la Reforma Escolar de Nicaragua.

**El Fondo Social de Bolivia.** Los fondos sociales generalmente incluyen varios tipos de subproyectos y por consiguiente el diseño de una evaluación puede abarcar diversos métodos. En el fondo Social de Bolivia, el patrón de implementación del proyecto dictó los métodos de evaluación. En el caso de la educación, ya se habían identificado las escuelas que recibirían la intervención, por consiguiente, no se podía emplear la aleatorización. En su lugar, se adoptaron métodos de pareo. En el caso de los proyectos de salud, se usaron métodos reflexivos, porque la intervención se implementaría en todos los centros de salud de la región (véase el Anexo 1.4).

**Uso de una variada combinación de componentes de evaluación: TRABAJAR II de Argentina.** La evaluación de TRABAJAR incluye una serie de componentes diseñados para evaluar la forma en que el proyecto está logrando sus objetivos de política. El primer componente se basa en los datos de una encuesta domiciliaria para evaluar las ganancias de ingreso que obtuvieron los participantes de TRABAJAR. El segundo componente supervisa la asignación de fondos del programa (focalización), vigilando los cambios en el tiempo como resultado de la reforma. Este componente genera retroalimentación dos veces al año, la que se usa para perfeccionar la focalización del programa. Entre los componentes adicionales de la evaluación se incluye un análisis de costos y beneficios de los proyectos de infraestructura y evaluaciones sociales diseñadas para proporcionar retroalimentación de la comunidad sobre la implementación del proyecto. Cada uno de estos componentes fue aplicado dos veces. Se planean tres componentes futuros. Nuevamente se aplicará la técnica de investigación de comparación pareada, para evaluar el impacto de la participación en el programa TRABAJAR sobre la actividad del mercado laboral. Se volverá a evaluar la calidad del proyecto de infraestructura, esta vez para los proyectos que fueron completados al menos un año atrás, para evaluar la durabilidad, mantenimiento y tasas de utilización. Por último, un componente cualitativo de investigación permitirá investigar las operaciones y procedimientos del programa, entrevistando a miembros del personal de los organismos que patrocinan proyectos y a beneficiarios del programa.

Los resultados de la evaluación proporcionan una clara orientación para la reforma de

políticas. El primer componente de evaluación revela que el programa TRABAJAR es muy exitoso en llegar a los pobres: la autoselección de los participantes mediante el ofrecimiento de bajos salarios es una estrategia que funciona en Argentina. Además, los participantes sí experimentan ganancias en su ingreso como resultado de la participación. El segundo componente permite concluir que la asignación geográfica de los fondos del programa ha mejorado en el tiempo; ahora el programa consigue dirigir mejor los fondos hacia las áreas pobres. Sin embargo, el proceso de evaluación permanente indica que el desempeño no es parejo y que incluso es persistentemente débil en algunas pocas provincias, conclusiones a las cuales en la actualidad se está prestando mayor atención de parte de las políticas. Los decepcionantes resultados de evaluación sobre la calidad de los proyectos de infraestructura han generado enormes efectos por parte del equipo del proyecto para mejorar los resultados en esta área a través de reformas de política. Se está insistiendo en realizar más visitas a terreno para evaluación y supervisión, sancionar a los organismos con resultados deficientes al finalizar el proyecto y fortalecer el manual de evaluación. Y por último, las evaluaciones sociales revelaron que es necesario prestar una mejor asistencia técnica a las ONG y municipalidades rurales durante la preparación e implementación de los proyectos, además de dar mayor publicidad y transparencia a la información sobre el programa TRABAJAR (Anexo 1.1).

**Reforma de autonomía escolar de Nicaragua.** En 1993, el gobierno nicaragüense adoptó medidas decisivas para implementar un importante esfuerzo de descentralización en el sector de la educación, otorgando autonomía presupuestaria y administrativa a un grupo de escuelas primarias y secundarias seleccionadas. La meta de las reformas es mejorar el aprendizaje de los alumnos: a medida que la administración de las escuelas sea más democrática y participativa, los patrones administrativos y de gasto de las escuelas locales se pueden orientar hacia esfuerzos que mejoren directamente la calidad pedagógica e impulsen el rendimiento de los alumnos. El efecto de esta reforma se ha evaluado usando una combinación de técnicas cuantitativas y cualitativas para examinar los resultados y el proceso de descentralización. El propósito del componente cualitativo es aclarar si en las escuelas se están efectivamente realizando las reformas administrativas y financieras previstas, y por qué y por qué no. El componente cuantitativo refuerza estos resultados al responder a la pregunta: "¿Los cambios administrativos y financieros en las escuelas realmente mejoran los resultados de aprendizaje en los niños?" Los resultados cualitativos indican que una implementación exitosa de las reformas depende en gran medida del marco y del entorno (es decir, el nivel de pobreza de la comunidad), mientras que los resultados cuantitativos indican que la mayor independencia en la toma de decisiones de las escuelas está asociado en forma significativa con un mejor rendimiento de los alumnos.

Se utilizaron metodologías y fuentes de datos diferentes, pero complementarias, para combinar con los enfoques. Por una parte, la evaluación cuantitativa siguió un diseño cuasi experimental, en el cual se comparan las calificaciones de las pruebas de una muestra de alumnos de escuelas autónomas (grupo de tratamiento) con los resultados de una muestra pareada de escuelas públicas no autónomas y escuelas privadas (grupo de comparación). Los datos para este componente de la evaluación se recopilaron de un panel de dos encuestas pareadas de escuelas-hogares y de las calificaciones de pruebas de rendimiento de

los alumnos. Por otra parte, el diseño de evaluación cualitativo consistía en una serie de entrevistas a informantes clave y conversaciones con grupos representativos con diferentes padres y miembros del personal de la escuela de una submuestra de escuelas autónomas y tradicionales incluida en el estudio cuantitativo.

