

Apéndice C

Seguimiento y evaluación: Notas técnicas y casos prácticos

Nota técnica C.1	Principales tipos de evaluaciones.....	1
Nota técnica C.2	Diseños de evaluación de impactos	3
Nota técnica C.3	Métodos de evaluación de impactos para políticas y programas de cobertura integral	4
Nota técnica C.4	Tipos de fuentes de datos para evaluación de impactos	4
Caso práctico C.1	Seguimiento de la evolución del Plan de Acción para la Erradicación de la Pobreza (PEAP) en Uganda.....	5
Caso práctico C.2	Plan propuesto para el seguimiento de la Estrategia de Lucha contra la Pobreza en Tanzania.....	14
Caso práctico C.3	Encuestas de opinión a ciudadanos como herramienta para la participación de la sociedad civil en la evaluación del desempeño del sector público: el caso de Bangalore, India	22
Caso práctico C.4	Evaluación de los beneficios para los pobres de la asistencia a los trabajadores: el programa Trabajar de Argentina.....	23
Caso práctico C.5	Evaluación del proyecto de extensión agrícola de Kenia.....	27
Caso práctico C.6	Evaluación de la reforma escolar de Nicaragua: un método cuantitativo-cualitativo combinado	30
Caso práctico C.7	Rendimiento escolar en las escuelas elementales de Filipinas: evaluación del impacto de cuatro experimentos	33
Notas.....		37

Tablas

C.1. Resumen de los casos prácticos de evaluación de impactos	7
C.2. Lista modificada de parámetros de seguimiento, Uganda	9
C.3. Parámetros propuestos para el seguimiento del DELP, Tanzania.....	15
C.4. Fuentes de información para el seguimiento, Tanzania.....	19

Figura

Figura C.1. Marco institucional para el seguimiento del DELP, Tanzania.....	21
---	----

Cuadro

Cuadro C.1. Objetivos del Plan de Acción para la Erradicación de la Pobreza (PEAP) de Uganda	8
--	---

Nota técnica C.1 Principales tipos de evaluaciones

La evaluación es un análisis sistemático de la relevancia, funcionamiento y resultados de programas y políticas, comparados con una serie de estándares explícitos o implícitos, y que tienen por objeto mejorar las intervenciones públicas. Los distintos tipos de evaluaciones abordan distintas cuestiones de evaluación. *Grosso modo*, estas cuestiones se pueden clasificar en tres categorías:

- Las cuestiones de **proceso** pretenden entender si un programa, o un componente específico del mismo está siendo implementado tal y como se concibió originalmente.
- Las cuestiones de resultados pretenden evaluar si la situación de los individuos o unidades familiares han cambiado o no en lo que respecta a sus resultados básicos (conocimientos, conducta, bienestar, etc.), y en qué medida es el programa el responsable de los cambios observados. Los resultados pueden cambiar por una serie de razones, muchas de las cuales pueden no tener nada que ver con el programa. Las cuestiones de *atribución* analizan si los cambios observados fueron causados por el programa o si se habrían producido de todos modos.
- Las cuestiones de **motivos** pretenden saber cuáles son los motivos de los procesos y resultados observados, y estudian cómo y por qué los resultados fueron los que han sido.

Estas cuestiones pueden, más o menos, equipararse a los tres principales tipos de evaluaciones: evaluación de proceso, evaluación de resultados y evaluación basada en teorías. A su vez, cada tipo de evaluación tiene un menú de posibles diseños de evaluación y de métodos de recogida de datos. Los diseños de evaluación son conjuntos de técnicas que se pueden utilizar, en diferentes combinaciones, para responder a las distintas cuestiones de evaluación. Los diseños de evaluación especifican las unidades de análisis (por ejemplo, unidades familiares, individuos, instalaciones, comunidades, etc.) y cómo van a seleccionarse (de manera oportunista o bien utilizando estrategias de muestreo sistemático); qué tipo de comparaciones van a hacerse (por ejemplo, comparaciones en tiempo y espacio, comparación de diferentes grupos, ninguna comparación, etc.); y el momento de la recogida de datos (por ejemplo, antes y después del programa, sólo inmediatamente después del programa, durante la implementación del programa, etc.).

Las **evaluaciones de proceso** valoran en qué medida se está implementando con eficacia una intervención pública; se centran en aspectos tales como quién está participando, qué actividades se ofrecen, qué medidas se han adoptado, cuáles son las prácticas del personal a cargo y cuáles las reacciones de los beneficiarios. Una evaluación de proceso se puede efectuar en el caso de surgir problemas tales como retrasos, exralimitación en los costos o detección de insatisfacción de los beneficiarios por el sistema de seguimiento, o bien realizarse periódicamente como sistema de alerta temprana. Las evaluaciones de proceso tienden a basarse en diseños de evaluación y métodos de consulta menos formales, como autoevaluación y dictámenes de expertos.

Las **evaluaciones de resultados** estudian qué les ha ocurrido a los individuos (u otras unidades de análisis) después de implementar una política o programa; se centran en los resultados de la intervención, como por ejemplo si la gente es más sana, mejor educada o menos vulnerable a los vaivenes adversos. Los diseños de las evaluaciones de resultados pueden variar en una serie de niveles de complejidad. En un extremo del espectro encontramos las evaluaciones de resultados que se limitan a evaluar si los participantes en un programa experimentaron algún cambio en los parámetros básicos del bienestar. Se trata básicamente de ejercicios de seguimiento. Los diseños de evaluación y los métodos de recogida y análisis de datos en este extremo del espectro tienden a ser relativamente sencillos y dan resultados rápidamente, aunque dejan lugar a interpretaciones divergentes en cuanto a la magnitud del cambio que puede atribuirse a determinada intervención. Por lo general, este tipo de evaluación observa sólo al grupo de participantes en el programa. No los compara con personas o comunidades que no participan en el programa ni se hace ningún esfuerzo por aislar los efectos del programa o política de otros eventos que se producen simultáneamente. La evaluación puede observar los resultados después de que la intervención haya estado en vigencia durante un tiempo o haya concluido, o bien antes y después de la evaluación. Los métodos de recogida y análisis de datos pueden ser cuantitativos (encuestas de prestación de servicios), cualitativos (entrevistas a informadores importantes o grupos ad hoc) o participativos (métodos de evaluación rápidos).

En el otro extremo del espectro encontramos las evaluaciones que abordan cuestiones de atribución empleando técnicas especiales (normalmente complejas) para averiguar cuál ha sido el beneficio neto de las intervenciones (véanse las notas técnicas 2 y 3). Normalmente, estas evaluaciones se denominan evaluaciones de impacto. Las evaluaciones de impacto valoran en qué medida las intervenciones públicas han producido los efectos previstos, y en qué medida los cambios observados en el bienestar de los individuos pueden atribuirse a determinado programa o política. Estiman la magnitud de los efectos de un programa o política, y asignan causalidad. Dicho análisis de causalidad es esencial para comprender el papel relativo de las intervenciones programáticas alternativas en la reducción de la pobreza y, por consiguiente, para diseñar estrategias de reducción de la pobreza adecuadas.

Las **evaluaciones basadas en teorías** analizan los vínculos entre factores causales, actividades y resultados, pretenden determinar si se ha producido un desglose y, de ser afirmativa la respuesta, dónde, cómo y cuándo. Presentan una teoría explícita o implícita acerca de cómo y por qué una intervención pública funcionará como una serie de micropasos, y los analizan secuencialmente para realizar un seguimiento del desarrollo de las hipótesis. Siguiendo la secuencia de pasos, este tipo de evaluación puede determinar si el proceso que va desde los factores causales hasta los resultados del programa ha fallado o no y, en su caso, dónde.

Nota técnica C.2 Diseños de evaluación de impactos

Los **diseños experimentales o aleatorizados** implican reunir a un conjunto de individuos (u otras unidades de análisis) con iguales derechos de participación y deseosos de participar en el programa, y dividirlo en dos grupos: los beneficiarios de la intervención (grupo de tratamiento) y los aún no beneficiarios de la misma (grupo de control). Por ejemplo, en algunos fondos sociales, los proyectos económicamente viables presentados por las comunidades son seleccionados aleatoriamente para recibir financiación durante la primera fase del proyecto (grupo de tratamiento), en tanto que los demás, a quienes la financiación se otorgará en una fase posterior, se pueden utilizar como grupo de control. Dado que los participantes en el programa son seleccionados aleatoriamente, toda diferencia con quienes no participan en el programa es puramente accidental. Por tal motivo, los diseños experimentados suelen ser considerados el método más confiable, capaz de proporcionar los resultados más fáciles de interpretar. No obstante, en la práctica este tipo de evaluación puede resultar difícil de implementar, entre otras cosas porque resulta difícil retrasar la prestación de beneficios a individuos con igual derecho a recibirlos (véase el caso práctico 7).

Otra opción es el **diseño cuasi experimental**. En aquellos casos en que la aleatorización no sea viable, se puede estructurar un grupo de comparación. Los dos métodos para crear un grupo de comparación son la comparación pareada y la comparación reflexiva. *Pareada* consiste en seleccionar no participantes que, en sus características esenciales, sean comparables a los participantes. Ello se efectúa sobre la base de unas pocas o de varias características, seleccionadas mediante técnicas estadísticas. Por ejemplo, para la evaluación de Trabajar, un programa de obras públicas de Argentina, se creó un grupo de comparación emparejando participantes en el programa con no participantes sobre la base de diversas características socioeconómicas, como escolarización, género, vivienda, percepciones subjetivas del bienestar y pertenencia a partidos políticos (véase el caso práctico 4). Las evaluaciones en las que se emplean métodos de comparación pareada suelen ser más fáciles y económicas de implementar que los diseños experimentales, aunque los resultados son menos confiables e interpretarlos resulta más difícil.

Otro tipo de diseño cuasi experimental es el denominado **comparación reflexiva**. En una comparación reflexiva, el modelo contrafactual se estructura sobre la base de la situación de los participantes del programa antes de implementar el mismo. Así, los participantes en el programa son comparados consigo mismos antes y después de la intervención, y constituyen a la vez el grupo de tratamiento y el grupo de comparación. Este tipo de diseño resulta especialmente útil en evaluaciones de intervenciones de cobertura integral, como políticas y programas nacionales en los que participa toda la población y en los que no hay ámbito para un grupo de control (véase el caso práctico 5). Sin embargo, este método presenta una importante desventaja: la situación de los participantes en el programa antes y después de la intervención puede cambiar debido a motivos que nada tienen que ver con el programa. Por ejemplo, los participantes en un programa de capacitación pueden haber mejorado sus perspectivas de empleo después del programa. Aunque esta mejora bien puede ser resultado del programa, también puede deberse al hecho de que la economía se esté recuperando de una crisis y que el empleo esté creciendo otra vez. Si no se realizan de manera cuidadosa, las reflexiones comparativas no podrán diferenciar entre el programa y otros efectos externos, lo que compromete la confiabilidad de los resultados.

Los diseños no experimentales se pueden utilizar cuando no existe posibilidad de seleccionar un grupo de control o un grupo de comparación. Los participantes en el programa pueden ser comparados con los no participantes empleando métodos estadísticos para cuantificar las diferencias entre ambos grupos. Empleando el análisis de regresión es posible “controlar” la edad, renta, género y otras características de los participantes. Al igual que con los métodos cuasi experimentales, este diseño de evaluación resulta relativamente económico y fácil de implementar, aunque la interpretación de los resultados no es sencilla y los resultados pueden ser menos confiables.

Nota técnica C.3 Métodos de evaluación de impactos para políticas y programas de cobertura integral

Cuando las políticas o programas afectan a la totalidad de la población, no suele ser posible identificar o estructurar un grupo de control, por lo que resulta considerablemente más difícil evaluar si dichas intervenciones han provocado o no cambios en los resultados. Se pueden utilizar diversos métodos.

Los modelos de equilibrio general calculable (CGE, por sus siglas en inglés) intentan comparar los resultados en las situaciones observada y contrafactual empleando simulaciones informáticas. Estos modelos pretenden realizar un seguimiento del funcionamiento de la economía real y por lo general se basan en matrices de evaluación social basadas en las cuentas nacionales, encuestas de gasto de unidades familiares y otros gastos de encuestas. Los modelos CGE simulan el elemento contrafactual, aunque la solidez del modelo dependerá enteramente de la calidad de los datos en que está basado y en la validez de las hipótesis. Esto podría ser problemático, ya que las bases de datos suelen estar incompletas y muchos de los parámetros necesarios no pueden ser estimados mediante métodos econométricos formales. Además, los modelos CGE requieren muchísimo tiempo, son engorrosos y su desarrollo, caro.

Las comparaciones “con y sin” comparan el comportamiento de variables básicas en una muestra de regiones o países en los que se ha aplicado el programa con el comportamiento de las mismas en países en los que no se ha aplicado (grupo de comparación). Así, este método utiliza las experiencias de los países en los que no se ha aplicado el programa como variable representativa de lo que hubiese ocurrido en los países en que sí se aplica. Una importante limitación de este concepto es que parte del supuesto de que sólo la adopción de determinada política o programa es la que diferencia a los países o regiones en los que se aplica el programa de las demás, y de que los factores externos afectan a ambos grupos por igual o bien que su impacto se puede identificar independientemente del impacto de la intervención.

Los métodos de control estadísticos consisten en regresiones que controlan las diferencias en las condiciones iniciales y en las políticas adoptadas en países o regiones en los que se implementó y no se implementó un programa. El método identifica las diferencias entre las regiones afectadas y no afectadas por el programa durante el período previo a la implementación del mismo y, a continuación, controla estadísticamente estas diferencias para identificar los efectos aislados de los programas en el período posterior a la implementación.

Fuente: Baker, Judy. 2000. “Evaluación del impacto de los proyectos de desarrollo en la pobreza: Manual para profesionales”. Directivas de desarrollo. Washington, D.C.: Banco Mundial.

Nota técnica C.4 Tipos de fuentes de datos para evaluación de impactos

Los conjuntos de datos longitudinales, o de muestreo, incluyen información sobre los mismos individuos (u otras unidades de análisis) en al menos dos momentos diferentes: uno previo a la intervención (los datos de referencia) y otro posterior. Los conjuntos de datos de muestreo son muy valiosos para la evaluación de programas, pero pueden ser caros y requerir una sustancial capacidad institucional (véanse el caso práctico 5 y el capítulo 1, “Medición y análisis de la pobreza”).

Los datos transversales recogen información de diferentes personas en diferentes momentos. Normalmente, las evaluaciones que utilizan este tipo de datos cuestan menos que los estudios que emplean información de más de un punto en el tiempo, pero como suele ser difícil distinguir si los cambios se debieron a las intervenciones o a otros factores, los resultados tienden a ser menos confiables, salvo en los diseños experimentales (véase el caso práctico 4).

Los datos de series temporales (históricos) recogen información sobre mediciones de resultados básicos a intervalos periódicos, tanto antes como después de la implementación del programa. Permiten analizar los cambios en las tendencias antes y después del programa. Sin embargo, para un análisis riguroso se requieren muchos puntos de datos antes y después de la implementación del programa. Las series temporales se emplean fundamentalmente para evaluar políticas y programas de cobertura integral o nacional.

Casos prácticos

Los casos prácticos C.1 y C.2 presentan ejemplos de sistemas nacionales de seguimiento de la pobreza, en tanto que el caso práctico C.3 presenta un ejemplo del uso de las encuestas de opinión ciudadana como herramienta de participación de la sociedad civil en la evaluación del rendimiento del sector público. Los casos prácticos C.4 a C.7 (adaptados de la obra de Judy Baker [2000], *Evaluación del impacto de los proyectos de desarrollo en la pobreza: Manual para profesionales*, Banco Mundial, Washington, D.C.) presentan ejemplos de evaluaciones del impacto de programas y proyectos sobre diferentes sectores. Muestran una amplia variedad de conceptos de diseño de evaluación, uso de datos, relevancia política de los resultados e impacto consiguiente sobre la capacidad de evaluación (véase la tabla C.1). Cada caso incluye una explicación de las ventajas y desventajas relativas de cada evaluación.

Caso práctico C.1 Seguimiento de la evolución del Plan de Acción para la Erradicación de la Pobreza (PEAP) en Uganda¹

C.1.1. Introducción

En 1995, las autoridades de Uganda iniciaron la formulación del Plan de Acción para la Erradicación de la Pobreza (PEAP) para asegurarse de que la reducción de la pobreza fuese el eje de su estrategia general de crecimiento y desarrollo. Este plan fue desarrollado mediante un proceso de consultas en el que participaron representantes del Estado y de la sociedad civil, así como de organizaciones donantes. El objetivo último del PEAP es erradicar la pobreza masiva: reducir el porcentaje de la población que vivía en la pobreza absoluta del 56 % (1992) al 10 %, y reducir el porcentaje de personas que vivían en pobreza relativa desde más del 85 % al 30 % hacia el año 2017.