Al usar técnicas de investigación cualitativas y cuantitativas, se generó una valiosa combinación de resultados útiles y pertinentes para las políticas. El trabajo cuantitativo proporcionó una amplia visión general y estadísticamente válida de las condiciones y resultados de las escuelas; en el trabajo cualitativo se perfeccionaron estas conclusiones con información sobre por qué algunos resultados esperados del programa de reforma han sido exitosos, mientras que otros han fracasado, lo que sirvió de base para los ajustes de la política. Además, como el trabajo cualitativo es más intuitivo, era más accesible y, por consiguiente, más interesante para el personal del Ministerio, lo que a su vez facilitó una rápida creación de capacidades para el proceso de evaluación dentro del Ministerio, aumentando su credibilidad (Anexo 1.11).

### **Aprovechamiento de las fuentes de datos existentes**

Las fuentes de datos existentes, como una encuesta domiciliaria nacional, censo, registros administrativos de un programa o datos municipales, pueden aportar valiosa información a los esfuerzos de evaluación. Aprovechar las fuentes existentes reduce la necesidad de recurrir a una costosa recopilación de datos para el exclusivo propósito de una evaluación, como se ilustra en el caso de la evaluación de Caminos Rurales de Vietnam. Además, aunque los datos existentes podrían no contener toda la información que idealmente se reuniría para los fines de la evaluación, hay técnicas de evaluación innovadoras que con frecuencia pueden compensar los datos faltantes, como se muestra en el Proyecto de Extensión Agrícola Nacional de Kenya.

**El Proyecto de caminos rurales de Vietnam.** Los datos que se usaron en esta evaluación se basaron en una combinación eficaz de los datos nacionales y locales existentes con las encuestas realizadas específicamente para los fines de la evaluación. La encuesta domiciliaria de la evaluación fue diseñada con eficiencia para repetir varias preguntas de la Encuesta Nacional sobre las Condiciones de Vida de Vietnam, de modo que, al basarse en información que es común para ambas encuestas, se pueden usar técnicas de regresión para estimar la posición de cada hogar en la distribución nacional del bienestar.

La evaluación se basa en gran medida en los datos a nivel de comunal recopilados anualmente por las comunas y que cubren aspectos demográficos, uso de la tierra y actividades productivas. Esta fuente de datos se ve aumentada con una encuesta a nivel comunal realizada para la evaluación. Se crearon dos bases de datos adicionales usando la información existente. Se estableció una extensa base de datos a nivel provincial para ayudar a comprender la selección de las provincias dentro del proyecto. Esta base de datos abarca todas las provincias de Vietnam y los datos cubren un gran número de variables socioeconómicas. Por último, también se creó una base de datos a nivel de proyecto para cada una de las áreas del proyecto en que se realizaron encuestas, para controlar por la

magnitud del proyecto y el método de implementación para evaluar el impacto del proyecto (Anexo 1.15).

**Proyecto de Extensión Nacional de Kenya (NEP).** Los resultados del Proyecto de Extensión Nacional de Kenya (NEP) han sido polémicos y forman parte de un debate mayor sobre la eficacia en función de los costos del enfoque de capacitación y visitas (T&V) frente a los servicios de extensión agrícolas. En el contexto de Kenya, se ha intensificado el debate debido, por una parte, los altísimos rendimientos estimados para la capacitación y visitas en un estudio (Bindlish, Evenson y Gbetibouo, 1993, 1997) y, por otra, la falta de resultados visibles convincentes, incluido el rendimiento deficiente de la agricultura de Kenya en los últimos años.

El desacuerdo con respecto a los resultados del Proyecto de Extensión Nacional de Kenya se mantiene en espera de los resultados de esta evaluación, la cual se diseñó para asumir un enfoque riguroso y empírico frente a la evaluación del desarrollo institucional del programa y su efecto en el rendimiento agrícola. En la evaluación se utiliza una combinación de métodos cualitativos y cuantitativos para plantear preguntas sumamente pertinentes para las políticas y revela serias falencias del programa: (a) el desarrollo institucional del NEP ha sido limitado y después de 15 años hay pocos progresos en la eficacia de sus servicios; (b) la calidad y cantidad de los servicios prestados es deficiente y (c) los servicios de extensión sólo tienen un pequeño efecto positivo en la eficiencia agrícola y ninguno en la productividad de las explotaciones agrícolas.

De los resultados de la evaluación se puede obtener una serie de conclusiones concretas para las políticas, muchos de las cuales son pertinentes para el diseño de futuros proyectos de extensión agrícola. En primer lugar, la evaluación revela la necesidad de mejorar la focalización de la capacitación y visitas, centrándose en las áreas y grupos en que el efecto probablemente sea mayor. Además, la asesoría se debe adaptar cuidadosamente a las demandas de los agricultores, tomando en cuenta las variaciones de las condiciones tecnológicas y económicas locales. Para poder lograr este nivel de focalización del servicio, se requieren flujos adecuados de una información oportuna y confiable y, por consiguiente, un sistema de supervisión y evaluación del programa que genere un flujo constante de retroalimentación desde los beneficiarios sobre el contenido del servicio. Para aumentar la eficiencia del programa, es probable que una presencia más discreta y menos intensa de la capacitación y visitas, pero de cobertura más amplia, sea más eficaz en función de los costos. El enfoque general del programa frente a la entrega del servicio con una sola metodología (visitas a las explotaciones agrícolas) para entregar mensajes estándares también limita la eficiencia del programa. Es probable que con la reforma institucional mejore la eficacia en la entrega del servicio. La descentralización del diseño del programa, incluidos mecanismos de integración que hagan participar al agricultor (como participación en los costos y organizaciones de agricultores), debe llegar a ser una parte integral del mecanismo de entrega. Por último, la recuperación de costos, aun cuando sea sólo parcial, proporcionaría los incentivos adecuados, abordaría los temas de la rendición de cuentas y control de calidad, haría que el servicio se guiara más por la demanda y fuera más sensible y proporcionaría cierto mejoramiento presupuestario (Anexo 1.8).