Se convinieron objetivos adicionales en cuatro áreas –macroeconomía, gestión pública, generación de renta y desarrollo humano– que se ampliaron en un conjunto de objetivos estratégicos (véase el cuadro C.1). La atención sanitaria primaria, la educación primaria, la extensión agrícola y las carreteras secundarias rurales fueron identificadas como áreas de pobreza de prioridad inicial para la asignación de recursos. La definición de metas y la selección de objetivos estratégicos y áreas prioritarias fueron procesos dinámicos, revisados periódicamente a la luz de información reciente, como por el ejemplo la procedente de la Uganda Participatory Poverty Assessment (UPPA, Evaluación Participativa de la Pobreza) realizada en 1998 y 1999, así como del sistema de seguimiento de la pobreza.

C.1.2. Sistema de seguimiento de la pobreza

La evolución en el cumplimiento de los objetivos se evalúa mediante un seguimiento continuo de la pobreza. Este proceso se inició como una actividad ad hoc y fue evolucionando gradualmente hacia un sistema de seguimiento descentralizado y participativo, con una clara definición de papeles y responsabilidades, incluidos mecanismos para la colaboración entre instituciones.

El sistema consta de tres instituciones básicas:

- La Oficina de Estadísticas (Uganda Bureau of Statistics, UBoS), que recoge, analiza y publica los datos procedentes de las encuestas a hogares.
- Los Departamentos de Estadísticas de los ministerios sectoriales, que recogen y analizan datos sectoriales procedentes de los sistemas de información para la gestión.
- La Poverty Monitoring Unit (Unidad de Seguimiento de la Pobreza), cuya principal función es relacionar a los productores de datos con los diseñadores de políticas. Recoge datos sobre la pobreza procedentes de diferentes fuentes, incluida la UBoS, los ministerios sectoriales y otras organizaciones e instituciones extragubernamentales; analiza los datos; divulga los resultados y debate las tendencias y perspectivas de la pobreza con representantes y organismos del Estado. En el futuro, la unidad se ampliará para incluir en el ámbito de su jurisdicción el análisis de políticas para la reducción de la pobreza. La unidad está integrada en el Ministerio de Finanzas, Planificación y Desarrollo Económico, organismo fundamental por su influencia política.

Además, hay otras instituciones –como las organizaciones no gubernamentales (ONG), centros académicos y de investigación, donantes– que juegan un papel importante (aunque todavía no sistemático) en la recogida y análisis de datos adicionales. También los diseñadores de políticas son parte importante del sistema en tanto que principales usuarios de los resultados (fundamentalmente en el nivel

central, aunque se ha llegado a la conclusión de que también pueden utilizarse datos estadísticos recogidos en el nivel local para la toma de decisiones locales).

Tabla C.1. Resumen de los casos prácticos de evaluación de impactos

Programa/ proyecto	País	Tipo de base de datos	Unidad de análisis	Parámetros de resultados	Método econométrico				Evaluación cualitativa	Ventajas
					Aleatorización	Compa- ración pareada	Compara- ción reflexiva	Variables instrumen- tales		
Educación										
Reforma para la autonomía escolar	Nicaragua	Encuestas de muestreo y evaluaciones cualitativas	Estudiantes, padres, docentes, directores	Notas de exámenes, grado de autonomía local en la toma de decisiones	No	Sí	Sí	No	Sí	Combinación cualitativa- cuantitativa
Intervención en el absentismo escolar	Filipinas	Valores de referencia y encuesta postinterven- ción	Estudiantes, aulas, docentes	Notas de exámenes y porcentaje de absentismo escolar	Sí	No	Sí	Sí	No	Análisis de costo- beneficio; generación de capacidades
Programas de trabajo										
Programa Trabajar	Argentina	Encuesta a hogares, censo, registros administrati- vos, evaluaciones sociales	Trabajadores , unidades familiares	Renta, canalización de recursos, costos	No	Sí	No	Sí	Sí	Uso prudente de fuentes de datos existentes, técnicas analíticas innovadoras
Agricultura										
Proyecto nacional de extensión agrícola	Kenia	Datos de muestreo, evaluaciones de beneficiarios	Unidades familiares, explotacione s agrícolas	Productivida d y eficiencia de las explotacione s agrícolas	No	No	Sí	No	No	Relevancia política de los resultados

Fuente: Adaptado de Baker, Judy. 2000. "Evaluación del impacto de los proyectos de desarrollo en la pobreza: Manual para profesionales". Directivas de desarrollo. Washington, D.C.: Banco Mundial.

El sistema está actualmente sujeto a una importante reforma, dirigida a:

- Incrementar la participación; es decir, promover una mayor implicación en las actividades de seguimiento en el nivel local y la colaboración entre la UboS, la Poverty Monitoring Unity, las ONG y los ministerios sectoriales para la recogida, análisis y divulgación de los datos. También se están reformando los vínculos entre los organismos de distrito y centrales que compaginan datos estadísticos.
- Desarrollar las capacidades, en especial la de seguimiento en el nivel local (distrito), y la de análisis y divulgación de datos en el nivel central, con el objeto de reducir el período entre la recogida y el análisis/divulgación de los datos.
- Definición de las funciones institucionales; es decir, establecimiento de funciones y responsabilidades claramente definidas, incluidos los mecanismos de colaboración.
- Armonizar la preparación de los informes de evolución; es decir, definición de un formato común para la generación de informes de progresos de los programas sectoriales y de lucha contra la pobreza.

Cuadro C.1. Objetivos del Plan de Acción para la Erradicación de la Pobreza (PEAP) de Uganda

Objetivo general: reducir el porcentaje de la población que vive en la pobreza absoluta desde el 56% al 10%, y reducir el porcentaje de la población que vive en pobreza relativa desde más del 85% al 30% hasta el 2017.

Meta 1. Implementación de políticas macroeconómicas que posibiliten un entorno favorable para la reducción de la pobreza

- Mantener un tipo de cambio estable que contribuya a la competitividad del sector exportador
- Mantener bajos niveles de inflación que faciliten la movilización del ahorro y la planificación a largo plazo
- Promover las inversiones del sector privado en áreas rurales
- Reducir la distorsión antiexportación de la política comercial para mejorar las perspectivas de las exportaciones
- Promover un crecimiento económico de amplia base
- Reducir la deuda externa hasta niveles sostenibles
- Reducir las disparidades de la pobreza entre los distritos
- Mejorar la participación económica y política de la mujer
- Ampliar la base impositiva
- Redireccionar el gasto público para vincularlo directamente a la erradicación de la pobreza

Meta 2: Creación de un marco institucional que promueva la reducción de la pobreza mediante una amplia participación, transparencia y obligación de rendir cuentas

- Reforzar la prestación eficaz y eficiente de servicios públicos al tiempo que se promueve la transparencia y obligación de rendir cuentas.
- Promover el crecimiento del sector privado incrementando las inversiones locales y extranjeras
- Reforzar los mecanismos de mantenimiento de la ley y el orden, y de administración de la justicia, y mejorar el acceso de los pobres a los servicios legales
- Reforzar el cumplimiento de los derechos humanos y la libertad, así como una buena gestión democrática
- Promover la participación de la comunidad en la planificación y prestación de servicios

Meta 3: Expansión de las oportunidades de ingresos de los pobres

- Crear una red de carreteras eficiente
- Transformar y modernizar la producción agrícola
- Garantizar la seguridad de la tenencia de la tierra, así como el acceso adecuado y el uso eficiente de la misma, sin perjudicar al medio ambiente
- Apoyar el desarrollo de mercados rurales: infraestructura, información y estándares de mercado
- Oferta de servicios financieros a los pobres mediante la promoción del crecimiento de instituciones de microfinanciación y de cajas rurales
- Incrementar la productividad laboral, dando prioridad al empleo de la mujer, a la reducción de la explotación infantil y a la protección de los derechos de los trabajadores
- Crear un entorno que posibilite el desarrollo de microempresas y pequeñas empresas

Meta 4: Mejorar la calidad de vida y el capital humano de los pobres

- Cumplir el mandato constitucional de atención sanitaria básica universal, mejorar la prestación de servicios sanitarios a toda la población sobre una base de racionalidad económica
- Proporcionar a toda la población fácil acceso al agua potable y mejorar la eficacia económica del suministro de agua
- Llegar a la educación primaria universal y mejorar la calidad de la educación
- Promover el acceso a la educación básica a niños vulnerables (por ejemplo, niños sin hogar)
- Promover la adquisición, uso y mantenimiento de la alfabetización funcional de todos los habitantes de Uganda

Una de las opciones consideradas para abordar algunos de estos temas es el establecimiento de una organización de campo para la UBoS. La organización de campo sería responsable de controlar el flujo de información hacia y desde la central; de apoyar el desarrollo de estadísticas en el nivel de distrito; de contratar, capacitar y supervisar al personal de campo; de programar el trabajo de campo; de recoger e introducir los datos; y de todas aquellas otras funciones asociadas con el trabajo de campo. Se establecerían seis regiones estadísticas. Cada región tendría una oficina zonal con un pequeño número de empleados permanentes (supervisor de zona, asistente estadístico y operador de introducción de datos), más supervisores y recogedores de datos que se contratarían temporalmente.

En octubre de 2001, el Ministerio de Finanzas, Planificación y Desarrollo Económico preparó Estrategia para el Seguimiento y Evaluación de la Pobreza, con el objeto de tratar algunos de estos temas y sus posibles soluciones.

Parámetros

La selección de parámetros ha sido un proceso iterativo. Originalmente, los parámetros se seleccionaron de acuerdo con el trabajo de grupos temáticos para realizar un seguimiento de los progresos en una serie de áreas: indigencia (pobreza de renta), estado de salud, educación, medio ambiente, infraestructura, gestión pública, empleo y acceso a la información, el mercado y el crédito. Esta primera lista de parámetros se percibió como demasiado larga, incompleta en algunas áreas, y no centrada en las prioridades. Como parte del Marco de Referencia del Gasto a Medio Plazo (MTEF, por sus iniciales en inglés), el Grupo de Trabajo sobre la Pobreza, compuesto por funcionarios públicos y representantes de la sociedad civil y de los donantes, redujeron la lista de parámetros (véase la tabla C.2). El contenido de esta lista será sometido a una segunda depuración para garantizar que sea coherente con el PEAP revisado. En este momento, prácticamente todos los parámetros son objeto de un seguimiento en el nivel nacional. Un subconjunto de los mismos es seguido en el nivel de distrito y/o regional. Los de educación son los únicos datos desglosados por género. Ello supone una importante limitación para un análisis completo de la pobreza, y se trata de un problema que se prevé que se solucionará en el futuro.

Aunque se han registrado algunos progresos en alinear los parámetros con las metas del PEAP, todavía hay áreas susceptibles de mejora. Varios indicadores están definidos en términos de número de casos. Las cifras reales son importantes, pero en muchos casos los porcentajes y coeficientes permiten que los parámetros resulten más útiles. Por ejemplo, la proporción de unidades de prestación de servicios de salud con medicamentos esenciales es un indicador más informativo que sólo el número de unidades. Otro problema es que algunos parámetros no son indicadores inequívocos de progreso. Es decir, no es posible determinar si la situación ha mejorado o no de acuerdo con dicho parámetro. Por ejemplo, un aumento en el gasto familiar en educación no es un parámetro inequívoco de mejora. Es posible que las unidades familiares estén gastando más en educación porque consumen más o porque tienen que pagar más, en cuyo caso estarían, de hecho, consumiendo menos. Por último, la lista no diferencia entre parámetros finales e intermedios, una diferenciación que resultaría útil para evaluar el progreso general.

Tabla C.2. Lista modificada de parámetros de seguimiento, Uganda

<i>Parámetros</i>	<i>Nivel previsto de desglose</i>
INDIGENCIA	
Porcentaje de la población por debajo de la línea de la pobreza	Nacional, regional, de distrito
Número de personas que viven en la pobreza absoluta	Nacional, regional
Participación de las familias en el gasto en alimentos	Nacional, regional
Porcentaje de la población que reside en viviendas con techo de paja	Nacional, regional
Coefficiente de dependencia	Nacional, regional, de distrito
Coefficiente de Gini	Nacional, rural/urbano
Consumo per cápita del 20 % más pobre	Nacional, regional, de distrito
PIB per cápita	Nacional
Coefficiente ahorro/PIB	Nacional
Ingresos per cápita y por distrito	De distrito
Seguridad y vulnerabilidad	
Porcentaje de hogares afectados por robos o conmoción civil	Nacional, regional
Número de personas desplazadas dentro del país	Nacional, regional
Número de civiles muertos como consecuencia de la insurgencia	Nacional, regional
Número de casos criminales denunciados	Nacional, regional

Parámetros	Nivel previsto de desglose
INDIGENCIA (continuación)	
Proporción de unidades familiares que han experimentado importantes vaivenes en la renta el año anterior	Nacional, regional
Proporción de refugiados y desplazados en la población del distrito	De distrito
Proporción de unidades familiares que sufren necesidades económicas que han vendido bienes	Nacional
Red de carreteras	
Kilómetros de carreteras abiertos	Nacional
Kilómetros de carreteras rehabilitados	Nacional
Porcentaje de distritos con más del 50 % de las carreteras en mal estado	Nacional, de distrito
Porcentaje de la superficie no cubierta por carreteras	Nacional, de distrito
Tierras	
Incidencia de la pobreza por propiedad y tenencia de la tierra	Nacional, de distrito
Agricultura	
Tasa de adopción de métodos agrícolas modernos	Nacional, de distrito
Tasas de rendimiento	Nacional, de distrito
Porcentaje de agricultores dedicados a cultivos de supervivencia	Nacional, de distrito
Mercados	
Disponibilidad de mercados por tipo	Nacional, de distrito
Acceso a los mercados	Nacional, de distrito
Volumen de bienes y servicios negociados en determinado mercado	Nacional, de distrito
Proporción de hogares en los que el precio de venta del principal producto agrícola es inferior al 50 % del precio en el mercado urbano	Nacional, de distrito
Productividad laboral y empleo	
Tasa de desempleo	Nacional, de distrito
Participación en la capacitación vocacional	Nacional, de distrito
Media de horas trabajadas por día	Nacional, de distrito
Crédito rural	
Crecimiento de la cartera de microcréditos	Nacional, de distrito
Porcentaje de la población con acceso a microcréditos	Nacional, de distrito
Crecimiento del ahorro	Nacional, de distrito
Gestión del crédito (uso efectivo)	Nacional, de distrito
Disponibilidad de servicios de microfinanciación	Nacional, rural/urbano
CALIDAD DE VIDA	
Parámetros de salud	
Incidencia de enfermedades	Nacional, de distrito
Cobertura de campañas de vacunación	Nacional, de distrito
Porcentaje de la población residente en un radio de 5 km hasta la unidad de servicios sanitarios más cercana	Nacional, de distrito
Gasto per cápita de las unidades familiares en salud	Nacional, de distrito
Número de unidades de prestación de servicios sanitarios con medicamentos esenciales	Nacional, de distrito
Número de distritos con más de 1.000 habitantes por cada profesional sanitario cualificado	Nacional, de distrito
Cobertura de la asistencia prenatal	Nacional, de distrito
Agua y saneamiento	
Porcentaje de la población residente dentro de un radio de 0,5 de distancia al agua potable, por región	Nacional, de distrito rural/urbano
Porcentaje de la población con letrinas adecuadas y en buen estado	Nacional, rural/urbano
Sistema seguro de vertido de residuos	Nacional
Parámetros de educación	
Coficiente neto de escolarización primaria	Nacional, de distrito, género
Proporción de alumnos de primaria que terminan más de 4 años de educación	Nacional, de distrito, género
Número de alumnos por docente cualificado	Nacional, de distrito, género

Parámetros	Nivel previsto de desglose
CALIDAD DE VIDA (continuación)	
Distancia a las escuelas	Nacional, de distrito, género
Número de alumnos por aula	Nacional, de distrito, género
Número de alumnos por libro de texto	Nacional, de distrito, género
Gasto per cápita de las unidades familiares en educación	Nacional, de distrito, género

Tabla C.2. Lista modificada de parámetros de seguimiento, Uganda (continuación)

Parámetros	Nivel previsto de desglose
MEDIO AMBIENTE	
Nivel de cumplimiento de las normas medioambientales	Nacional general
Medidas correctivas adoptadas por el organismo nacional de gestión medioambiental	
Porcentaje de la población que aplica métodos de uso de la tierra sostenibles	
Asignación presupuestaria a los programas medioambientales de parte de las administraciones locales	
Porcentaje de tierras registradas en el catastro en los distritos	
GESTIÓN PÚBLICA Y TRANSPARENCIA	
Nivel de concienciación de derechos entre la población	Nacional
Porcentaje de casos denunciados resueltos	Nacional
Número de personas encarceladas más allá del período especificado por la ley	Nacional
Número de casos retrasados en los tribunales	Nacional
Casos de corrupción denunciados a diferentes niveles	Nacional
Programas exitosos en la erradicación de la pobreza	Nacional, de distrito
Número de casos de corrupción/estafa y abuso de poder resultantes en condenas	Nacional

Recogida de datos

Entre las principales fuentes de datos necesarios para el seguimiento se incluyen las encuestas a hogares, los sistemas de información para la gestión y los estudios cualitativos.