## Costos y financiamiento

Es indudable que una evaluación del impacto implica muchos costos, lo que explica por qué algunos países se muestran reticentes a financiar tales estudios. Estos costos incluyen la recopilación de datos y el valor del tiempo del personal para todos los miembros del equipo de evaluación. El financiamiento de una evaluación del impacto puede provenir de dentro del proyecto, de otros recursos gubernamentales, de una subvención para investigación o de un donante externo. La información sobre una muestra de evaluaciones del Banco Mundial indica que aunque muchos países asumieron la mayor parte de los costos de la evaluación, la implementación exitosa de una evaluación de impacto requirió considerables recursos externos más allá de los previstos en un préstamo o crédito para un proyecto. Estos recursos provinieron de una combinación de las siguientes fuentes: (a) préstamo o financiamiento de crédito del Banco Mundial para la recopilación y el procesamiento de datos; (b) el gobierno, a través de los sueldos pagados al personal local asignado al esfuerzo de evaluación (como se explica en el cuadro 4.1, estos costos de personal no se incluyeron en el cálculo de los costos de las evaluaciones realizadas para los casos que aquí se examinan, debido a las limitaciones de datos,); (c) subvenciones para investigación del Banco Mundial y subvenciones de donantes bilaterales que financiaron la asistencia técnica de asesores con los conocimientos especializados específicos requeridos para la evaluación y (d) el presupuesto de gastos generales del Banco Mundial, a través del tiempo del personal proporcionado para guiar la evaluación de impacto y que con frecuencia participa activamente en el trabajo analítico.

Aunque en pocas evaluaciones de impacto se documenta el costo de realizar el trabajo, en el Cuadro 4.2 se proporciona estimaciones de costos para una muestra de evaluaciones de impacto en que participó el Banco Mundial. Sin embargo, estas estimaciones de costos no incluyen el valor del tiempo del personal que aportaron al proceso las contrapartidas del país cliente (que puede ser significativo), porque no se contaba con esta información. Como referencia, en los 8 casos anteriores, fue habitual que hasta cinco miembros del personal estuvieran asignados al esfuerzo de evaluación durante varios años, un nivel de esfuerzo suficiente como para elevar significativamente el costo de la evaluación en muchos de los casos.

**Cuadro 4.2 Resumen de costos estimados de diversas evaluaciones de impacto del Banco Mundial**

<i>Proyecto</i>	<i>Costo estimado de la evaluación (US\$)<sup>a</sup></i>	<i>Costo como % del costo total del proyecto<sup>b</sup></i>	<i>Costo como % del préstamo del BIRF o crédito de la AIF<sup>b</sup></i>	<i>Desglose de los costos de evaluación (%)</i>			
				<i>Viaje<sup>c</sup></i>	<i>Personal del Banco Mundial</i>	<i>Asesores</i>	<i>Recopilación de datos</i>
Nicaragua Administración basada en las escuelas	495.000	1,26%	1,5%	8,1%	18,1%	39,0%	34,8%
El Salvador Administración basada en las escuelas	443.000	0,60%	1,3%	7,7%	7,4%	25,8%	59,2%
Colombia Programa de vales	266.000	0,20%	0,3%	9,4%	9,8%	21,8%	59,0%
Honduras Fondo Social	263.000	0,23%	0,9%	3,0%	11,5%	53,2%	32,3%
Nicaragua Fondo Social	449.000	0,30%	0,8%	4,9%	33,0%	7,8%	55,7%
Bolivia Fondo Social	878.000	0,50%	1,4%	3,4%	14,6%	12,9%	69,1%
Trinidad y Tobago Capacitación de Jóvenes	238.000	0,80%	1,2%	7,6%	11,5%	17,9%	63,1%
<b>PROMEDIO</b>	<b>433.000</b>	<b>0,56%</b>	<b>1,0%</b>	<b>6,3%</b>	<b>15,1%</b>	<b>25,5%</b>	<b>53,3%</b>

a. Este costo no incluye el costo del personal local de contrapartida no financiados con el préstamo o crédito. Las cifras se refieren al período dentro del cual se seleccionaron los proyectos de la muestra de evaluaciones y no al financiamiento total proporcionado alguna vez por el Banco y otros a esas instituciones.

b. Estos costos, como porcentaje del préstamo o crédito o del proyecto, se presentan sólo a modo de referencia. En muchos casos, el financiamiento real de la evaluación se obtuvo de fuentes ajenas al financiamiento del proyecto.

c. Las estimaciones de costo de viajes incluyen el viaje de la misión del personal del Banco Mundial y de asesores internacionales a los países clientes, al igual que viajes de contrapartidas del país cliente, en especial para participar en las sesiones de estrategia y talleres analíticos con asesores internacionales y personal del Banco Mundial.

*Fuente:* Archivos de Proyectos del Banco Mundial.

El costo estimado promedio para la evaluación del impacto fue de US\$433.000. Esto refleja una gama de US\$263.000 para la evaluación de un programa de capacitación en habilidades vocacionales para jóvenes desempleados de Trinidad y Tobago y US\$878.000 para la evaluación del Fondo de Inversión Social de Bolivia. El gasto en las evaluaciones de impacto para los proyectos a continuación refleja, en promedio, un 0,6% del costo total del proyecto (que a veces incluía financiamiento de varios donantes) o un 1,3% del costo del préstamo del Banco Internacional de Reconstrucción y Fomento (BIRF) o del crédito de la Asociación Internacional de Fomento (AIF). Los componentes más costosos de las

evaluaciones indicadas a continuación fueron la recopilación de datos y los asesores locales e internacionales. En muchos de los casos, los costos de viajes incluyeron viajes del personal local para reunirse con personal del Banco Mundial e investigadores para sesiones de estrategia y capacitación, porque la creación de capacidades entre el personal del país cliente constituía un objetivo clave. Los dos ejemplos siguientes para las evaluaciones de impacto de proyectos en Trinidad y Tobago y Bolivia ilustran algunas de las fluctuaciones que pueden surgir en los costos de los programas.

La evaluación del programa de capacitación en habilidades vocacionales de Trinidad y Tobago aprovechó una encuesta nacional sobre ingreso y empleo para tomar muestras de los graduados del programa y crear un grupo de comparación de un subconjunto de la muestra nacional. Además, el equipo de evaluación ayudó a diseñar y usar los datos administrativos disponibles de registros de postulantes al programa, de modo que se contaba con datos previos a la intervención y no se requirió empadronamiento para ubicar a los graduados del programa. El tamaño total de la muestra para cada uno de los tres estudios de seguimiento fue de aproximadamente 2500 jóvenes, incluidos los grupos de tratamiento y de comparación. En la encuesta sólo se aplicó un breve cuestionario y sólo a los graduados del programa. Trinidad y Tobago es un país pequeño, las comunidades son relativamente accesibles por tierra y el idioma común es el inglés en el país y entre los graduados del programa.

La evaluación del Fondo Social de Bolivia (FIS) utilizó sus propias encuestas básicas y de seguimiento de los grupos de tratamiento y comparación para evaluar las intervenciones en salud, educación, agua y servicios sanitarios. No había encuestas nacionales disponibles con las cuales realizar los análisis o efectuar sobremuestreos y por ello la evaluación se llevó toda la carga de la recopilación de datos. La muestra de los grupos de tratamiento y de comparación se componía de 7.000 hogares y 300 establecimientos entrevistados en la encuesta básica de 1993 y en la encuesta de seguimiento de 1998.

En Bolivia, los instrumentos de recopilación de datos para la evaluación del impacto incluyeron laboratorios portátiles para realizar pruebas de calidad del agua en laboratorio, pruebas de rendimiento y ocho cuestionarios para declarantes de los hogares y establecimientos. Los ocho cuestionarios constaban de dos cuestionarios domiciliarios (uno para el declarante principal y uno para mujeres en edad fértil), un cuestionario para la comunidad, cuatro diferentes cuestionarios de centros de salud para los distintos tipos de éstos (los que fluctúan desde pequeños consultorios locales hasta hospitales) y un cuestionario para escuelas, para el director de éstas. Para evaluar la focalización, la evaluación incluyó una medida de la pobreza basada en el consumo, lo que planteó la necesidad de recopilar datos detallados sobre el consumo de los hogares, como asimismo datos regionales sobre precios de las comunidades. El trabajo en terreno se realizó en áreas rurales, donde se ubica la mayoría de los proyectos del FIS, e incluyó una muestra aleatoria de hogares rurales a los cuales con frecuencia sólo se podía acceder a pie o a caballo. Por último, los cuestionarios tuvieron que prepararse y aplicar en español, quechua y aymará.

## Asuntos de economía política

Existen varios aspectos de política económica que influyen no sólo en si se lleva a cabo una evaluación, sino también en la forma en que se implementa. La decisión de realizar una evaluación depende mucho de un sólido apoyo político. Muchos gobiernos no aprecian el valor de efectuar proyectos de evaluación y por lo tanto no desean invertir recursos en ellos. Además puede haber reticencia a permitir la ejecución de una evaluación independiente que pueda revelar resultados contrarios a la política del gobierno, especialmente en regímenes autoritarios o cerrados. Sin embargo, los gobiernos más abiertos podrían considerar que las evaluaciones y la divulgación de los resultados forman parte importante del proceso democrático.

Las evaluaciones también son sensibles al cambio político. Tres de los ocho estudios de efectos indicados en el Cuadro 4.2 se cancelaron debido a problemas de economía política. Las rotaciones de regímenes o de puestos clave dentro de un departamento gubernamental de contrapartida y los cambios en estrategias de política pueden afectar no sólo el esfuerzo de evaluación, sino más fundamentalmente la implementación del programa que se está evaluando. Un ejemplo de este tipo de riesgo lo experimentó un equipo que trabajaba en el diseño y evaluación del impacto de un programa piloto de administración basada en escuelas del Perú, como parte de un proceso de educación primaria financiado por el Banco Mundial. El equipo, compuesto por funcionarios del Ministerio de Educación, personal del Banco Mundial y asesores internacionales y locales, había trabajado por más de un año en la elaboración de los modelos de administración basada en escuelas que se probarían, en el establecimiento de un diseño experimental, en el diseño de instrumentos de encuestas y pruebas de rendimiento y en la recopilación de datos básicos sobre las características de las escuelas y el rendimiento escolar. Justo antes de la introducción de la prueba piloto en las escuelas seleccionadas en forma aleatoria, funcionarios gubernamentales de alto nivel cancelaron el experimento de administración basada en las escuelas en una reacción ante las repercusiones políticas percibidas del programa piloto. Varios años después se introdujo una reforma similar, pero sin el beneficio de una prueba piloto ni una evaluación.