Las **encuestas a hogares** son planificadas e implementadas de forma centralizada por la UBoS tras consultas o participación limitadas en el nivel de distrito. El papel de los distritos está siendo actualmente analizado, con el objetivo de generar una capacidad local y de promover el rápido acceso a la información específica de cada distrito para que éstos puedan utilizarla en la planificación, implementación y seguimiento de sus programas y políticas. La UBoS y el Ministerio de Administraciones Locales están trabajando en un sistema que posibilite hacer partícipes a la Unidades de Planificación de Distrito en la recogida de datos, con el objeto de asegurar de que en el nivel local se utilicen datos estadísticos e información cualitativa relevante en el seguimiento del desempeño. Dicho sistema podría complementar el sistema de recogida de datos en unidades familiares que se gestiona de forma centralizada. Las encuestas a hogares para el seguimiento de la pobreza incluyen:

Las **Encuestas Integradas de Hogares (IHS)**, que recogen datos sobre las características de la unidad familiar y de la vivienda; las rentas y gastos de la unidad familiar; los bienes, préstamos y ahorros; la producción agrícola; y el estado sanitario y nutricional de los niños. La IHS realizada en 1992 proporcionó la información de referencia sobre 10.000 hogares de todo el país. El cuestionario de la encuesta fue revisado sobre la base de las aportaciones de la Evaluación Participativa de la Pobreza de Uganda, y ahora incluye preguntas sobre temas tales como la seguridad de la unidad familiar. La encuesta revisada - la Encuesta Nacional de Hogares de Uganda- fue realizada en 2001/2002, y está prevista una nueva ronda para 2003/2004.

Las **encuestas de seguimiento**, que recogen información similar a la de las IHS, aunque con una muestra más pequeña (5.000 unidades familiares) y un cuestionario más corto (que incluye un módulo relativo al consumo). Se realizaron anualmente desde 1992-93 hasta 1997.

Las **Encuestas Demográficas y Sanitarias (DHS)** recogen información sobre la salud materno-infantil, vacunación, acceso al tratamiento sanitario, incidencia de las principales enfermedades, etc. Los datos de referencia se obtuvieron a partir de la DHS de 1995, y en el año 2001 se realizó una encuesta de seguimiento.

Además, el Censo de Población de 2002 proporcionará información actualizada acerca de la estructura demográfica de la población: edad, estado civil, etnia, religión, número de integrantes de la unidad familiar, coeficientes de dependencia, etc. Otras encuestas, como la Encuesta de Seguimiento del Gasto Público (véase el capítulo 6, "Gasto público") y la Encuesta Nacional de Prestación de Servicios, han proporcionado información muy útil, aunque aún no se han convertido en parte del sistema de seguimiento ordinario. La Encuesta Nacional de Prestación de Servicios recoge información acerca del uso de los servicios públicos y de la satisfacción con los mismos. Esta encuesta se realizó a modo experimental en 1996, y tres años después se aplicó en el nivel nacional (por el momento es responsable de la misma el Ministerio de Servicios Públicos, aunque en el futuro está previsto que pase al ámbito de la UBoS). Recientemente finalizó una encuesta sobre instalaciones sanitarias.²

Los **sistemas de información para la gestión (MIS)** recogen datos sectoriales sobre productos, acceso a servicios y, en medida más limitada, calidad de los mismos. Por ejemplo, en el sector de la salud los MIS recogen información sobre el número de instalaciones sanitarias por tipo (públicas o privadas); el número de camas; las instalaciones que ofrecen servicios esenciales; la dotación de recursos humanos; y las principales causas de las patologías. En el ámbito de la educación, el Ministerio de Educación y Deportes efectúa un censo anual para recoger información local de escolarización, número de docentes, materiales de enseñanza, instalaciones y finanzas.

Se han identificado una serie de problemas relacionados con el uso de los datos de los MIS. En primer lugar, la información es incompleta. En 1996, por ejemplo, el censo de la educación tuvo un índice de respuestas del 60 % entre las instituciones asistidas por el Estado, y del 30 % de parte de las escuelas privadas. En segundo lugar, los datos no son de fiar. En el sector de la educación, basarse en la información aportada por los directores es una cuestión problemática, ya que suele exagerarse el número de alumnos para obtener más subvenciones. Se han implementado comprobaciones aleatorias para descubrir datos de escolarización exagerados. En el sector de la salud, en muchas áreas (en especial las zonas rurales remotas) el instrumental de diagnóstico, la capacidad del personal y la infraestructura de comunicaciones son limitados, por lo que se produce una importante carencia de información sobre incidencia de enfermedades. En tercer lugar está la cuestión de puntualidad. La llegada de datos desde los distritos es lenta, lo que retrasa el análisis, compilación y divulgación de los datos desde el organismo central. Por ejemplo, los resúmenes estadísticos de la educación se publican un año y medio después de haber sido recogidos los datos. Por tal motivo, estos datos no se utilizan para la prestación y planificación de los servicios.

Uno de los motivos del mal desempeño de los MIS es su alto nivel de centralización. Los distritos están obligados a recoger información sin ser partícipes en el proceso de seguimiento. Por consiguiente, carecen de incentivos para asegurar una recogida puntual de datos confiables. Las iniciativas tendentes a corregir esta situación comprenden actividades en el nivel de distrito y en el nivel sectorial en los ministerios sectoriales centrales. Entre las actividades en el nivel de distrito se incluye la implementación del Sistema de Información de Recursos del Distrito (DRIS). El DRIS es la segunda fase de un primer intento de recoger datos sobre servicios sociales y sobre la infraestructura relevante en todos los distritos (el Estudio del Perfil de Dotación de Recursos del Distrito, o DREPS). Este sistema establece un vínculo directo entre los distritos y la UboS, y está centrado en un gran número de variables, como administración, prestación de servicios e infraestructura.

En cuanto a los **estudios cualitativos**, el Proyecto de Evaluación Participativa de la Pobreza de Uganda (UPPAP) es la principal fuente de datos cualitativos para el seguimiento del PEAP. Se trata de un proyecto trienal cuyo objeto es incorporar la percepción de los pobres al diálogo local y nacional para la reducción de la pobreza, así como comprender en mayor profundidad las tendencias emergentes de los datos cuantitativos. El trabajo de campo de la evaluación participativa de la pobreza más reciente se realizó a finales de 2001. El UPPAP es fruto de la colaboración entre el Estado, los donantes, las autoridades de los nueve distritos en los que se implementó el proyecto, y Oxfam, el organismo implementador.³ En el ámbito del Estado, el UPPAP es responsabilidad del Ministerio de Finanzas, Planificación y Desarrollo Económico, bajo la Unidad de Seguimiento y Análisis de la Pobreza.

La agricultura y la gestión pública son dos áreas en las que se requiere trabajo adicional. A pesar del alto porcentaje de pobres dedicados a actividades agrícolas, no existen muchos datos disponibles sobre este sector. El Programa de Modernización de la Agricultura, un elemento básico del tercer pilar del PEAP, no define parámetros. Por su parte, el DRIS no recoge datos de la agricultura. De manera similar, existe poca información útil sobre asuntos de gestión pública, a excepción de la Encuesta Nacional de Integridad y de los informes de la Comisión de Derechos Humanos. La Inspección General de Gestión Pública ha recogido gran cantidad de datos, pero debido a que esta organización carece de personal suficiente no se dispone de resúmenes de datos estadísticos. Los parámetros del PEAP en este terreno no han sido definidos.

Análisis y divulgación de datos, y opinión

El análisis de los datos es realizado fundamentalmente en el nivel central por la UBoS. En el nivel de distrito, el análisis de datos es limitado. Unos pocos distritos han puesto en marcha sus propios sistemas de seguimiento en el marco del Proyecto de Desarrollo de las Administraciones Locales (LGDP). Su principal objetivo es reforzar la planificación participativa y el desarrollo de los sistemas de elaboración de presupuestos y seguimiento en el nivel de distrito. Recientemente, el proyecto se ha ampliado desde los 5 a 9 distritos originales a otros 30.

Cada dos años, la Unidad de Seguimiento de la Pobreza presenta un Informe del Estado de la Pobreza para evaluar la evolución y las dificultades en la implementación del PEAP. Presenta un cuadro general de los progresos en pos a las metas del PEAP, y además refleja el estado de las iniciativas para la erradicación de la pobreza, incluidas las asignaciones presupuestarias. Esta información constituye la base para identificar lagunas, problemas y áreas prioritarias. El próximo informe está previsto para 2003.

Los informes son divulgados en el nivel nacional y de distrito, y se utilizan para la revisión del PEAP, el MTEF y las reseñas sectoriales. La Unidad de Seguimiento de la Pobreza y el Grupo de Trabajo sobre la Pobreza (PWG) son responsables de asegurar que los resultados del seguimiento se apliquen a los procesos de diseño de políticas y de asignación de presupuestos. Específicamente, el PWG, en el que colaboran funcionarios públicos y representantes de las ONG y del mundo académico, debe asegurar que los datos recogidos con el sistema de seguimiento de la pobreza sean tomados en consideración y utilizados por el grupo de trabajo del sector pertinente en los procesos del MTEF y del presupuesto. Asimismo, el PWG presenta recomendaciones sobre la asignación general de recursos presupuestarios para la reducción de la pobreza y sobre otras políticas presupuestarias que afectan a los pobres. A pesar de estas iniciativas, el desempeño del gasto público todavía se mide mayormente en términos de factores de producción y actividades más que de aportaciones a la reducción de la pobreza. La evolución hacia las metas aún desempeña un papel limitado en el proceso sectorial de asignación de presupuestos. Es necesario desarrollar un sistema de incentivos que vincule la asignación de recursos y la evaluación del desempeño a las aportaciones a los resultados del PEAP.

C.1.3 Generación de capacidad estadística

Las actividades del sistema de seguimiento de la pobreza están apoyadas por un importante programa que tiene por objeto modernizar los sistemas estadísticos de Uganda. La principal meta del programa es generar capacidad nacional para recoger, procesar, almacenar y divulgar información estadística que pueda utilizarse en el seguimiento y evaluación de la pobreza tanto en el nivel nacional como de distrito. El programa está centrado en el refuerzo de la capacidad del UBoS para elaborar un programa estadístico básico que permita el seguimiento periódico y puntual de los objetivos nacionales de desarrollo. El proyecto pretende crear una nueva infraestructura informática que permita integrar los sistemas de información para la gestión. Está diseñado para asegurar que todos los datos recogidos directamente por el UBoS o recibidos por este organismo como datos secundarios desde otras fuentes, queden centralizados en un formato común que facilite el libre acceso a los datos por parte de los usuarios, sea en forma impresa o electrónica. El Almacén Central de Datos guarda todos los datos en un formato listo para ser utilizado, lo que garantiza que todas las tablas y análisis estén basados en la misma fuente de datos. El sistema incorpora también datos macroeconómicos y tablas de resultados, susceptibles de ser utilizados para elaborar informes y boletines, o bien para su divulgación por medios electrónicos.

Otro terreno en el que se está haciendo hincapié es en el refuerzo de la capacidad de la UboS para administrar encuestas a hogares. Entre las principales actividades se incluye un programa trienal de Estrategia Integrada de Encuestas a Unidades Familiares) y la creación de un grupo de trabajo básico de equipos itinerantes. Estos equipos se utilizarán tanto para administrar las encuestas de la UboS como para centralizar el soporte técnico a los distritos que planifiquen sus propias encuestas. La UBoS realizará un estudio piloto de una encuesta de seguimiento de parámetros sencillos que podrá ser administrada por las autoridades de distritos para satisfacer sus necesidades de información.

El programa también apoya la administración repetida de una Encuesta Nacional de Prestación de Servicios anual. En esta encuesta se emplea un pequeño cuestionario y una amplia muestra (aproximadamente 20.000 hogares) que permita desglosar los resultados en el nivel de distrito. Incorpora una serie de características del Cuestionario sobre Indicadores Básicos del Bienestar (CWIQ), incluido el uso de lápiz óptico para agilizar el proceso de introducción de datos. Esta encuesta será gradualmente generalizada y quedará a cargo de la UboS. Las futuras rondas de esta encuesta se complementarán con entrevistas a grupos ad hoc.

Fuentes:

Hauge, Arild. 2001. "Strengthening Capacity for Monitoring and Evaluation in Uganda: A Results Based Management Perspective. Documento n° 8 de la Serie de documentos de trabajo sobre fortalecimiento de la capacidad de evaluación. Banco Mundial, Departamento de Evaluación de Operaciones, Washington, D.C.

Republic of Uganda. 1997. "Poverty Eradication Action Plan: A National Challenge for Uganda". Ministry of Finance, Planning, and Economic Development. Kampala.

———. 1999. "Uganda Poverty Status Report, 1999". Ministry of Finance, Planning, and Economic Development. Kampala.

———. 1999. "Five Year Strategy for Poverty Monitoring and Policy Analysis". Planning and Poverty Eradication Section, Ministry of Finance, Planning and Economic Development, Kampala

Banco Mundial. 1999. "Uganda's Integrated Information Management System: A New Approach in Statistical Capacity-Building". Africa Region Findings n° 142, Washington, D.C.⁴

Caso práctico C.2 Plan propuesto para el seguimiento de la Estrategia de Lucha contra la Pobreza en Tanzania

C.2.1 Introducción

La Estrategia de Lucha contra la Pobreza de Tanzania está basada en anteriores estrategias orientadas a solucionar la cuestión de la pobreza y promover el desarrollo humano. En ella están consolidadas una serie de anteriores estrategias a medio y largo plazo, como Visión 2025, la Estrategia Nacional para la Erradicación de la Pobreza de 1997 y la Estrategia de Ayuda a Tanzania. La ELP define un plan centrado en tres metas fundamentales:

- reducir la indigencia;
- mejorar la calidad de vida y el bienestar social; y
- alcanzar y mantener un entorno favorable al desarrollo.

La preparación de la ELP se ha caracterizado por una amplia participación de los interesados. Durante el proceso, y a través de seminarios regionales, se recogieron los puntos de vista de los principales implicados, incluidas las administraciones y comunidades locales, y la sociedad civil. También las metas tentativas, las prioridades y las iniciativas se discutieron en un seminario nacional en el que participaron funcionarios de las administraciones central y regionales, organizaciones del sector público, donantes y medios de comunicación.

C.2.2 Seguimiento de la Estrategia de Lucha contra la Pobreza

El Documento de Estrategia de Lucha contra la Pobreza (DELP) presentó un plan provisional para el seguimiento y evolución de la estrategia. Desde la publicación de la ELP, dicho plan ha sido depurado y continuará siéndolo a medida que, durante la implementación, se vayan aprendiendo nuevas lecciones. El presente caso práctico destaca tres aspectos de la experiencia de Tanzania: selección de parámetros, fuentes de datos y marco institucional previsto para el seguimiento y evaluación de la ELP. No se tratan aquí otras actividades, como estudios participativos, comunicación de los resultados y promoción de la participación.

Selección de parámetros

El sistema de seguimiento y evaluación (SyE) incluye una serie de parámetros finales e intermedios. Los parámetros finales fueron seleccionados en una lista más amplia de parámetros de la pobreza y el bienestar elaborada a través de un proceso de consultas. Estos parámetros se utilizarán para realizar un seguimiento de la evolución en pos de los principales objetivos de la estrategia. Los parámetros intermedios se emplearán para realizar un seguimiento de la implementación de la estrategia en lo que respecta a recursos asignados y a los bienes y servicios generados a través de las intervenciones políticas fundamentales. Dado que se reconoce la dificultad de medir determinados parámetros finales a intervalos frecuentes, el sistema de seguimiento incluye también una serie de parámetros representativos susceptibles de seguimiento anual. Por ejemplo, un objetivo de la ELP es reducir la indigencia. Por ello, es importante realizar un seguimiento periódico del porcentaje de población que vive por debajo de la línea de la pobreza. Sin embargo, en el caso de Tanzania (como ocurre en muchos otros países), la recogida de datos de renta o gasto a intervalos frecuentes no es viable, por lo que se decidió incluir los parámetros de propiedad de bienes domésticos y de materiales de construcción de las viviendas (que pueden observarse anualmente) como representativos de la indigencia.