En Venezuela, la evaluación de un programa de salud y nutrición materno-infantil se rediseñó tres veces con tres contrapartidas diferentes del cliente a medida que el gobierno transfería la responsabilidad de la evaluación de un organismo a otro. Cada cambio estuvo acompañado de una renegociación de contrato con la empresa del sector privado que se había identificado para realizar la recolección de datos y la mayor parte de los análisis para la evaluación. Cuando la legitimidad de la tercera contrapartida del gobierno comenzó a ser cuestionada, la empresa anuló el contrato y se abandonó la evaluación. Estos incidentes se produjeron durante un período de cambio político caracterizado por varias reorganizaciones de gabinete, las que terminaron con el fracaso del gobierno elegido que servía como contrapartida del proyecto. De modo que la evaluación difícilmente era la única que se veía afectada por las repercusiones de la inestabilidad política. Sin embargo, en los casos de Perú y Venezuela, es sensato reflexionar sobre la cantidad de recursos destinados a un esfuerzo que nunca se materializó. Un ejemplo menos dramático del efecto del cambio político en las estrategias de evaluación proviene de El Salvador, donde el éxito reconocido

de una reforma introducida en escuelas rurales hizo que el gobierno introdujera una reforma educacional similar en todas las escuelas urbanas de una sola vez, en lugar de introducirla gradualmente en las escuelas en el lapso de tiempo planificado originalmente. Esta decisión no permitió usar un diseño experimental y sólo se pudo usar una comparación reflexiva menos sólida como la única opción de diseño viable.

## Bibliografía

La palabra *procesado* es una descripción de los trabajos reproducidos informalmente que podrían no estar disponibles habitualmente en bibliotecas.

Abadie, A., J. Angrist y G. Imbens. 1998. "Instrumental Variables Estimation of Quartile Treatment Effects." National Bureau of Economic Research, Serie de documentos de trabajo, N° 229:1–28.

Atkinson, Anthony. 1987. "On the Measurement of Poverty." *Econometrica* 55: 749–64.

Bamberger, Michael. 2000. *Integrating Quantitative and Qualitative Methods in Development Research*. Washington, D.C. Banco Mundial

Barnes, Carolyn. 1966. "Assets and the Impact of Microenterprise Finance Programs." USAID AIMS Project Brief 6. Washington, D.C.: USAID

Barnes, Carolyn, y Erica Keogh. 1999. "An Assessment of the Impact of Zambuko's Microenterprise Program in Zimbabwe: Baseline Findings." USAID AIMS Project Brief 23. Washington, D.C.: USAID

Benus, Jacob, Neelima Grover y Recep Varcin. 1998. *Turkey: Impact of Active Labor Programs*. Bethesda, MD: Abt Associates.

Benus, Jacob, Neelima Grover, Jiri Berkovsky y Jan Rehak. 1998. *Czech Republic: Impact of Active Labor Programs*. Bethesda, MD: Abt Associates.

Besharov, Douglas J., Peter Germanis y Peter H. Rossi. 1997. *Evaluating Welfare Reform: A Guide for Scholars and Practitioners*. College Park, Md.: University of Maryland.

Bindlish, V., R. Evenson y Mathurin Gbetibouo. 1993. *Evaluation of T & V-Based Extension in Burkina Faso*. Washington, D.C. Banco Mundial

Bloom, Howard S., Larry L. Orr, George Cave, Stephen H. Bell, Fred Doolittle y Winston Lin. 1994. *The National JTPA Study: Overview, Impacts, Benefits, and Costs of Title II-A*. Bethesda, Md.: Abt Associates.

Blundell, Richard W., y R. J. Smith. 1993. "Simultaneous Microeconomic Models with Censoring or Qualitative Dependent Variables." En G. S. Maddala, C. R. Rao y H. D. Vinod, eds., *Handbook of Statistics*. Vol. 11. Amsterdam: Holanda del Norte.

Bourguignon, Francois, Jaime de Melo y Akiko Suwa. 1991. "Distributional Effects of Adjustment Policies: Simulations for Archetype Economies in Africa and Latin America." *World Bank Economic Review* 5 (2): 339–66.

- Burtless, Gary. 1995. "The Case for Randomized Field Trials in Economic and Policy Research." *Journal of Economic Perspectives* 9 (Primavera): 63–84.
- Card, David, y Philip K. Robins. 1996. "Do Financial Incentives Encourage Welfare Recipients to Work? Evidence from a Randomized Evaluation of the Self-Sufficiency Project." National Bureau of Economic, Documento de estudio 5701, agosto. Cambridge, Mass.: NBER.
- Carvalho, Soniya, y Howard White. 1994. "Indicators for Monitoring Poverty Reduction." Documentos de trabajo 254 del Banco Mundial. Washington, D.C.: Banco Mundial
- Chen, Martha Alter, y Donald Snodgrass. 1999. "An Assessment of the Impact of Sewa Bank in India: Baseline Findings." USAID AIMS Project, procesado. Washington, D.C. USAID
- Cohen, Monique, y Gary Gaile. 1997. "Highlights and Recommendations of the Virtual Meeting of the CGAP Working Group on Impact Assessment Methodologies." USAID AIMS Project. Washington, D.C.: Management Systems International.
- . 1998. "Highlights and Recommendations of the Second Virtual Meeting of the CGAP Working Group on Impact Assessment Methodologies." USAID AIMS Project. Washington, D.C.: Management Systems International.
- Dar, Amit, e Indermit S. Gill. 1998. "Evaluating Retraining Programs in OECD Countries: Lessons Learned." *The World Bank Research Observer* 13 (febrero): 79–101.
- Dar, Amit, y Zafiris Tzannatos. 1999. "Active Labor Market Programs: A Review of Evidence from Evaluations." Documento de trabajo 9901 de protección social del Banco Mundial. Washington, D.C.: Banco Mundial
- Dehejia, Rajeev H., y Sadek Wahba. 1998. "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs." NBER, Serie de documentos de trabajo 6586. Cambridge, Mass.: NBER <http://www.nber.org/papers/w6586>.
- Dennis, Michael L., y Robert F. Boruch. 1989. "Randomized Experiments for Planning and Testing Projects in Developing Countries: Threshold Conditions." *Evaluation Review* 13 (junio): 292–309.
- Diagne, Aliou, y Manfred Zeller. 1998. "Determinants of Access to Credit and Its Impact on Income and Food Security in Malawi." Manuscrito presentado al Comité de revisión de publicaciones del Instituto Internacional de Investigaciones sobre Políticas Alimentarias (IFPRI) para su consideración como un Informe de Investigación IFPRI.
- Diop, F., A. Yazbeck, y R. Bitran. 1995. "The Impact of Alternative Cost Recovery

- Schemes on Access and Equity in Niger.” *Health Policy and Planning* 10 (3): 223–40.
- Donecker, Jane, y Michael Green. 1998. *Impact Assessment in Multilateral Development Institutions*. Londres: Department for International Development.
- Dunn, Elizabeth. 1999. “Microfinance Clients in Lima, Peru: Baseline Report for AIMS Core Impact Assessment.” USAID AIMS Project, procesado. Washington, D.C.: USAID
- Edgecomb, Elaine L. y Carter Garber. 1998. “Practitioner-Led Impact Assessment: A Test in Honduras.” USAID AIMS. Washington, D.C.: USAID.
- Ezemenari, Kene, Anders Ruqvist y K. Subbarao. 1999. “Impact Evaluation: A Note on Concepts and Methods.” World Bank Poverty Reduction and Economic Management Network, procesado. Washington, D.C.: Banco Mundial.
- Foster, James, J. Greer y Erik Thorbecke. 1984. “A Class of Decomposable Poverty Measures.” *Econometrica* 52: 761–65.
- Friedlander, Daniel, y Gayle Hamilton. 1993. *The Saturation Work Initiative Model in San Diego: A Five-Year Follow-Up Study*. Nueva York: Manpower Demonstration Research Corporation.
- Friedlander, Daniel, y Philip K. Robins. 1995. “Evaluating Program Evaluations: New Evidence on Commonly Used Nonexperimental Methods.” *American Economic Review* 85 (septiembre): 923–37.
- Friedlander, Daniel, David H. Greenberg y Philip K. Robins. 1997. “Evaluating Government Training Programs for the Economically Disadvantaged.” *Journal of Economic Literature* 35 (diciembre): 1809–55.
- Fuller, Bruce, y Magdalena Rivarola. 1998. *Nicaragua’s Experiment to Decentralize Schools: Views of Parents, Teachers and Directors*. Serie de documentos de trabajo sobre Evaluación de reformas de la educación, documento N° 5. Washington, D.C.: Banco Mundial.
- Gaile, Gary L., y Jenifer Foster. 1996. “Review of Methodological Approaches to the Study of the Impact of MicroEnterprise Credit Programs.” USAID AIMS Brief 2. Washington, D.C.: USAID
- Glewwe, Paul, Michael Kremer y Sylvie Moulin. 1998. “Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya.” Development Research Group, Banco Mundial, 28 de octubre
- Goodman, Margaret, Samuel Morley, Gabriel Siri y Elaine Zuckerman. 1997. *Social Investment Funds in Latin America: Past Performance and Future Role*. Washington,

D.C. Banco Interamericano de Desarrollo. Marzo.

Gobierno de Dinamarca. 1995. *Methods for Evaluation of Poverty Oriented Aid Interventions*. Copenhague: Ministerio de Asuntos Exteriores.

Greene, W. H. 1997. *Econometric Analysis*. Hemel Hempstead, New Jersey: Prentice Hall Press.

Greenberg, David, y Mark Shroder. 1997. *The Digest of Social Experiments*, segunda edición. Washington, D.C.: The Urban Institute Press.

Grosh, Margaret E., y Juan Muñoz. 1996. "A Manual for Planning and Implementing the Living Standards Measurement Study Survey." LSMS, Documento de trabajo N° 126. Banco Mundial, Washington, D.C.

Grossman, Jean Baldwin. 1994. "Evaluating Social Policies: Principles and U.S. Experience." *The World Bank Research Observer* 9 (julio): 159–80.

Habicht, J. P., C. G. Victoria y J. P. Vaughan. 1999. "Evaluation Designs for Adequacy, Plausibility and Probability of Public Health Programme Performance and Impact." *International Journal of Epidemiology* 28: 10–18.

Harrell, Adele, Martha Burt, Harry Hatry, Shelli Rossman, Jeffrey Roth y William Sabol. 1996. *Evaluation Strategies for Human Service Programs: A Guide for Policymakers and Providers*. Washington, D.C.: Urban Institute.

Heckman, James. 1997. "Instrumental Variables. A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *Journal of Human Resources* 32 (3): 441–61.

Heckman, James, R. Lalonde y J. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." En O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, vol. III. Nueva York: Elsevier Science Publishing Co.

Heckman, James, y Richard Robb. 1985. "Alternative Methods of Evaluating the Impact of Interventions: An Overview." *Journal of Econometrics* 30: 239–67.

Heckman, James J., y Jeffrey A. Smith. 1995. "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9 (2): 85–110.

Heckman, J., H. Ichimura, J. Smith y P. Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66: 1017–99.

Hicks, Norman. 1998. *Measuring the Poverty Impact of Projects in LAC*. Región de América Latina y el Caribe del Banco Mundial, procesado. Washington, D.C.:

Banco Mundial.

Holder, Harold D., Andrew J. Treno, Robert F. Saltz y Joel W. Grube. 1997. "Summing Up: Recommendations and Experiences for Evaluation of Community-Level Prevention Programs." *Evaluation Review* 21 (abril): 268–78.

Hollister, Robinson G., y Jennifer Hill. 1995. *Problems in the Evaluation of Community-Wide Initiatives*. Nueva York: Russell Sage Foundation.

Hulme, David. 1997. "Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice." Manchester, R.U.: Institute for Development Policy and Management, University of Manchester.

Instituto Internacional de Investigaciones sobre Políticas Alimentarias. 1998. *Programa Nacional de Educación, Salud, y Alimentación (PROGRESA): A Proposal for Evaluation* (con apéndice técnico). Washington, D.C.: IFPRI.