Como puede verse en la tabla C.3, los parámetros propuestos pueden categorizarse en cuatro áreas coincidentes con los objetivos de la ELP: indigencia, calidad de vida y bienestar social (salud, educación, vulnerabilidad y bienestar social), estabilidad macroeconómica y gestión pública. La idoneidad de los parámetros en cuanto a relevancia, claridad, confiabilidad, puntualidad y combinación equilibrada entre parámetros finales e intermedios, varía enormemente entre las distintas áreas.

La ELP selecciona un parámetro final nacional específico de la indigencia, la incidencia de la indigencia medida sobre la base de la línea nacional de la pobreza, y una meta más ambiciosa que la de los Objetivos de desarrollo del milenio: reducir a la mitad la incidencia de la pobreza hasta 2010 en lugar de 2015. La incidencia de la pobreza se desglosará por áreas rurales y urbanas. Aunque sigue siendo en gran medida un fenómeno rural, la indigencia se está convirtiendo cada vez más en un problema urbano. Tal y como se ha mencionado, dado que la indigencia no se mide todos los años, se han identificado parámetros representativos. Se está desarrollando un modelo de encuesta capaz de recoger esta información anualmente. Los parámetros intermedios seleccionados son pertinentes al contexto de Tanzania y, por consiguiente, deben estar disponibles anualmente para el proceso de análisis de la ELP. Existe una buena combinación de parámetros intermedios y finales.

Tabla C.3. Parámetros propuestos para el seguimiento del DELP, Tanzania

<i>Objetivos</i>	<i>Parámetros finales</i>	<i>Parámetros intermedios</i>
1. Reducir la indigencia	<ul style="list-style-type: none"> • Incidencia de la pobreza • Propiedad de bienes domésticos (parámetro representativo) • Tipo de materiales de construcción de las viviendas, como por ejemplo suelos, paredes y techos (parámetro representativo) 	<ul style="list-style-type: none"> • Crecimiento real del PIB • Inversión (física y humana) • Productividad de la inversión • Crecimiento del valor agregado de la agricultura • Desarrollo de la estrategia del sector privado • Producción estacional de principales cultivo alimentarios y comerciales • Kilómetros de carreteras rurales rehabilitadas • Asignación real y presupuestada a carreteras rurales • Asignación real y presupuestada a extensión agrícola

2. Mejorar la calidad de vida y el bienestar social

- | | | |
|-------------------------------------|--|---|
| A. Salud, supervivencia y nutrición | <ul style="list-style-type: none"> • Índice de mortalidad de bebés y menores de 5 años • Porcentaje de niños menores de 2 años vacunados (sarampión, triple) • Índice de embarazadas seropositivas • Mortalidad materna • Esperanza de vida • Índice de mortalidad de menores de 5 años por malaria • Porcentaje de enfermedad/patologías • Porcentaje de hogares con acceso a agua potable • Incidencia del raquitismo • Incidencia de las enfermedades de deterioro progresivo | <ul style="list-style-type: none"> • Porcentaje de distritos con campañas activas de concienciación sobre el VIH/SIDA • Porcentaje de nacimientos asistidos por personal cualificado • Prácticas de amamantamiento • Implementación de programas de control de la malaria • Implementación del programa Administración Integrada de Enfermedades Infantiles • Asignación real y presupuestada para atención sanitaria primaria • Asignación real y presupuestada para VIH/SIDA • Asignación real y presupuestada para agua y saneamiento • Asignación real y presupuestada para educación básica |
| B. Educación | <ul style="list-style-type: none"> • Tasa de analfabetismo • Igualdad de géneros en la educación primaria y secundaria • Porcentaje de niños en edad escolar que terminan la educación primaria • Índice neto de escolarización primaria • Índice bruto de escolarización • Índice de absentismo escolar • Índice de transición de la educación primaria a la secundaria • Proporción de alumnos de 7º grado que aprueban con una nota específica en exámenes normales | |
| C. Vulnerabilidad | <ul style="list-style-type: none"> • Generación de capacidades en todas las comunidades que necesitan programas de redes de protección social | <ul style="list-style-type: none"> • Establecimiento de base de datos de grupos vulnerables • Producción de cultivos resistentes a la sequía en todas las áreas propensas a la sequía promovidas • Planes de riesgo comunitarios en todas las potenciales áreas de riego |
| D. Bienestar social | <ul style="list-style-type: none"> • Implementación plena de la Estrategia de Lucha contra la Pobreza | <ul style="list-style-type: none"> • Implementación plena del programa de reforma de la administración local • Coeficiente entre casos sentenciados y denunciados ante los tribunales • Tiempo medio que se tarda en resolver litigios comerciales • Relación entre sesiones reales y sesiones previstas de los tribunales de apelaciones • Número de seminarios de ELP celebrados, índices de participación y composición de los comités • Divulgación de informes |

Tabla C.3. Parámetros propuestos para el seguimiento del DELP, Tanzania (continuación)

<i>Objetivos</i>	<i>Parámetros finales</i>	<i>Parámetros intermedios</i>
3. Alcanzar y mantener un entorno favorable al desarrollo		
A. Estabilidad macroeconómica	<ul style="list-style-type: none"> • Tasa de inflación 	<ul style="list-style-type: none"> • Equilibrio presupuestario • Reservas internacionales oficiales brutas • Tipo de cambio • Balance por cuenta corriente
B. Gestión pública	<ul style="list-style-type: none"> • Número de votos presupuestarios administrados a través de los Sistemas Integrados de Información para la Gestión Financiera (IFM) • Compromisos de gastos y atrasos registrados a través de los IFM • Alcance y magnitud de la corrupción • Honestidad y transparencia del sistema contable • Sistema de gestión pública eficaz y efectivamente descentralizado • Refuerzo de la eficacia profesional y económica del sistema de gestión pública • Mejora de la capacidad, motivación y rendimiento de la gestión pública • Mejora de la administración presupuestaria en el nivel central e inferiores 	<ul style="list-style-type: none"> • Implementación de los IFM en todos los ministerios y subtesorerías • Desarrollo y aprobación de planes de acción específicos contra la corrupción en los ministerios de Agricultura y Cooperativas, Educación y Cultura, Sanidad, Aguas y CSD, basados en la Estrategia Nacional Anticorrupción • Desarrollo y aprobación de módulos de mejora de rendimiento para los sectores prioritarios • Preparación puntual de presupuestos en todos los niveles • Pluralismo institucional en la prestación de servicios públicos

Los parámetros de salud, supervivencia y nutrición reflejan el objetivo a largo plazo de aumentar la esperanza de vida a 52 años hasta 2010. No obstante, podrían definirse algunos indicadores intermedios que proporcionen más información. Por ejemplo, no están bien definidas las implementaciones del programa de control de la malaria y del programa Administración Integrada de Enfermedades Infantiles, salvo que hagan referencia a listas específicas de parámetros contenidos en otros documentos, como los planes de seguimiento de sectores o de programas. Si no hacen referencia a parámetros definidos en otro lado, deben definirse con mayor claridad. Por ejemplo, el parámetro de implementación del programa de control de la malaria podría convertirse en proporción de instalaciones de atención sanitaria primaria y secundaria con suministro periódico de medicamentos antipaludismo de primera y segunda línea. Del mismo modo, el porcentaje de instalaciones de atención sanitaria primaria y secundaria con personal cualificado en el marco del programa Administración Integrada de Enfermedades Infantiles (IMCI) es un parámetro mucho mejor que el si el programa IMCI se ha implementado o no. El factor tiempo también podría ser problemático en el caso de algunos de los parámetros finales y unos pocos intermedios (por ejemplo, las prácticas de amamantamiento), ya que la principal fuente de datos de estos parámetros será la Encuesta demográfica y sanitaria que se prevé que se realizará durante el período de implementación de la ELP.

Los parámetros y metas de la educación son relevantes para el objetivo de erradicar el analfabetismo hasta el año 2010: aumentar la escolarización primaria bruta al 85 %, aumentar el índice de transición de la educación primaria a la secundaria desde el 15 al 21 %; reducir el índice de deserción de la educación primaria desde el 6,6 al 3 %, y aumentar la escolarización primaria neta desde el 57 al 70 %. La mayoría de los parámetros finales están claramente definidos, a excepción del de igualdad de género, que podría medirse con respecto a los índices de escolarización primaria bruta, escolarización primaria neta, analfabetismo u otro. La lista de parámetros intermedios parece ser incompleta: los parámetros de resultados del gasto público, como el número de alumnos por docente, disponibilidad de libros de texto, porcentaje de aulas rehabilitadas y tiempo medio de viaje hasta la escuela, podrían facilitar información útil para comprender las tendencias de los parámetros finales. Es probable que los parámetros de la educación estén disponibles a intervalos frecuentes, ya que se obtienen a través de los sistemas habituales de recogida de datos del Ministerio de Educación. Los datos de escolarización se validarán con la información procedente del censo del año 2002.

Los parámetros de vulnerabilidad están menos desarrollados. Reflejan las iniciativas políticas que se implementarán en este sector, pero realizan un seguimiento de las actividades más que de los resultados. No se incluye ningún "parámetro final". También sería necesario desarrollar parámetros intermedios más específicos. Por ejemplo, el porcentaje de campesinos en regiones propensas a la sequía que se dedican a cultivos resistentes a la sequía podría ser un parámetro más adecuado que si se ha promovido o no la producción de cultivos resistentes a la sequía. Del mismo modo, un indicador del uso de la base de datos sobre grupos vulnerables podría ser más útil que el desarrollo (o no) de dicha base de datos.

Por su parte, los parámetros de bienestar social reflejan también la dificultad de especificar parámetros mensurables. Bajo este encabezado se aborda una multitud de problemas. Estos parámetros intentan reflejar la evolución en la transferencia de la responsabilidad de los servicios básicos a las autoridades locales; el acceso a la justicia, la eficacia y transparencia del sistema administrativo; y el nivel de participación de todos los implicados en el proceso de la ELP, pero proporcionarán sólo un cuadro muy parcial de los progresos en estos asuntos. Se trata de un área que requiere un desarrollo más profundo de objetivos, parámetros y metas.

Los parámetros macroeconómicos y de gestión pública intentan medir en qué medida se ha conseguido un entorno favorable al desarrollo. Específicamente, en el aspecto macroeconómico, la ELP pretende alcanzar una tasa de inflación más o menos acorde con la inflación prevista de los principales socios comerciales de Tanzania. Esto complementa el objetivo de alcanzar un crecimiento anual del PIB del 6 % en los próximos tres años. Ello constituiría la base para cumplir los objetivos de reducción de la pobreza a medio y largo plazo. Los parámetros macroeconómicos intermedios y finales que se proponen proporcionarán a intervalos frecuentes información relevante para el seguimiento de la evolución del objetivo de estabilidad.

En lo que respecta a la gestión pública, los principales objetivos son mejorar el desempeño del sector público, incluida la prestación de servicios públicos, minimizar la filtración de recursos y promover la rendición de cuentas. En general, los parámetros propuestos para el seguimiento de estos objetivos no están bien desarrollados. A diferencia de otras áreas, no todos los elementos que aparecen como "parámetros finales" son parámetros. Por ejemplo, "un sistema de gestión pública eficaz y efectivamente descentralizado" es un objetivo y no un parámetro de descentralización. Existen otros varios elementos que no son parámetros mensurables. Por ejemplo, el "alcance y magnitud de la corrupción" no especifica cómo va a medirse la corrupción. Sería necesario identificar parámetros pertinentes para Tanzania. El seguimiento de los problemas de género es parte integral del sistema de seguimiento de la ELP. Los parámetros de salud y educación, como tasa de mortalidad infantil y de menores de 5 años, porcentajes de vacunación, índices de escolarización e índices de transición de la educación primaria a la secundaria, deberían desglosarse por género. Además, el sistema de seguimiento incluye parámetros específicos de género, como índice de embarazadas seropositivas, mortalidad materna y porcentaje de partos asistidos por personal cualificado.

Los parámetros de salud y educación también se desglosarán por áreas urbanas y rurales, y por regiones administrativas. Ello reviste gran importancia, considerando la amplia variación geográfica en las condiciones sociales del país. Por ejemplo, la mortalidad infantil y de niños menores de 5 años es tres veces mayor en la región más pauperizada que en la menos pauperizada.

Una de las principales dificultades consiste en seleccionar un número manejable de parámetros que proporcionen información relevante y suficiente para evaluar la evolución de la ELP. En el caso de Tanzania, ha sido un proceso iterativo. El sistema de seguimiento empezó con aproximadamente 111 parámetros generales en el nivel nacional; en la actualidad, son unos 70. El proceso de depuración de la lista de parámetros continuará a medida que los funcionarios del Estado y sus colaboradores aprendan cuáles son los más útiles y cuáles faltan en la lista.

Tabla C.4. Fuentes de información para el seguimiento, Tanzania

Tipo de parámetro	Fuente de datos de referencia	Frecuencia de seguimiento y fuente de datos
Número de pobres	Estimaciones preliminares: Actualización de la Encuesta de Presupuestos de Unidades Familiares (HBS) 1991-92: HBS 2000-01	No están previstas HBS adicionales durante el período de implementación del DELP
Parámetros de renta representativos de la indigencia	HBS de 2000-01 Censo de 2002	Las encuestas anuales de seguimiento de la pobreza permitirán medir los parámetros representativos de la renta
Parámetros macroeconómicos	Cuentas nacionales y Encuesta Económica preparada por la Oficina Nacional de Estadísticas y la Comisión de Planificación	Actualizaciones anuales de las mismas fuentes
Infraestructura rural	Informes sectoriales de carreteras preparados por el Ministerio de Obras Públicas y el Ministerio de Administraciones Públicas Regionales y Locales	Misma fuente, frecuencia no especificada
Salud	Encuesta de Salud Reproductiva e Infantil de Tanzania de 1999 (DHS provisional)	Censo 2002 DHS (se prevé administrarla durante el período de implementación de este DELP) Sistema de Información Sanitaria (actualizaciones anuales de la cobertura de las campañas de vacunación)
Porcentaje de distritos con campañas activas de concienciación sobre el SIDA	Programa nacional de control del SIDA	Misma fuente, frecuencia no especificada
Estado nutricional de la infancia	Encuesta de Salud Reproductiva e Infantil de Tanzania de 1999 (nueva estimación a partir de la próxima DHS)	Seguimiento en el nivel comunitario y seguimiento de rutina en los centros de salud Las encuestas anuales de seguimiento de la pobreza podrían incluir también un módulo antropométrico
Parámetros de educación	Sistema habitual de recogida de datos del Ministerio de Educación Distribución geográfica de escuelas	Seguimiento anual empleando datos administrativos. El censo de 2002 permitirá una comprobación cruzada de los datos administrativos de escolarización
Asignación de recursos	Procesos PER, MTEF y presupuesto anual	Misma fuente; reuniones de revisión trimestrales

Fuentes de datos

En algunos casos resultó difícil calcular cifras de referencia fiables para los parámetros seleccionados, debido a la carencia de datos recientes. Los datos más recientes sobre el consumo utilizados para estimar la incidencia de la indigencia procedían de la Encuesta de Presupuestos de Unidades Familiares (HBS) de 1991-92. Se estimaron los datos de referencia del año 2000 y se definieron tentativas extrapolando los resultados de la encuesta de 1991-92 en función de las estimaciones de población procedente del listado de la HBS de 2000-01. No obstante, se observaron problemas metodológicos con estas estimaciones, que se revisarán de acuerdo con los resultados de la HBS de 2000-01.

Otras importantes fuentes de datos de referencias fueron la Encuesta de Salud Reproductiva e Infantil de Tanzania (TRCHS) de 1999, que proporcionó parámetros de salud y nutrición; los datos administrativos para los parámetros de educación; y las cuentas nacionales y Encuesta Económica para los parámetros macroeconómicos (véase la tabla C.4). Los datos del censo de 2002 y de la HBS de 2000-01 ayudarán a validar la confiabilidad de los datos administrativos sobre escolarización, y a actualizar las cifras de mortalidad del TRCHS de 1999. Se trata de una verificación importante, ya que la TRCHS es en realidad una DHS provisional que emplea una muestra más pequeña que la DHS. La mayoría de los parámetros serán objeto de un seguimiento anual, a excepción de los de incidencia de la pobreza y de algunos parámetros de salud y nutrición. Tal y como se ha mencionado anteriormente, el índice de pobreza sería sustituido por un conjunto de parámetros representativos de la renta, con datos de referencia calculados a partir de la HBS de 2000-01 y del censo de 2002, susceptibles de seguimiento a través de una encuesta anual de seguimiento de la pobreza. Se adoptó la decisión de no realizar una encuesta específica de seguimiento de la pobreza cada año, y en su lugar desarrollar un módulo especial que contenga los parámetros representativos de la renta, que será incluido en cualquier encuesta a hogares que se administre en determinado año. El seguimiento de los parámetros de salud y nutrición se efectuará al menos una vez durante el período trienal, en función de la disponibilidad de resultados de la próxima DHS.