Jalan, Jyotsna, y Martin Ravallion. 1998a. "Income Gains from Workfare: Estimates for Argentina's TRABAJAR Program Using Matching Methods." Washington, D.C. Development Research Group, Banco Mundial.

———. 1998b. "Transfer Benefits from Workfare: A Matching Estimate for Argentina." Washington, D.C.: Development Research Group, Banco Mundial.

Jamison, Dean T., Barbara Serle, Klaus Galda y Stephen P. Heyneman. 1981. "Improving Elementary Mathematics Education in Nicaragua: An Experimental Study of the Impact of Textbooks and Radio on Achievement." *Journal of Educational Psychology* 73 (4): 556–67.

Karoly, Lynn A., Peter W. Greenwood, Susan S. Everingham, Jill Houbé, M. Rebecca Kilburn, C. Peter Rydell, Matthew Sanders y James Chiesa. 1998. *Investing in Our Children: What We Know and Don't Know about the Costs and Benefits of Early Childhood Interventions*. Santa Monica, Calif.: Rand.

Kemple, James J., Fred Doolittle y John W. Wallace. 1993. *The National JTPA Study: Site Characteristics and Participation Patterns*. Nueva York: Manpower Demonstration Research Corporation.

Khandker, Shahidur R. 1998. *Fighting Poverty with Microcredit: Experience in Bangladesh*. Nueva York: Oxford University Press para el Banco Mundial.

Killick, Tony. 1995. *IMF Programmes in Developing Countries, Design and Impact*. Londres: Inglaterra.

King, Elizabeth, y Berk Ozler. 1998. *What's Decentralization Got to Do with Learning?*

*The Case of Nicaragua's School Autonomy Reform*. Serie de documentos de trabajo sobre Evaluación de reformas de la educación, documento N° 9. Washington, D.C.: Banco Mundial.

King, Elizabeth, Berk Ozler y Laura Rawlings. 1999. *Nicaragua's School Autonomy Reform: Fact or Fiction?* Washington, D.C.: Banco Mundial.

Kish, Leslie. 1965. *Survey Sampling*. Nueva York: John Wiley and Sons.

Levinson, James F., Beatrice Lorge Rogers, Kristin M. Hicks, Thomas Schaetzel, Lisa Troy y Collette Young. 1998. *Monitoring and Evaluation: A Guidebook for Nutrition Project Managers in Developing Countries*. Medford, Mass.: International Food and Nutrition Center, Tufts University School of Nutrition Science and Policy, marzo.

Manski, Charles, e Irwin Garfinkel, eds. 1992. *Evaluating Welfare and Training Programs*. Cambridge, Mass.: Harvard University Press.

Manski, Charles, y Steven Lerman. 1977. "The Estimation of Choice Probabilities from Choice-Based Samples." *Econometrica* 45: 1977–88.

McKay, H., A. McKay, L. Siniestra, H. Gomez y P. Lloreda. 1978. "Improving Cognitive Ability in Chronically Deprived Children" *Science* 200(21): 270–78.

Meyer, Bruce D. 1995. "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistics*, 13 (abril): 151–161.

Miles, Matthew B., y A. Michael Huberman. 1994. *Qualitative Data Analysis*. Londres: Sage Publications.

MkNelly, Barbara, y Karen Lippold. 1998. "Practitioner-Led Impact Assessment: A Test in Mali." USAID AIMS Brief 21. Washington, D.C.: USAID

MkNelly, Barbara, y Christopher Dunford (en colaboración con el Program in International Nutrition, University of California, Davis). 1998. "Impact of Credit with Education on Mothers' and Their Young Children's Nutrition: Lower Pra Rural Bank Credit with Education Program in Ghana." *Freedom from Hunger*, Documento de Estudio N° 4, marzo.

Moffitt, Robert. 1991. "Program Evaluation with Nonexperimental Data." *Evaluation Review* 15 (3): 291–314.

Mohr, Lawrence B. 1995. *Impact Analysis for Program Evaluation*, segunda edición.

Thousand Oaks, Calif.: Sage Publications.

- . 1999. “The Qualitative Method of Impact Analysis.” Procesado.
- Morduch, Jonathan. 1998a. “The Microfinance Schism.” Harvard Institute for International Development, Documento de Análisis de Desarrollo 626, febrero.
- . 1998b. “Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh.” Harvard University Institute for International Development, procesado, junio.
- . 1999. “The Microfinance Promise.” *Journal of Economic Literature* 37(4): 1569–1614.
- Newman, John, Menno Pradhan, Laura Rawlings, Ramiro Coa y Jose Luis Evia. 2000. “An Impact Evaluation of Education, Health, and Water Supply Investments of the Bolivia Social Investment Fund.” Procesado.
- Newman, John, Laura Rawlings y Paul Gertler. 1994. “Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries.” *The World Bank Research Observer* 9 (Julio): 181–202.
- Operations Evaluation Department, Banco Mundial. 1994. “An Overview of Monitoring and Evaluation in the World Bank.” 30 de junio.
- Poppele, J. S. Summarto, y L. Pritchett. 1999. “Social Impacts of the Indonesia Crisis: New Data and Policy Implications.” Social Monitoring Early Response Unit, Banco Mundial, Washington, D.C. Procesado.
- Pradhan, Menno, Laura Rawlings y Geert Ridder. 1998. “The Bolivian Social Investment Fund: An Analysis of Baseline Data for Impact Evaluation.” *World Bank Economic Review* 12 (3): 457–82.
- Ravallion, Martin. 1994. *Poverty Comparisons, Fundamentals in Pure and Applied Economics*, vol. 56. Harwood Academic Publishers.
- . 1999. *Monitoring Targeting Performance When Decentralized Allocations to the Poor Are Unobserved*. Banco Mundial, Washington, D.C. Procesado.
- Ravallion, Martin, y Quentin Wodon. 1998. *Evaluating a Targeted Social Program When Placement Is Decentralized*. Documento de Trabajo de Estudio de Políticas del Banco Mundial, 1945. Washington, D.C.: Banco Mundial.
- Ravallion, Martin, Dominique van de Walle y Madhur Gautam. 1995. “Testing a Social Safety Net.” *Journal of Public Economics* 57: 175–99.
- Rawlings, Laura. 2000. “Assessing Educational Management and Quality in Nicaragua.” En

Bamberger: *Integrating Quantitative and Qualitative Methods in Development Research*. Washington, D.C.: Banco Mundial.