Medidas institucionales previstas para el seguimiento y evaluación

El marco institucional propuesto para el seguimiento de la ELP es el resultado de un amplio proceso de consultas entre los diferentes interesados. En primer lugar, las propuestas se expusieron en una reunión de interesados, en la que participaron representantes del Estado, la sociedad civil, las ONG, el sector privado y las instituciones académicas y de investigación. Estas propuestas fueron debatidas en una segunda reunión en la que participaron funcionarios de organismos multilaterales y bilaterales, el Ministerio de Finanzas, la Comisión de Planificación y la Oficina Nacional de Estadísticas. La reunión fue organizada por la Oficina del Vicepresidente, en su papel de coordinador de la preparación del DELP. En noviembre de 2001 se acordó un Plan Maestro para el Seguimiento de la Pobreza.

El organismo fundamental previsto del sistema de SyE es el Comité Coordinador Nacional de Seguimiento de la Pobreza (NPMSC, por sus siglas en inglés). Su función será proporcionar orientación general sobre el seguimiento de la ELP y garantizar que el sondeo de opiniones del sistema de seguimiento quede incorporado al diseño de políticas nacional. El comité incluye representantes de la administración pública, del sector privado, de las ONG y de la sociedad civil. La División de Erradicación de la Pobreza (PED) de la Oficina del Vicepresidente actuará como secretaria.

Tal como puede verse en la figura C.1, en sus tareas el NPMSC recibe la colaboración de cuatro grupos de trabajo, coordinados por la PED.

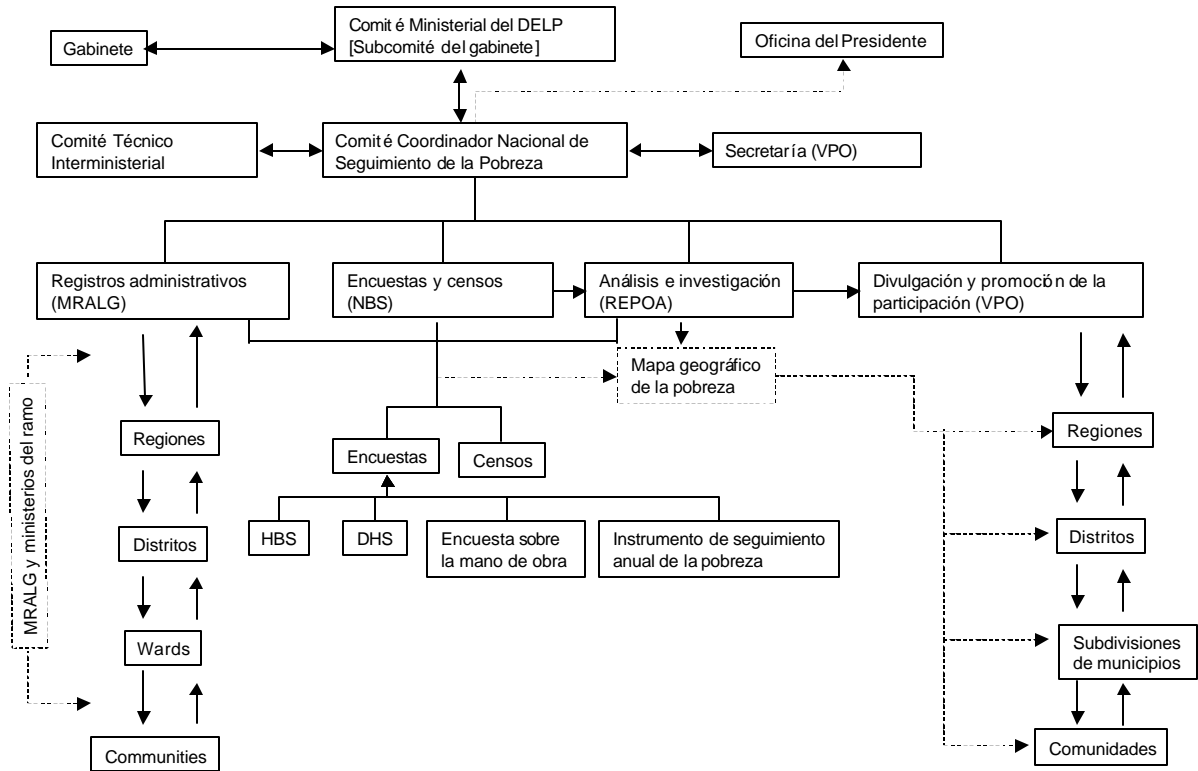
El Grupo de Trabajo de Encuestas y Censos, dirigido por la Oficina Nacional de Estadísticas, es responsable de realizar las grandes encuestas a hogares y del censo, así como de coordinar las actividades de almacenamiento de datos a través de la iniciativa Base de Datos Socioeconómicos.

El Grupo de Trabajo de Datos de Rutina, dirigido por el Ministerio de Administraciones Públicas Regionales y Locales (MRALG), es responsable de coordinar y administrar la recogida de datos sectoriales de los diversos ministerios, así como de los datos recogidos a través de los sistemas administrativos de las unidades oficiales descentralizadas.

El Grupo de Trabajo de Investigación y Análisis, dirigido por la Comisión de Planificación de la Presidencia y el Grupo de Investigación de Mitigación de la Pobreza (REPOA), son responsables de coordinar estudios e iniciativas especiales, como por ejemplo la elaboración de un mapa geográfico de la pobreza.

El Grupo de Trabajo de Divulgación, Sensibilización y Promoción de la Participación Pública, dirigido por la Oficina del Vicepresidente, está a cargo de la divulgación de actividades a todos los niveles, y de asegurar que los puntos de vista de las administraciones locales queden reflejados en el sistema de seguimiento.

Figura C.1. Marco institucional para el seguimiento del DELP, Tanzania



Está previsto que el Comité Coordinador Nacional de Seguimiento de la Pobreza desempeñe un papel fundamental como vínculo entre los diseñadores de políticas y el sistema de seguimiento, y sirva de enlace con el Comité Ministerial del DELP a través de la Oficina del Vicepresidente. Este comité, constituido por varios ministros y por el Gobernador del Banco de Tanzania, se formó con el objeto de orientar el proceso de preparación e implementación del DELP. Tiene el apoyo del Comité Técnico Interministerial, coordinado por el Ministerio de Finanzas, y está formado por funcionarios de las Oficinas del Presidente y del Vicepresidente, de la Comisión de Planificación, del Banco de Tanzania y de diversos ministerios sectoriales.

El marco institucional propuesto para el seguimiento y evaluación de la ELP podría ser un eslabón adecuado entre los productores y los usuarios de los datos, aunque también podría suponer una serie de problemas. En primer lugar, el MRALG y la Oficina Nacional de Estadísticas deben tener una sólida capacidad institucional para cumplir satisfactoriamente sus funciones de coordinación. En segundo lugar, el papel del Ministerio de Finanzas (MOF) como coordinador del seguimiento del gasto público no queda claramente reflejado en el marco actual. Sería importante establecer mecanismos de coordinación entre el MOF y el MRALG, que está a cargo de la coordinación de todos los datos administrativos para el seguimiento. En tercer lugar, es necesaria la plena implementación de la reforma de la administración pública local en curso para garantizar un adecuado flujo de datos administrativos desde los diferentes escalafones del Estado. Se espera que esta reforma aclare la división de responsabilidades de los sistemas de información para la gestión entre el MRALG y los ministerios sectoriales (Educación y Salud, por ejemplo), que por el momento no queda clara. Las demoras en la implementación de la reforma podrían conllevar la duplicación de esfuerzos o la pérdida de información. Por último, el marco institucional propuesto define una estructura bastante clara para el seguimiento y evaluación en el nivel nacional, aunque las medidas en el nivel regional y de distrito son menos claras. En la medida en que las iniciativas de descentralización devuelvan la capacidad de toma de decisiones a las administraciones públicas en el nivel local, es fundamental la implementación de una estructura de seguimiento y evaluación en ese nivel. En general, es importante que los incentivos para la recogida, análisis y difusión de información exacta y puntual sean homogéneos en todos los niveles del sistema de seguimiento.

Fuentes:

United Republic of Tanzania, The. 2000. "Poverty Reduction Strategy Paper". Dar Es Salaam
 _____. 2001. "Poverty Monitoring Master Plan." Dar Es Salaam.

Caso práctico C.3 Encuestas de opinión a ciudadanos como herramienta para la participación de la sociedad civil en la evaluación del desempeño del sector público: el caso de Bangalore, India

En Bangalore, India, una ONG ha realizado encuestas de opinión de los ciudadanos centradas en los servicios prestados por las autoridades municipales, como agua, electricidad, recogida de residuos y hospitales. En el marco de estas encuestas, se pregunta a los ciudadanos si están satisfechos con estos servicios públicos, cuáles son los aspectos más o menos satisfactorios, si los funcionarios son serviciales y si ha sido necesario sobornarlos para obtener estos servicios.

Los objetivos de la encuesta son:

- generar un flujo de opiniones ciudadanas sobre los servicios públicos y asignar a cada organismo municipal una puntuación general de su desempeño;
- identificar los servicios específicos que se prestan de manera adecuada o deficiente;
- identificar el alcance y profundidad de la corrupción;
- catalizar a los ciudadanos para que tengan más iniciativa;
- proporcionar una herramienta de diagnóstico a los departamentos municipales para que sus directivos puedan comprender mejor el desempeño de sus organismos e identificar aquellos aspectos de los servicios susceptibles de mejora; y
- fomentar y estimular a los organismos públicos para que sean más transparentes y orientados a los usuarios.

Las encuestas de Bangalore han clasificado todas los organismos administrativos municipales sobre la base del nivel de satisfacción de los ciudadanos con su prestación de servicios. Los hospitales y bancos recibieron las más altas puntuaciones. El organismo de desarrollo urbano (con el más alto nivel de corrupción), la más baja.

Los resultados de las encuestas han sido ampliamente divulgados y han sido objeto de una constante cobertura por parte de los medios de comunicación. Se han celebrado seminarios para facilitar los resultados a los grupos de ciudadanos y otras ONG. Aunque los resultados no suponían una novedad para ellos, proporcionaban pruebas fundamentadas y permitían señalar las áreas problemáticas. Además, los resultados han estimulado la participación ciudadana y la formación de grupos de residentes.

La ONG que realizó las encuestas presentó informes detallados a los directivos de todas las agencias oficiales encargadas de la prestación de servicios. La mayoría de los directivos recibieron con frialdad los resultados, aunque otros respondieron muy bien, como el director del organismo de desarrollo urbano, que a consecuencia de la encuesta promovió un acuerdo de colaboración con los grupos de ciudadanos y ONG. Ello llevó a innovaciones en la prestación de servicios y a la creación de un nuevo sistema para atender las reclamaciones de los usuarios. Con la ayuda de las ONG, se organizaron programas de capacitación de funcionarios y se estableció un grupo de colaboración para divulgar la información y actuar como vigilante.

Desde entonces se han realizado encuestas similares en otras ciudades de la India, como Madrás, Bombay, Calcuta y Pune. Ello ha permitido realizar publicar comparaciones de distintas ciudades.

Fuente: Adaptado de MacKay, Keith y Sulley Gariba, editores. 2000. *The Role of Civil Society in Assessing Public Sector Performance in Ghana: Proceedings of a Workshop*. Fortalecimiento de la Capacidad de Evaluación, Departamento de Evaluación de Operaciones, Banco Mundial, Washington, D.C.

Caso práctico C.4 Evaluación de los beneficios para los pobres de la asistencia a los trabajadores: el programa Trabajar de Argentina

C.4.1 Introducción

El programa Trabajar de Argentina tiene por objeto reducir la pobreza mediante la generación de oportunidades de empleo y, simultáneamente, la mejora de la infraestructura social de las comunidades pobres. El programa ofrece salarios relativamente bajos para atraer como participantes sólo a trabajadores pobres y desempleados, que se "autoseleccionarían" para participar. (Si desea información más detallada acerca de éste y de otros programas de obras públicas, consulte el capítulo 18, "Protección social"). Los proyectos de infraestructura en los que trabajan los participantes son propuestos por las administraciones locales y las ONG, que asumen los costos no salariales del proyecto. Los proyectos son aprobados en el nivel regional, siguiendo las directrices del gobierno central.

El programa ha evolucionado en el transcurso del tiempo, y ha incorporado las lecciones aprendidas. Trabajar I, un programa piloto, fue lanzado en 1996 como respuesta a la crisis económica y a índices de desempleo superiores al 17 %. Trabajar II se inició en 1997 como versión ampliada y corregida del programa piloto, en tanto que Trabajar III se inició en 1998. Trabajar II incluía una serie de reformas previstas para mejorar la canalización del proyecto hacia los beneficiarios previstos: los criterios de asignación presupuestaria del gobierno central dieron mayor peso a los parámetros provinciales de pobreza y desempleo, y a las propuestas de proyectos procedentes de las áreas pobres. En el nivel local se hicieron esfuerzos por reforzar la capacidad de los organismos municipales para ayudar a las áreas pobres a elaborar proyectos y elevar la calidad de las infraestructuras construidas en el marco del programa.

C.4.2. Diseño de la evaluación

La iniciativa de evaluación se inició durante los preparativos de Trabajar II. El objeto de la evaluación fue determinar si el programa cumplía o no sus objetivos, así como señalar en aquellas áreas en las que las reformas podrían incrementar su eficacia. La evaluación constó de una serie de componentes independientes que valoraron: a) los incrementos netos de renta devengados por los participantes en el programa; b) la asignación de recursos del programa entre las regiones (canalización); c) la calidad de los proyectos de infraestructura financiados. Además, el estudio analizó el papel de las comunidades y de las ONG en los resultados del proyecto.

Dos de los componentes de la evaluación reflejaron técnicas óptimas. El primer componente de la evaluación de Trabajar, la valoración de los incrementos netos de renta, supusieron una mejora con respecto a las evaluaciones convencionales de programas de obras públicas, que normalmente miden el incremento de la renta de los participantes simplemente como salarios *brutos* devengados, al estimar los incrementos *netos* de la renta. Basándose en datos de encuestas a hogares nuevos y ya existentes, y utilizando los adelantos más recientes en las técnicas de comparación pareada, el estudio tomó en cuenta los ingresos no percibidos (los ingresos que no devengaron los participantes por haberse incorporado al programa Trabajar), lo cual dio como resultado una estimación más exacta –e inferior– del incremento neto de la renta de los participantes.

El segundo componente, el estudio de los resultados de la canalización de recursos, introdujo una nueva técnica para evaluar la asignación de la financiación del programa cuando no se observa la incidencia del gasto público en el nivel local.

El diseño general de la evaluación también presentó una combinación óptima de componentes y técnicas de investigación –desde el análisis cuantitativo hasta la preparación de visitas de campo para la evaluación social– que proporcionaron un flujo integrado y puntual de resultados.

C.4.3. Técnicas de recogida y análisis de datos

La **evaluación del incremento neto de la renta** de los participantes del programa se basó en dos fuentes de datos: una encuesta nacional de niveles de vida (Encuesta de Desarrollo Social, o EDS) y una encuesta a los participantes de Trabajar realizada específicamente a efectos de evaluación.⁵ Estas encuestas fueron realizadas en agosto y septiembre de 1997 por la oficina nacional de estadísticas, utilizando el mismo cuestionario y los mismos equipos encuestadores. La encuesta para la EDS cubrió al 85 % de la población nacional, y se omitieron algunas áreas rurales y comunidades muy pequeñas. La muestra para la encuesta a los participantes de Trabajar se elaboró a partir de una muestra aleatoria de los proyectos de Trabajar II incluidos en la muestra de la EDS. Generó datos sobre 2.802 personas que en ese momento participaban en el programa (el total de participantes en Trabajar II entre mayo de 1997 y enero de 1998 alcanzó las 65.321 personas).

Para generar el grupo de control para la comparación pareada a partir de la EDS, el estudio empleó una técnica denominada índice de propensión. El emparejamiento ideal sería entre dos individuos, uno de la muestra de participantes y otro del grupo de control, para quienes las variables (x) de predicción de participación en el programa son idénticas. El problema básico en la comparación pareada es que resulta impracticable considerando el elevando número de variables contenidas en (x). No obstante, pueden calcularse las coincidencias en el índice de propensión de cada individuo, que es sencillamente la probabilidad de participar condicionada a (x).⁶ Los datos sobre la renta en la comparación entre el grupo de control de no participantes permitió estimar los ingresos no percibidos por los participantes en Trabajar II. A continuación, se calculó la renta neta obtenida por la participación en el programa restando los ingresos no percibidos del salario total del programa.

Sorprendentemente, el **análisis de canalización de recursos** no requirió ninguna recogida especial de datos. Se basó en los datos de la oficina ministerial del proyecto sobre asignaciones de financiación por departamento geográfico. Se basó asimismo en un índice de pobreza de cada departamento, calculado a partir del censo de 1991 como proporción de hogares con “Necesidades básicas insatisfechas”.⁷ Para analizar la incidencia de la canalización, se efectuó una regresión de los datos del gasto público por área geográfica –departamento, en este caso– sobre los índices de pobreza geográficos correspondientes. El coeficiente resultante estimó de manera homogénea un “diferencial de canalización” resultante de la diferencia entre el promedio de asignaciones del programa a los pobres y a los no pobres. Seguidamente, el diferencial nacional de canalización pudo descomponerse en diversos componentes debidos al mecanismo de canalización del gobierno central (financiación de asignaciones entre departamentos) y a la canalización realizada en el nivel provincial.

El **análisis de la calidad de las infraestructuras** fue un análisis de costo-beneficio de dos fases de los proyectos de infraestructura de Trabajar. En la primera fase, se tomó una muestra de 50 proyectos concluidos del programa Trabajar I, a los que se asignó una puntuación general de calidad basada en parámetros de seis categorías: técnica, institucional, medioambiental, socioeconómica, supervisión, operaciones y mantenimiento, y se realizó análisis de costo-beneficio donde procedía (no escuelas ni centros de salud). En la segunda fase, un año después, se efectuó un estudio de seguimiento de 120 proyectos de Trabajar II, que analizó el impacto de las reformas sobre la calidad de las infraestructuras.

Durante la preparación de los proyectos de Trabajar I y Trabajar II se realizaron **evaluaciones socioeconómicas**. Proporcionaron información sobre cuestiones de implementación de los proyectos, como el papel de las ONG, la disponibilidad de asistencia técnica en la preparación y construcción de proyectos, y la selección de beneficiarios. Fueron realizadas por sociólogos e incluyeron grupos ad hoc y entrevistas.

C.4.4. Resultados

Impacto del programa. Tomar en cuenta los ingresos no percibidos es importante para obtener un cuadro exacto de las ventajas de un programa de obras públicas. Los participantes en el programa no podían permitirse estar desempleados en ausencia del programa; en consecuencia, ciertos ingresos no se percibieron como consecuencia de la participación en el programa. Los ingresos no percibidos se calculan observando las rentas de los no participantes y “emparejándolos” con los de los participantes en el programa. Las estimaciones obtenidas mediante el método de comparación pareada muestran que obviar los ingresos no percibidos sobrestima enormemente la media de incremento de renta del programa, que se calcula que es aproximadamente la mitad de los salarios percibidos a través del programa. La evaluación también dejó en evidencia que la distribución de los beneficios fue decididamente favorable a los pobres, ya que el 80 % de los participantes en el programa se contaban entre el 20 % más pobre de la población, incluso después de reducir el incremento de la renta restándole los ingresos no percibidos. La participación femenina en el programa es baja (15 %), aunque el incremento neto de la renta es prácticamente idéntico para los hombres y mujeres participantes en Trabajar.

Desempeño de la canalización. El desempeño mejoró notablemente como resultado de las reformas de Trabajar II. Entre Trabajar I y Trabajar II, la asignación implícita de recursos a los hogares pobres se multiplicó por siete. Una tercera parte de esta evolución fue el resultado de una mejor canalización en el nivel central, y las dos terceras partes consecuencia de lo mismo en el nivel provincial. No obstante, se observaron significativas diferencias en los resultados de la canalización entre las provincias. Un departamento con un 40 % de sus habitantes clasificados como pobres podría esperar recibir entre cero hasta cinco veces la asignación media departamental, dependiendo de la provincia a la que perteneciese. Además, el desempeño de la canalización tendía a ser peor en las provincias más pobres.

Calidad de los proyectos de infraestructura. Se observó que la calidad era adecuada, pero lamentablemente las reformas de Trabajar II no conllevaron una mejora significativa. En parte, el motivo fue la brusca expansión del programa, que dificultó el cumplimiento de ciertas normas operativas que habían sido especificadas de antemano. No obstante, los proyectos de infraestructura de Trabajar II cumplieron mejor las necesidades prioritarias de la comunidad. La evaluación social reveló la necesidad de una mejor asistencia técnica a las ONG y municipios rurales, así como una mayor publicidad y transparencia de la información sobre el programa Trabajar.

C.4.5. Implicaciones políticas

Los resultados de la evaluación proporcionaron pruebas claras de que los participantes en el programa Trabajar provenían mayormente de entre los pobres. La autoselección de participantes conseguida ofreciendo salarios bajos es una estrategia que funciona en Argentina, y como resultado de la participación los participantes obtuvieron incrementos de la renta (aunque el beneficio neto es inferior al salario bruto debido a los ingresos no percibidos). Al parecer, el programa no discrimina a las participantes femeninas. Las reformas de Trabajar II han conseguido mejorar los resultados de la canalización geográfica. Ahora, el programa es más eficaz en la canalización de fondos a las áreas pobres. No obstante, el desempeño es variable y persistentemente deficiente en algunas provincias que merecerían mayor atención. Por último, los decepcionantes resultados en la calidad de los proyectos de infraestructura han generado iniciativas de mejora de los procedimientos operativos por parte del equipo del proyecto: insistir en más visitas *in situ* para la evaluación y supervisión, penalizar a las agencias que incumplan los requisitos del proyecto y reforzar el manual de evaluación.

C.4.6 Costos y administración de la evaluación

Costos. El costo de realizar la encuesta de Trabajar para el estudio del incremento de la renta y para el procesamiento de datos fue de aproximadamente 350.000 dólares. Las dos evaluaciones de costo-beneficio de la calidad de los proyectos costó unos 10.000 cada una, al igual que las evaluaciones sociales, con lo que se estima que el gasto total de la evaluación giró en torno a los 390.000 dólares.

Administración. La evaluación fue implementada conjuntamente por el personal del Banco Mundial y por el equipo del proyecto argentino. En sus diferentes fases, la tarea de evaluación requirió la coordinación con diversos organismos oficiales, incluida la agencia estadística, el Ministerio de Trabajo (oficinas de campo) y la división de análisis de políticas del Ministerio de Desarrollo Social.

C.4.7. Lecciones aprendidas y futuros trabajos

Importancia de contabilizar los ingresos no percibidos para evaluar los beneficios del programa de obras públicas. Los ingresos no percibidos representan aproximadamente la mitad de los salarios brutos devengados por los participantes en el programa de obras públicas de Argentina. El resultado sugiere que los métodos de evaluación convencionales (que utilizan sólo los salarios brutos) infravaloran sustancialmente los beneficios netos y, por consiguiente, también en qué medida los participantes serían pobres si el programa no existiese.

Utilidad de los métodos de comparación con índice de propensión. Los índices de propensión permiten comparaciones confiables entre una muestra de participantes y otra de no participantes (grupo de control).

Uso prudente de las fuentes de datos nacionales existentes. A menudo, las fuentes de datos existentes, como el censo nacional o las encuestas a hogares, pueden proporcionar valiosa información para las iniciativas de evaluación. El recurrir a las fuentes existentes reduce la necesidad de costosas recogidas de datos ad hoc. Las innovadoras técnicas de evaluación pueden compensar la carencia de datos, como lo refleja la evaluación de los resultados de la canalización de recursos geográfica de Trabajar.

Amplia variedad de componentes de evaluación. El diseño de la evaluación de Trabajar ejemplifica una combinación eficaz de herramientas y técnicas de evaluación. Los análisis de datos de encuestas, las visitas de campo y las evaluaciones sociales se utilizaron conjuntamente para generar una amplia variedad de resultados que proporcionan valiosa información sobre la eficacia del proyecto y que señalaron las áreas susceptibles de reforma.

Puntualidad de los resultados. Muchos de los componentes de la evaluación fueron explícitamente diseñados teniendo en cuenta el ciclo del proyecto, y sincronizados para generar resultados durante las fases de preparación del mismo, de modo que estos resultados pudiesen utilizarse para aportar información al futuro diseño. En la actualidad, numerosos componentes generan periódicamente datos, en un proceso continuo de seguimiento del proyecto.

Futuros trabajos. Están previstos tres estudios: la técnica de investigación de comparación pareada volverá a aplicarse para evaluar el impacto de la participación en el mercado de trabajo; se volverá a evaluar la calidad de los proyectos de infraestructura; y un componente de investigación cualitativa examinará las operaciones y procedimientos del programa.

Fuentes y bibliografía adicional:

Jalan, Jyotsna y Martin Ravallion. 1999. "Income Gains to the Poor from Workfare: Estimates for Argentina's Trabajar Program". Documento de trabajo de Investigación de Políticas nº 2149. Banco Mundial, Grupo de Investigación sobre Desarrollo Económico, Washington, D.C.

Ravallion, Martin. 1999. "Monitoring Targeting Performance When Decentralized Allocations to the Poor Are Unobserved". Documento de trabajo de Investigación de Políticas nº 2080. Banco Mundial, Grupo de Investigación sobre Desarrollo Económico, Washington, D.C.

Caso práctico C.5 Evaluación del proyecto de extensión agrícola de Kenia

C.5.1 Introducción

Descripción del proyecto. El primer Proyecto Nacional de Extensión (NEP I) de Kenia implantó en 1983 el sistema de capacitación y visitas (T&V) para la gestión de los servicios de extensión agrícola. El proyecto tenía el doble objetivo de desarrollar las instituciones y de prestar servicios de extensión a los campesinos, con el objeto de incrementar la productividad agrícola. En 1991 fue seguido por el NEP II, orientado a consolidar los beneficios obtenidos con NEP I mediante un incremento del contacto directo con los campesinos, la mejora de la relevancia de la información y las tecnologías de extensión, la capacitación del personal y de los campesinos y la promoción del desarrollo institucional.

Evaluación de impacto. El desempeño del sistema de extensión keniano ha sido controvertido, y es parte de un debate más amplio acerca de la rentabilidad de aplicar el método T&V a la extensión agraria. A pesar de la intensidad del debate y a las importantes inversiones realizadas, se han hecho muy pocos intentos rigurosos de medir el impacto del método T&V. En el caso de Kenia, el debate ha sido especialmente marcado debido a que un estudio previo del método T&V había estimado una rentabilidad muy alta, que no fue confirmada por resultados convincentemente visibles, entre los que se incluyó el rendimiento deficiente de la agricultura del país en años recientes.

El Departamento de Evaluación de Operaciones (DEO) del Banco Mundial realizó entre 1997 y 1999 una evaluación del sistema de extensión de Kenia. Utilizando un marco de gestión basado en resultados, la evaluación analiza el impacto de los servicios del proyecto en la productividad y eficiencia agraria. Asimismo, desarrolla indicadores de los productos (por ejemplo, frecuencia y calidad de los contactos) y resultados (la concienciación de los campesinos y la adopción de nuevas técnicas) del programa.

C.5.2 Diseño de la evaluación

La estrategia de la evaluación es un ejemplo de técnicas óptimas por el hecho de emplear una amplia variedad de métodos de evaluación y de aprovechar los datos existentes.⁸ Se basó tanto en métodos cuantitativos como cualitativos, de modo que las rigurosas observaciones empíricas sobre el impacto del programa pudieron complementarse con evaluaciones de beneficiarios y entrevistas del equipo que dejaron en evidencia problemas prácticos en el proceso de evaluación. Además, el estudio aplicó el método de valoración contingente para deducir la disposición de los campesinos a pagar por los servicios de extensión.⁹ La evaluación cuantitativa se vio complicada por el hecho de que el sistema T&V se introdujo a escala nacional, lo que impidió la comparación entre los "con programa" y los "sin programa" (grupo de control). Por consiguiente, la metodología de evaluación pretendía aprovechar los datos de producción agrícola familiar anteriores al proyecto con el objeto de realizar comparaciones limitadas de "antes y después" empleando métodos de datos de muestreo. Para ello, los datos existentes de unidades familiares se complementaron con una nueva encuesta para constituir una muestra. No se pudieron realizar las evaluaciones de beneficiarios previstas para este estudio, aunque la evaluación se basó en las observaciones relevantes de dos evaluaciones de beneficiarios recientes realizadas en Kenia.

C.5.3 Técnicas de recogida y análisis de datos

El método de evaluación se basó en diversas fuentes de datos cualitativos y cuantitativos existentes. La evaluación cuantitativa se basó en gran medida en una encuesta a hogares de 1998, administrada por el DOE. Esta encuesta generó los datos de muestreo mediante nuevas visitas a tantas unidades familiares encuestadas en 1990 como pudieron encontrarse (una encuesta a hogares realizada por el Departamento Técnico para África del Banco Mundial), que a su vez se basó en una submuestra de la Encuesta de Presupuestos de Unidades Familiares (HBS) Rurales de 1982.¹⁰

El estudio evaluó el desarrollo institucional basándose en una encuesta del personal de extensión, en diversas reseñas recientes del servicio de extensión realizadas o encargadas por el Ministerio de Agricultura, y en discusiones individuales y de grupos ad hoc con el personal de extensión.

Asimismo, el estudio empleó dos evaluaciones de beneficiarios recientes: un estudio de 1997 realizado por Actionaid de Kenia que recabó las opiniones de usuarios y potenciales usuarios de los servicios de extensión del país; y una Evaluación Participativa de la Pobreza de 1994, realizada conjuntamente por el Gobierno de Kenia, la African Medical and Research Foundation, la British Overseas Development Administration, UNICEF y el Banco Mundial, que sondeó la opinión acerca de los servicios públicos, incluidos los de extensión. La calidad y cantidad de los servicios prestados se evaluó empleando una combinación de las observaciones de las evaluaciones participativas (de beneficiarios) y de las encuestas del personal de extensión, así como con indicadores del alcance, naturaleza y frecuencia de los contactos entre los agentes de extensión y los campesinos obtenidos a partir de la encuesta del DEO de 1998. Se utilizaron métodos de valoración contingente para sondear directamente la disposición de los campesinos a pagar en concepto de los servicios de extensión.

Los datos de la encuesta también se utilizaron para evaluar los resultados del programa, medidos en términos de concienciación de los campesinos y adopción de las recomendaciones de extensión. Los resultados del programa —sus efectos reales sobre la producción agrícola de Kenia— se evaluaron correlacionando la prestación de servicios de extensión con los cambios en la productividad y eficiencia en el nivel de explotaciones agrícolas. Basados en los datos de muestreo de unidades familiares, estos impactos se calcularon empleando el Análisis por envoltura de datos, una técnica no paramétrica, con el objeto de medir los cambios en la eficiencia y productividad de los campesinos en el transcurso del tiempo, conjuntamente con un análisis econométrico que midió el impacto de la prestación de servicios de extensión sobre la producción agrícola.

C.5.4 Resultados

El desarrollo institucional de NEP I y NEP II ha sido limitado. Al cabo de 15 años, la efectividad de los servicios de extensión ha mejorado muy poco. Aunque recientemente se ha producido una sana reformulación de los métodos de extensión, en general el programa de extensión ha carecido de la visión estratégica necesaria para un futuro desarrollo. La gestión del sistema sigue siendo deficiente, y los sistemas de información prácticamente no existen. La calidad y cantidad de la prestación de servicios es deficiente. Tanto los beneficiarios como el personal de extensión manifiestan que las visitas son infrecuentes e ineficaces. Aunque sigue existiendo una demanda no satisfecha de servicios técnicamente útiles, el servicio público de extensión se ha mantenido centrado en sencillos y básicos mensajes agronómicos. No obstante, el método adoptado —contactos de alta intensidad con un número limitado de campesinos— es adecuado para proporcionar información más técnica. El resultado ha sido un sistema de prestación de servicio caro e ineficaz. El análisis demostró que las actividades de extensión apenas influían en la concienciación y en la adopción de recomendaciones, lo que indica un limitado potencial de impacto. En cuanto al impacto real sobre la producción y eficiencia agrícola, los datos indicaron un pequeño impacto positivo de los servicios de extensión en la capacidad de los campesinos de obtener el máximo de producción a partir de los recursos disponibles (eficiencia técnica). No obstante, no se observó ningún efecto sobre la eficiencia en la asignación de recursos (el uso de los recursos considerando los precios de mercado) ni sobre la eficiencia económica (la combinación de la eficiencia técnica con la eficiencia en la asignación de recursos). Asimismo, no pudo determinarse un impacto significativo de la prestación de servicios de extensión sobre la productividad en el nivel de explotación agrícola. Sin embargo, los datos dejaron en evidencia que el impacto fue relativamente mayor en las áreas previamente menos productivas, donde al parecer la falta de conocimientos era mayor. Estas observaciones coincidieron con los resultados de la valoración contingente: una vasta mayoría de campesinos, tanto beneficiarios como no beneficiarios, estaban dispuestos a pagar por el asesoramiento, lo cual indica una demanda insatisfecha. No obstante, el valor percibido del servicio, en términos de los importes ofrecidos, era muy inferior a lo que la administración invertía para su prestación.