- Rebien, Claus C. 1997. "Development Assistance Evaluation and the Foundations of Program Evaluation." *Evaluation Review* 21 (agosto): 438–60.
- Revenga, Ana, Michelle Riboud y Hong Tan. 1994. "The Impact of Mexico's Retraining Program on Employment and Wages." *World Bank Economic Review* 8 (2): 247–77.
- Rivers, Douglas, y Quang H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39: 347–66.
- Robinson, Sherman. 1991. "Macroeconomics, Financial Variables, and Computable General Equilibrium Models." *World Development* 19(11): 1509–25.
- Rosenbaum, P., y D. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70: 41–55.
- . 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score." *American Statistician* 39: 35–39.
- Rossmann, Gretchen B., y Bruce L. Wilson. 1985. "Numbers and Words: Combining Quantitative and Qualitative Methods in a Single Large-Scale Evaluation Study." *Evaluation Review* 9 (3): 627–43.
- Sahn, David, Paul Dorosh y Stephen Younger. 1996. "Exchange Rate, Fiscal and Agricultural Policies in Africa: Does Adjustment Hurt the Poor?" *World Development* 24(4): 719–47.
- Sebstad, Jennifer, y Gregory Chen. 1996. "Overview of Studies on the Impact of MicroEnterprise Credit." USAID AIMS, junio. Washington, D.C.: USAID
- Sharma, Manohar, y Gertrud Schrieder. 1998. "Impact of Finance on Food Security and Poverty Alleviation—A Review and Synthesis of Empirical Evidence." Documento presentado en Workshop on Innovations in Micro-Finance for the Rural Poor: Exchange of Knowledge and Implications for Policy, Ghana, noviembre.
- Smith, William. 1998. *Group Based Assessment of Change: Method and Results*. Hanoi, Vietnam: ActionAid.
- Subbarao, K., Kene Ezemenari, John Randa y Gloria Rubio. 1999. *Impact Evaluation in FY98 Bank Projects: A Review*. World Bank Poverty Reduction and Economic Management Network, procesado, enero.

Tan, J.P., J. Lane y G. Lassibille. 1999. "Schooling Outcomes in Philippine Elementary Schools: Evaluation of the Impact of Four Experiments." *World Bank Economic Review*, septiembre.

Taschereau, Suzanne. 1998. *Evaluating the Impact of Training and Institutional Development Programs, a Collaborative Approach*. Economic Development Institute of the World Bank, enero.

USAID. 1984. "Study of Low Cost Housing." Nueva Delhi: USAID. Junio.

———. 1995. "Assessing the Impacts of MicroEnterprise Interventions: A Framework for Analysis." MicroEnterprise Development Brief, junio.

Valadez, J., y M. Bamberger, ed. 1994. "Monitoring and Evaluating Social Programs in Developing Countries." Economic Development Institute of the World Bank Series, Banco Mundial, Washington, D.C.

van de Walle, Dominique. 1999. "Assessing the Poverty Impact of Rural Road Projects." Banco Mundial. Procesado.

Walker, Ian, Rafael del Cid, Fidel Ordóñez y Florencia Rodríguez. 1999. "Evaluación Ex-Post del Fondo Hondureño de Inversión Social (FHIS 2)." ESA Consultores, Tegucigalpa, Honduras.

Weiss, Carol H. 1998. *Evaluation*. Upper Saddle River, New Jersey: Prentice Hall.

Wilson, Sandra Jo. 1998. "A Multi-Method Proposal to Evaluate the Impact of CGIAR-Instigated Activities on Poverty Reduction." Documento presentado en la Conferencia Anual de la American Evaluation Association, noviembre.

Banco Mundial, Poverty and Human Resources Division, Policy Research Department. 1998a. *Impact Evaluation of Education Projects Involving Decentralization and Privatization*. Documento de trabajo. Washington, D.C.: Banco Mundial.

———, Learning and Leadership Center and Operations Policy Department. 1996. *Handbook on Economic Analysis of Investment Operations*. Washington, D.C.: Banco Mundial.

———, Operations and Evaluation Department. 1998a. *1998 Annual Review of Development Effectiveness*. Washington, D.C.: Banco Mundial.

———, Poverty Reduction and Economic Management, Latin America and Caribbean Region. 1998b. *Measuring the Poverty Impact of Projects in LAC*. Washington, D.C.: Banco Mundial.

- , Operations and Evaluation Department. 1999. *World Bank Agricultural Extension Projects in Kenya: An Impact Evaluation*. Informe N° 19523. Washington, D.C.: Banco Mundial.
- , Poverty Reduction and Economic Management Unit, 1999. *Strategy for Impact Evaluation*. Washington, D.C.: Banco Mundial.
- Wouters, A. 1995. "Improving Quality through Cost Recovery in Niger." *Health Policy and Planning* 10 (3): 257–70.
- Zeller, Manfred, Akhter Ahmed, Suresh Babu, Sumiter Broca, Aliou Diagne y Manohar Sharma. 1996. "Security of the Poor: Methodologies for a Multicountry Research Project." IFPRI, Documento de trabajo 11.