C.5.5 Implicaciones políticas

La evaluación del Servicio de Extensión de Kenia es de destacar en lo que respecta a la variedad de conclusiones políticas prácticas que pueden extraerse a partir de sus resultados, muchos de los cuales son relevantes para el diseño de futuros proyectos de extensión agraria. En primer lugar, la evaluación dejó en evidencia la necesidad de mejorar la canalización de los servicios de extensión, al centrarse en áreas y grupos en los cuales es mayor la diferencia entre las prácticas medias y las óptimas, donde por consiguiente el impacto tiene mejores probabilidades de ser significativo. Además, el asesoramiento debe adaptarse cuidadosamente a la demanda de los campesinos, y tener en cuenta las variaciones en las condiciones tecnológicas y económicas locales. El alcanzar un alto nivel de adaptación requiere flujos permanentes y puntuales de información adecuada y confiable, así como un sistema de evaluación y seguimiento que aporte periódicamente la opinión de los beneficiarios sobre el contenido del servicio.

Para incrementar la eficiencia del programa lo más probable es que se requiera una presencia menos nutrida e intensa de agentes de extensión con una mayor cobertura. No hay suficientes innovaciones técnicas como para justificar una alta frecuencia de visitas, y existe una demanda no satisfecha de quienes actualmente no reciben los servicios. También el concepto generalista de prestación de servicios del programa, basado predominantemente en una única metodología (visitas a explotaciones agrícolas) para comunicar mensajes sencillos, limita la eficiencia del programa. Hoy son populares los programas de radio, los campesinos más jóvenes tienen más educación y están comenzando a aparecer prestadores alternativos de servicios (ONG) en la Kenia rural. Es muy probable que un concepto flexible y pluralista de la prestación de servicios, en especial alguno que emplee los medios de comunicación de menor coste, consiga mejorar la rentabilidad del programa.

Por último, las principales observaciones apuntaron a la necesidad de una reforma institucional. El objetivo central de la institución debe ser el campesino. La descentralización del diseño del programa, incluidos los mecanismos participativos que den voz al campesino (como las cooperativas agrarias), deben convertirse en parte integral del mecanismo de prestación del servicio. La sustentabilidad financiera es fundamental. La magnitud e intensidad del servicio debe basarse en las lagunas tecnológicas y de conocimientos existentes, así como en el ritmo de implantación de las nuevas tecnologías. La recuperación de costos, incluso parcial, supone varias ventajas: ofrece los incentivos adecuados, aborda los problemas de responsabilidades y control de calidad, hace que el servicio resulte más sensible a la demanda, y constituye un cierto alivio presupuestario. Este tipo de medidas de descentralización institucional aún no han sido ensayadas en Kenia, ni tampoco en muchos programas de extensión agraria de África y del resto del mundo.

C.5.6 Costos y administración de la evaluación

Costos. El costo total de la evaluación ascendió a aproximadamente 350.000 dólares, que cubrieron la recogida y procesamiento de los datos de las encuestas a hogares (65.000, aunque probablemente sea una cifra inferior al costo real); los datos de la encuesta realizados por el personal de extensión y el informe de consultoría (12.500 \$); otros costos de recogida de datos (12.500 \$). A ello deberían sumarse aproximadamente 160.000 \$ en concepto de tiempo y viajes del personal del Banco Mundial para el procesamiento y análisis de datos, así como para la preparación del informe, con el objeto de reflejar el costo real del estudio.

Administración. Para mantener la objetividad y dissociar el trabajo de la encuesta tanto del servicio de extensión oficial como del Banco Mundial, la encuesta a las unidades familiares fue implementada por el Tegemeo Institute de la Egerton University, un instituto de investigación independiente de Kenia. El análisis fue realizado por el personal del Banco Mundial.

C.5.7 Lecciones aprendidas

La combinación de una evaluación basada en teoría con un marco basado en resultados puede constituir una sólida base para evaluar el impacto de las intervenciones del proyecto, en especial cuando existen muchos factores susceptibles de afectar a los resultados previstos. El diseño de esta evaluación permitió medir los parámetros esenciales en fases críticas del ciclo del proyecto, al vincular los factores causales del mismo a los resultados previstos con el objeto de obtener suficientes pruebas del impacto.

Una evaluación empírica exige una supervisión constante e intensa. Una buena evaluación se puede simplificar significativamente con un sistema de seguimiento y evaluación eficaz y de alta calidad, especialmente con buenos datos de referencia. Muy pocas veces se dispone de los recursos adecuados para estas actividades. Esta evaluación también se benefició enormemente del hecho de tener acceso a algunos datos (aunque limitados) correspondientes a la época anterior al proyecto, así como a fuentes de datos independientes que se emplearon a efectos comparativos.

Es importante una validación cruzada de las conclusiones empleando diferentes métodos analíticos y fuentes de datos para obtener una evidencia creíble. Los problemas de datos e implementaciones imperfectas imponen límites al grado de confianza con que los métodos individuales pueden proporcionar respuestas a interrogantes fundamentales de la evaluación. Las evaluaciones cualitativas y cuantitativas se complementan perfectamente entre sí. La experiencia obtenida de esta evaluación indica que incluso en ausencia de evaluaciones participativas de los beneficiarios, es posible incluir preguntas adecuadamente formuladas en los cuestionarios para obtener tanto información cualitativa como cuantitativa. Esta información puede proporcionar datos útiles para complementar las evaluaciones cuantitativas.

Aplicada correctamente, la valoración contingente puede ser una herramienta muy útil, especialmente en la evaluación del valor del servicio público existente. Los resultados de la aplicación en esta evaluación son alentadores, y las respuestas parecen ser racionales y razonables.

Fuente: Gautam, Madhur. 1999. "World Bank Agricultural Extension Projects in Kenya: An Impact Evaluation". Informe nº 19253, Banco Mundial, Departamento de Evaluación de Operaciones, Washington, D.C.

Caso práctico C.6 Evaluación de la reforma escolar de Nicaragua: un método cuantitativo-cualitativo combinado

C.6.1 Introducción

Descripción del proyecto. En 1991, el gobierno nicaragüense implantó una amplia reforma de su sistema de educación pública: descentralizó la gestión de las escuelas (decisiones sobre personal, presupuestos, planes de estudios y pedagogía) y transfirió las responsabilidades de financiación al nivel local.

Las reformas se implantaron gradualmente, a partir de un decreto de 1991 que creó las asociaciones de padres en todas las escuelas públicas. Más tarde, en 1993, un programa piloto aplicado en 20 escuelas secundarias seleccionadas, transformó estas asociaciones en consejos de administración escolares, con mayor responsabilidad en los aspectos de personal, presupuestos, planes de estudios y pedagogía. En 1995 los consejos de administración escolares funcionaban en 100 escuelas secundarias y más de 300 primarias, que se incorporaron al programa a través de un proceso de autoselección que implicó una petición de docentes y directores.

Se esperaba que a medida que la gestión de las escuelas se hacía más democrática y participativa, con un incremento de los ingresos generados localmente, los patrones de gasto apoyasen mejor las iniciativas tendentes a mejorar la pedagogía y el rendimiento escolar.

Evaluación de impacto. La evaluación de las reformas de autonomía escolar de Nicaragua representa uno de los primeros esfuerzos sistemáticos de evaluar el impacto de la descentralización de las instituciones educativas en el rendimiento escolar. El diseño es innovador por el hecho de combinar tanto métodos de valoración cualitativos como cuantitativos. El componente cuantitativo es singular, ya que incluye un módulo separado para evaluar los procesos de toma de decisiones en la escuela. La evaluación también muestra las técnicas óptimas para los casos en los que no existen datos de referencia, y cuando la aplicación selectiva (no aleatoria) de las reformas descarga cualquier diseño de evaluación experimental.

El objeto del componente cualitativo de la evaluación fue determinar si las reformas previstas en la gestión y financiación se aplicaron efectivamente en las escuelas, así como sondear la opinión sobre el proceso de la reforma de los diversos implicados. El componente cuantitativo tenía por objeto responder al siguiente interrogante: los cambios en la gestión y financiación escolar, ¿generaron realmente un mejor rendimiento escolar de los alumnos? Los resultados cualitativos demostraron que la implementación satisfactoria de las reformas dependía en gran medida del contexto y el entorno de la escuela (es decir, el nivel de pobreza de la comunidad), en tanto que los resultados cuantitativos sugirieron que la mayor capacidad de toma de decisiones por parte de las escuelas estaba significativamente asociada a la mejora del rendimiento escolar.

C.6.2 Diseño de la evaluación

El diseño de la evaluación de la reforma de la autonomía escolar de Nicaragua está basado en una comparación pareada, en la cual los datos de una muestra representativa de las escuelas participantes en el proceso de reforma fueron comparados con datos procedentes de una muestra de escuelas que no participan. La muestra de escuelas no participantes se seleccionó de tal modo que se asemejara lo más posible a las características de las que sí participan, con lo que se proporciona el elemento contrafactual. Este diseño se eligió porque la ausencia de datos de referencia descartaba una evaluación de "antes y después", y porque las reformas no se aplicaron aleatoriamente a las escuelas, lo que descartaba un diseño de evaluación experimental.

C.6.3 Técnicas de recogida y análisis de datos

El estudio cualitativo se basó en datos procedentes de una muestra de 12 escuelas (9 que implementaron la reforma y 3 que no; estas últimas constituyeron el grupo de control).¹¹ La muestra de 12 escuelas se eligió de tal modo que estuviesen representadas tanto las primarias como las secundarias y tanto las rurales como las urbanas, así como escuelas con diferentes grados de autonomía real en la toma de decisiones. Se realizaron un total de 82 entrevistas y reuniones con grupos ad hoc, centrados en evaluar cómo entendían y consideraban el proceso de descentralización los directores, integrantes de los consejos, padres y docentes. Todas las entrevistas fueron realizadas por nicaragüenses, capacitados mediante simulación de entrevistas y pruebas piloto para utilizar una serie de preguntas guiadas sin respuestas sugeridas. Las entrevistas se grabaron en cinta, se transcribieron y se redujeron a una transcripción de dos a cuatro páginas, que posteriormente fueron analizadas para identificar conjuntos discretos de información y temáticas fundamentales que surgieron en todas las escuelas e implicados, y entre las escuelas participantes en la reforma y el grupo de control.

La recogida de datos cuantitativos constó de dos componentes: una encuesta de muestreo de escuelas, aplicada en dos rondas (noviembre-diciembre de 1995 y abril-agosto de 1997), y exámenes de rendimiento escolar de los alumnos de dichas escuelas, que tuvieron lugar en noviembre de 1996. La encuesta de las escuelas recogió datos sobre escolarización, índices de repetición y deserción, recursos físicos y humanos, toma de decisiones en la escuela, y características de los directores, docentes y estudiantes (y sus familiares) de la escuela. El módulo de toma de decisiones es único, y presenta una serie de 25 preguntas formuladas con el objeto de calibrar en qué medida, y cómo, la reforma realmente incrementó la capacidad de toma de decisiones por parte de las escuelas. La encuesta abarcó 116 escuelas secundarias (73 que adoptaron la reforma y 43 que no, que constituían el grupo de control), y 126 escuelas primarias (80 que adoptaron la reforma y 46 que no). También en este caso los grupos de control se seleccionaron para asemejarse a las características de las escuelas que adoptaron la reforma. Asimismo, la encuesta recogió datos sobre 400 docentes, 182 integrantes de los consejos y 3.000 estudiantes y sus padres, y se seleccionaron aleatoriamente entre 10 y 15 estudiantes en cada escuela.

El análisis de datos cuantitativos empleó técnicas de regresión para estimar una función de producción educativa. Analizó el impacto del grado de autonomía escolar sobre los niveles de rendimiento escolar, y se controlaron los factores causales escolares y las características del alumno y de su unidad familiar. El análisis midió el efecto de la autonomía tanto *de jure* como *de facto*. La definición de la autonomía *de jure* fue sencillamente si la escuela se había incorporado legalmente o no a la reforma, en tanto que la autonomía *de facto* midió el grado de autonomía real alcanzado por la escuela como porcentaje de las 25 decisiones más importantes adoptadas por la propia escuela. Se preveía que la autonomía *de facto* variaría entre las escuelas porque las reformas fueron implantadas gradualmente (de modo que las escuelas de la muestra se encontraban en distintas fases del proceso de reforma), y además porque la capacidad de implementar satisfactoriamente las reformas variaba en función del contexto de la escuela, tal y como se observó en el estudio cualitativo).

C.6.4 Resultados

El estudio cualitativo dejó en evidencia que los cambios políticos en el nivel central no siempre conllevaron cambios en el nivel local. En general, las reformas se asociaron a un incremento de la participación de los padres, así como a mejoras en la gestión y el liderazgo. Pero el grado de éxito con que fueron implementadas las reformas varió en función del contexto de cada escuela. De particular importancia fueron el grado de empobrecimiento de la comunidad circundante (en las comunidades pobres resultó difícil obtener financiación local para la escuela) y el grado de cohesión del personal escolar: allí donde los principales protagonistas, como los docentes, no se sintieron integrados en el proceso de reforma, la descentralización fue limitada. Los diseñadores de políticas suelen ignorar los contextos locales altamente variables en los que se implantan los nuevos programas. Los resultados de Nicaragua señalaron que la meta de incrementar la financiación local para las escuelas era de improbable cumplimiento en las comunidades pobres.

El estudio cuantitativo reforzó la observación de que las escuelas incorporadas a la reforma adoptaban efectivamente más de sus propias decisiones, en especial en lo relativo a cuestiones pedagógicas y de personal. La autonomía *de jure*, es decir, si la escuela había firmado formalmente el contrato de la reforma, no se tradujo necesariamente en un mayor nivel de toma de decisiones por la escuela, ni tampoco afectó a todas las escuelas de igual modo. El grado de autonomía alcanzado dependió del nivel de pobreza de la comunidad y de cuánto tiempo venía participando la escuela en el proceso de la reforma. Los resultados de la regresión demostraron que la autonomía *de jure* tuvo poca influencia sobre el rendimiento escolar, aunque la autonomía *de facto* estaba significativamente asociada a la mejora del rendimiento escolar. Además, las simulaciones indican que la mayor autonomía escolar tuvo un mayor efecto sobre el rendimiento escolar que otras intervenciones típicas, como aumentar el número de libro de textos, ampliar la capacitación docente y reducir el número de alumnos por aula.

C.6.5 Implicaciones políticas

Los resultados de la evaluación proporcionaron pruebas concretas de que la reforma escolar de Nicaragua generó resultados tangibles. De hecho, las escuelas integradas en la reforma adoptaron más decisiones localmente y la mejora del proceso de toma de decisiones local conllevó una mejora del rendimiento escolar.

Los resultados también señalaron áreas susceptibles de mejora, y el Ministerio de Educación decidió una serie de cambios en el programa de la reforma escolar. Ahora, el programa pone mayor énfasis en la capacitación de los docentes y en la promoción de los aspectos pedagógicos de la reforma. Asimismo, en respuesta a los problemas de financiación de las comunidades pobres, el ministerio desarrolló un mecanismo de subsidios basado en un mapa de la pobreza. Por último, las ventajas tangibles de esta evaluación llevaron a que el ministerio incorporase un componente de evaluación permanente al programa de la reforma.

C.6.6 Costos y administración de la evaluación

Costos. El costo total de la evaluación ascendió a aproximadamente 495.000 dólares, menos del 1,5 % del crédito otorgado por el Banco Mundial en apoyo de las reformas.¹² De este costo total de la evaluación, el 39 % se invirtió en el soporte técnico aportado por consultores independientes, el 35 % en la recogida de datos, el 18 % en tiempo del personal del Banco Mundial y el 8 % en desplazamientos.

Administración. La evaluación fue realizada conjuntamente por el Ministerio de Educación de Nicaragua, el Banco Mundial e investigadores de la Harvard School of Education.

C.6.7 Lecciones aprendidas

Valor del concepto de método combinado. La utilización de técnicas de investigación cualitativas y cuantitativas generó una valiosa combinación de útiles resultados relevantes para la política. El trabajo cuantitativo aportó un panorama amplio y estadísticamente representativo de las condiciones y resultados de las escuelas. Por su parte, el trabajo cualitativo reforzó estos resultados con información que explicó por qué se habían alcanzado algunos resultados previstos del programa de la reforma y por qué otros no, lo que se empleó como guía para ajustar la política. Además, al ser más intuitivo, el trabajo cualitativo fue más accesible y llamó la atención del personal del ministerio, lo que a su vez facilitó una rápida generación de capacidades y reforzó la credibilidad del proceso de evaluación dentro del ministerio.

Importancia de la generación de capacidades locales. La generación de capacidades locales fue un proceso costoso y requirió un contacto y coordinación permanentes entre el personal nicaragüense, sus colegas del Banco Mundial y los consultores externos. Sin embargo, la contrapartida fue el rápido desarrollo del control y la responsabilidad locales por el proceso de evaluación, lo que a su vez promovió un alto grado de aceptación de los resultados de la evaluación, fuesen positivos o negativos. Estos resultados de la evaluación proporcionaron información directa de la reforma a medida que evolucionaba. El impacto político de la evaluación también se vio reforzado por un cohesionado equipo local, en el que los evaluadores y diseñadores de política trabajaron en estrecha colaboración, y por el apoyo del Ministerio de Educación.

Fuentes y bibliografía adicional:

- 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 impacto de las reformas educacionales, documento n° 5. Banco Mundial, Grupo de Investigación sobre Desarrollo Económico, Washington, D.C.
- 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 impacto de las reformas educacionales, documento n° 9. Banco Mundial, Grupo de Investigación sobre Desarrollo Económico, Washington, D.C.
- King, Elizabeth, Berk Ozler y Laura Rawlings. 1999. "Nicaragua's School Autonomy Reform: Fact or Fiction?". Serie de documentos de trabajo sobre Evaluación de impacto de las reformas educacionales, documento n° 19. Banco Mundial, Grupo de Investigación sobre Desarrollo Económico, Washington, D.C.
- Nicaragua Reform Evaluation Team. 1996. "Nicaragua's School Autonomy Reform: A First Look". Documento de trabajo n° 1 de la serie de Evaluación de impacto de las reformas en la educación. Banco Mundial, División de Pobreza y Recursos Humanos, Departamento de Investigación de Política, Washington, D.C.
- . 1996. "1995 and 1997 Questionnaires, Nicaragua School Autonomy Reform". Serie de documentos de trabajo sobre Evaluación de impacto de las reformas educacionales, documento n° 7. Banco Mundial, Grupo de Investigación sobre Desarrollo Económico, Washington, D.C.
- Rawlings, Laura B. 2000. "Evaluating Nicaragua's School-Based Management Reform". En Michael Bamberger, editor, *Integrating Quantitative and Qualitative Methods in Development Research*. Washington, D.C.: Banco Mundial.

Caso práctico C.7 Rendimiento escolar en las escuelas elementales de Filipinas: evaluación del impacto de cuatro experimentos

C.7.1 Introducción

Descripción del proyecto. En la mayoría de los países en desarrollo, los altos índices de deserción escolar y el aprendizaje inadecuado de los alumnos de la educación primaria son un problema que preocupa a quienes diseñan las políticas. Tal era el caso de Filipinas: casi una cuarta parte de los niños filipinos abandonaban sus estudios antes de terminar el sexto grado, y quienes lo hacían a menudo dominaban menos de la mitad del material que se les había enseñado. En el bienio 1990-92, las autoridades implementaron el Plan de Intervención contra la Deserción (DIP, por sus siglas en inglés) con el objeto de encarar estos temas. Se realizaron cuatro experimentos: se proporcionaron materiales de aprendizaje multinivel (MLM) y almuerzos escolares (SL), y cada uno de estos planes se combinaron con una colaboración entre padres y docentes (PTP). Los materiales de aprendizaje multinivel permiten a los docentes adaptar el ritmo de enseñanza a las diferentes necesidades de los alumnos, y es mucho más económico que la alimentación escolar. La colaboración entre padres y docentes no cuesta prácticamente nada, pero puede ayudar al aprendizaje escolar tanto en casa como en la escuela.

Evaluación de impacto. La evaluación es de destacar por el hecho de que pretendía explícitamente generar capacidades en el país, de tal modo que se convirtiese en componente integral de nuevas iniciativas y de que los requisitos de datos se considerasen antes que después de la implementación del programa. Otra importante aportación de la evaluación fue comprobar la solidez de los resultados con diferentes métodos econométricos. Por último, el análisis costo-beneficio aplicado al final fue importante por el hecho de que reconocía explícitamente que unos resultados significativos no son suficientes como para justificar una intervención: las intervenciones económicas todavía son mejores que las caras.

C.7.2 Diseño de la evaluación

El principal objetivo de la investigación fue evaluar el impacto de cuatro distintas intervenciones sobre el índice de deserción y el rendimiento escolar. El diseño de la evaluación estuvo condicionado tanto por necesidades pragmáticas como programáticas. El equipo del DIP siguió un proceso de selección de escuelas que constó de tres fases:

- Se identificaron dos distritos de baja renta en cada una de las cinco regiones del país.
- En cada distrito, el equipo seleccionó tres escuelas que a) tuviesen todos los grados escolares, con un aula por grado; b) tuviesen un alto índice de deserción; y c) no tuviesen un programa de alimentación escolar.
- En un distrito, las opciones existentes eran MLM, MLM-PTP o nada; en el otro, SL, SL-PTP o nada.

Las tres escuelas de cada distrito fueron asignadas al grupo de control o a una de las dos intervenciones sobre una base aleatoria.

En todas las aulas de las 30 escuelas se administraron, antes y después de la intervención, exámenes de tres asignaturas (matemáticas, tagalo e inglés), al principio y al final de los años lectivos 1991 y 1992.

C.7.3 Técnicas de recogida y análisis de datos

La recogida de datos de referencia se inició en 1990-91, y las intervenciones se implementaron en 1991-92. Se recogió información detallada sobre 25 escuelas, unos 180 docentes y aproximadamente 4.000 alumnos en cada uno de los dos años lectivos. Aunque el cuestionario era muy detallada, muchos de los detalles resultaron ser innecesarios: se utilizó sólo un pequeño subconjunto de la información, lo que sugiere que las molestias que conllevó el proceso de evaluación podrían reducirse.

Los datos se estructuraron de tal modo que fuese posible realizar un seguimiento de alumnos y escuelas durante los dos años. Lamentablemente, los identificadores de los alumnos resultaron no ser únicos. Debe destacarse de que este hecho se desconocía *a priori*, y que quedó en evidencia sólo después de que, tras seis meses de trabajo, se observasen incoherencias internas. La recuperación de los identificadores originales del Departamento de Educación de Filipinas no fue posible. Felizmente fue posible rescatar los datos de los alumnos de primer grado, lo que permitió cierto análisis longitudinal.

La estructura del procedimiento de muestreo presentó ciertos problemas econométricos para los índices de deserción y los resultados de los exámenes. En cada caso existen dos conjuntos de controles obvios: uno es el grupo de control de las escuelas; el otro, la encuesta de datos de referencia realizada en el año que precedió a la intervención. Los autores manejaron estos controles de diferentes modos.

En el análisis de los índices de deserción, es natural utilizar un método de diferencia en la diferencia y comparar en cambio en el índice medio de intervención de cada curso sujeto a la intervención entre los dos años con el cambio en el índice medio de intervención de los cursos de control. No obstante, inmediatamente surgieron dos problemas. En primer lugar, los resultados –aunque bastante voluminosos– sólo fueron significativos para la intervención MLM, posiblemente el resultado de muestras de pequeño tamaño. Esto no es extraño en este tipo de procedimiento, en parte se debe a la falta de fondos para experimentos a gran escala en el contexto de un país en desarrollo. En segundo lugar, una rápida comprobación de si las características y resultados de los alumnos eran efectivamente las mismas en las escuelas en el año que precedió a las intervenciones, sugirió que existían algunas diferencias significativas en las características. Estos dos factores llevaron a los autores a comprobar la solidez de los resultados mediante técnicas de regresión logística que controlaban las características personales y los antecedentes familiares. El resultado fundamental no experimentó cambios. No obstante, la técnica de regresión reveló un importante factor causal indirecto de la deserción, que era el bajo rendimiento escolar. Naturalmente, esto llevó al segundo conjunto de análisis, centrado en el rendimiento.

En la evaluación del impacto de la intervención sobre rendimiento escolar de un alumno, en determinada escuela y en determinado momento, se presentaron varios problemas econométricos: tomar en cuenta la correlación de la agrupación de errores que pudiera existir para los estudiantes de las mismas aulas y escuelas; reflejar la heterogeneidad no observada; y solucionar la distorsión de la selección (si los estudiantes con menor rendimiento escolar tienen más probabilidades de abandonar los estudios, entonces las estimaciones de los efectos del programa podrían distorsionarse al alza).

Para solucionar el primer problema se optó por aplicar una corrección de Huber-White a los errores típicos. El segundo problema podría haberse solucionado, en principio, en el nivel individual utilizando la diferencia en las notas de los exámenes como variable independiente. No obstante, los autores arguyeron que ello sería inadecuado, ya que presuponía que el valor del coeficiente de rendimiento escolar era de 1, lo cual no estaba validado por los exámenes. Por consiguiente, utilizaron el rendimiento escolar del primer período como variable explicativa, aunque esto provocó el problema de la correlación con el término de error, o distorsión de regresor endógeno. Esto se solucionó utilizando como variable instrumental de la nota del examen en cada asignatura previo a la intervención las notas de los exámenes en las demás asignaturas (también previos a la intervención). No obstante, los autores destacaron que la reducción de la distorsión se consiguió a expensas de reducir la eficiencia, y por consiguiente expusieron los resultados tanto de los mínimos cuadrados como de las variables instrumentales. Los autores utilizaron los efectos fijos de escuela y docente para controlar la heterogeneidad no observada en el entorno de aprendizaje y las condiciones del aula.

El tercer problema, la distorsión de la selección, se ha tratado abundantemente en la literatura y no existe para el mismo una solución plenamente aceptada. Obviamente, el rendimiento escolar individual está condicionado a la decisión de no abandonar los estudios. Aunque a menudo este problema ha sido abordado empleando el procedimiento de dos fases de Heckman (Heckman 1976, 1979), resulta ser bastante insatisfactorio por tres motivos: su insensibilidad a la hipótesis de la distribución normal, la selección e idoneidad de las variables adecuadas a utilizar en la primera fase, y su frecuente apoyo en la identificación a través de la no linealidad de la primera fase. Lamentablemente, aún no existe consenso sobre una alternativa adecuada. Una solución propuesta por Krueger (1997) consiste en imputar las notas de los exámenes de los estudiantes que abandonan la muestra asignándolos a los años en los que estuvieron ausentes de la muestra su percentil de exámenes más reciente. Los autores comunicaron tres conjuntos de resultados: la regresión simple de los resultados comparados con la intervención, el método de Krueger, el procedimiento y de Heckman y los resultados de las variables instrumentales obtenidos mediante el procedimiento de Krueger.

C.7.4 Resultados

El estudio reveló que el efecto positivo de los materiales de aprendizaje multinivel –en especial existiendo una colaboración entre padres y docentes– sobre los índices de deserción y rendimiento escolar era sólido con respecto a diferentes especificaciones. El efecto de los almuerzos escolares fue, en general, débil. Un interesante componente del estudio fue el análisis de costo-beneficio, que fundamentó el argumento de que los resultados significativos no reflejan totalmente la realidad. Un cálculo directo tanto de los costos directos como de los indirectos (oportunidad) de los programas llevó a la conclusión de que el método MLM era a la vez efectivo y rentable.

Sin embargo, es probable que se sobrestime la falta de eficacia de los almuerzos escolares: es posible que un método mejor canalizado de programas de alimentación escolar –dirigido, por ejemplo, sólo a los alumnos desnutridos o desfavorecidos– resultase ser más rentable. Además, dado que el período de tiempo entre la implementación y la evaluación del programa fue bastante breve, la evaluación no pudo abordar el impacto a largo plazo de las intervenciones.

C.7.5 Administración

La recogida de datos estuvo a cargo de la Oficina de Educación Elemental del Departamento de Educación, Cultura y Deportes de Filipinas. El análisis fue realizado por un empleado del Banco Mundial y dos investigadores académicos.

C.7.6 Lecciones aprendidas

Esta evaluación permitió aprender varias lecciones. Una de las más importantes es que puede perderse un gran volumen de vital información longitudinal si los identificadores de las observaciones (en este caso, los alumnos) no son únicos en el transcurso del tiempo. Una segunda lección es que se utilizó muy poca de la información recogida en las encuestas detalladas, y de que se podrían haber reducido sustancialmente las molestias a los encuestados. En tercer lugar, el estudio destacó el valor de los diferentes métodos econométricos y las ventajas de encontrar resultados homogéneos empleando distintas técnicas. En cuarto lugar, el estudio fue ejemplar por el hecho de haber identificado y valorado el costo de las diferentes intervenciones. Y por último, aunque durante el estudio se cometieron errores, los autores destacaron que el objetivo principal fue generar capacidad de evaluación en Filipinas.

Fuentes y bibliografía adicional:

- Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models". *Annals of Economic and Social Measurement* 5:475-492.
- Heckman, James J. 1979. "Sample Selection as a Specification Error". *Econometrica* 47:153-161.
- Krueger, Alan. 1997. "Experimental Estimates of Educational Production Function". NBER Working Paper 6051, Cambridge, MA
- 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* 13(3):493-508.

Notas

1. Una primera versión de este caso práctico fue preparada por Margaret Kakande, Head, Poverty Monitoring Unit, Ministry of Finance, Planning and Economic Development, Government of Uganda, y Kimberley McClean, Head International Projects, Aus Health International.
2. Para obtener más información acerca de la Encuesta Nacional de Prestación de Servicios, véase el sitio Web del Departamento de Investigación del Banco Mundial sobre prestación de servicios públicos en [http://www.worldbank.org/research/projects/publicspending/tools/tools.htm#Quantitative Service Delivery](http://www.worldbank.org/research/projects/publicspending/tools/tools.htm#Quantitative%20Service%20Delivery).
3. Véase <http://www.uppap.or.ug/>.
4. Puede consultarse en <http://www.worldbank.org/afr/findings/english/find142.htm>.
5. La EDS fue financiada por un proyecto apoyado por el Banco Mundial distinto de Trabajar. Fue diseñada para mejorar la calidad de la información sobre el bienestar de las familias de Argentina, en especial en el área de acceso a los servicios sociales y a los programas sociales del Estado.
6. El cuestionario de la EDS es muy exhaustivo y recoge datos detallados sobre las características de la unidad familiar que ayudan a predecir la participación en el programa y facilitan el uso de la técnica de índice de propensión. El índice de propensión se calcula para cada observación en las muestras de participantes y del grupo de control utilizando modelos de unidades logarítmicas (logit) estándar.
7. Se trata de un índice compuesto que incluye número de residentes por metro cuadrado en las viviendas, instalaciones de saneamiento, calidad de la vivienda, logros educativos de los adultos, escolarización de niños, empleo y dependencia (coeficiente entre integrantes de la familia que trabajan y que no trabajan). Aunque este índice era en cierto modo obsoleto, tenía la ventaja de que el parámetro de pobreza departamental no estaba influido por las intervenciones del programa Trabajar.
8. No se realizó ningún intento de estudiar el impacto sobre el bienestar de las familias, susceptible de haberse visto afectado por una serie de factores que van mucho más allá del ámbito de las actividades de T&V.
9. El método de valoración contingente encuesta a los individuos sobre la valoración del uso y del no uso de bienes y servicios públicos y privados. Se pide a los entrevistados que manifiesten su disposición a pagar para evitar un cambio hipotético en la prestación de bienes o servicios; es decir, el resultado "contingente". En este caso, se preguntó a los campesinos cuánto estarían dispuestos a pagar para obtener servicios continuados de extensión agrícola en caso de que el Gobierno dejase de prestárselos.
10. Estas tres encuestas generaron un conjunto de datos de muestreo de aproximadamente 300 unidades familiares. Las encuestas incluían la composición demográfica de las familias, las características de las explotaciones agrícolas y los datos de insumo-producto de la producción agrícola. Las encuestas de 1990 y 1998 también recogieron información sobre el contacto con los servicios de extensión, incluidos el aprendizaje y adopción de los mensajes de extensión.
11. En realidad se recogieron datos de 18 escuelas, pero sólo 12 de éstas se incluyeron en el análisis cualitativo debido a los retrasos en la obtención de las transcripciones preparadas. Por ello, se adoptó la decisión de concentrar el grueso del análisis en las escuelas que implantaron la reforma, lo que proporcionó material más relevantes para el análisis.
12. Este total no incluye el costo de los equipos del Ministerio de Educación de Nicaragua.