

El análisis experimental de la ayuda al desarrollo

La evaluación de lo que funciona y lo que no funciona

José García Montalvo (Ed.)

**EL ANÁLISIS EXPERIMENTAL
DE LA AYUDA AL DESARROLLO**

EL ANÁLISIS EXPERIMENTAL
DE LA AYUDA AL DESARROLLO
LA EVALUACIÓN DE LO QUE FUNCIONA
Y LO QUE NO FUNCIONA

Edición a cargo de:
José García Montalvo

Fundación **BBVA**

La decisión de la Fundación BBVA de publicar el presente libro no implica responsabilidad alguna sobre su contenido ni sobre la inclusión, dentro de esta obra, de documentos o información complementaria facilitada por los autores.

No se permite la reproducción total o parcial de esta publicación, incluido el diseño de la cubierta, ni su incorporación a un sistema informático, ni su transmisión por cualquier forma o medio, sea electrónico, mecánico, reprográfico, fotoquímico, óptico, de grabación u otro sin permiso previo y por escrito del titular del *copyright*.

© los autores, 2008
© de esta edición / *of this edition*: Fundación BBVA, 2008

EDITA / PUBLISHED BY:
Fundación BBVA
Plaza de San Nicolás, 4. 48005 Bilbao

ISBN: 978-84-96515-79-6
DEPÓSITO LEGAL / LEGAL DEPOSIT NO.: M-49.530-2008

EDICIÓN Y PRODUCCIÓN: Atlántida Grupo Editor
COMPOSICIÓN Y MAQUETACIÓN: Márvel, S. L.
IMPRESIÓN Y ENCUADERNACIÓN: Rógar, S. A.

Impreso en España - *Printed in Spain*

Los libros editados por la Fundación BBVA están elaborados con papel 100% reciclado, fabricado a partir de fibras celulósicas recuperadas (papel usado) y no de celulosa virgen, cumpliendo los estándares medioambientales exigidos por la actual legislación.

El proceso de producción de este papel se ha realizado conforme a las regulaciones y leyes medioambientales europeas y ha merecido los distintivos Nordic Swan y Ángel Azul.

Índice

Agradecimientos	9
Introducción, <i>José García Montalvo</i>	11
1. La ayuda al desarrollo: su eficacia y métodos experimentales para su evaluación <i>José García Montalvo</i>	17
2. Cambios en los hábitos y los sistemas sanitarios: algunas evidencias a partir de evaluaciones aleatorizadas <i>Michael Kremer</i>	49
3. La evaluación de las intervenciones educativas: evidencia a partir de experimentos aleatorizados <i>Esther Duflo</i>	75
4. La evaluación de las intervenciones financieras en países en vías de desarrollo: ¿pueden los bancos rurales reducir la pobreza? <i>Rohini Pande</i>	103

5. El valor de la evaluación interdisciplinar: el análisis de programas de desarrollo basados en la comunidad <i>Vijayendra Rao</i>	129
Bibliografía adicional	146
Índice de cuadros	159
Índice de gráficos	161
Índice alfabético	163
Nota sobre los autores	167

Agradecimientos

El presente volumen tiene como origen las ponencias presentadas en la jornada sobre Análisis Científico de la Ayuda al Desarrollo que, patrocinada por la Fundación BBVA, tuvo lugar en su sede de Madrid el 22 de enero de 2007. Por lo tanto, en primer lugar desearía agradecer la confianza que el director de la Fundación, Rafael Pardo, tuvo en el proyecto desde el primer momento. En segundo lugar, agradecer también la colaboración del equipo de la Fundación que organizó un evento que resultó ser un éxito de público. Más de trescientas personas habían mostrado su interés por acudir al acto, hasta que una semana antes de su celebración tuvo que interrumpirse la aceptación de solicitudes por limitaciones de espacio. En tercer lugar, debo agradecer la disponibilidad de los ponentes no sólo para participar en la jornada, sino también para preparar presentaciones específicas para dicho evento, teniendo en cuenta sus apretadas agendas internacionales. Por tanto, muchas gracias a Michael, Esther, Rohini y Viju. Agradecer también al Ministerio de Educación y Ciencia su apoyo a través del proyecto SEJ2007-64340. Finalmente, deseo dedicar este volumen a mi madre, María Jesús Montalvo, por su desinteresado trabajo durante muchos años, desde el voluntariado de Cruz Roja, en pos de los más desfavorecidos. Ella fue la inspiración última para organizar la jornada que ha dado lugar a esta publicación.

Varios meses después de la jornada sobre Análisis Científico de la Ayuda al Desarrollo, el Gobierno de España estableció en el Banco Mundial el SIEF (Spanish Trust Fund for Impact Evaluation), el mayor fondo (10,4 millones de euros) para la evaluación experimental de la ayuda al desarrollo. Quiero pensar que esta jornada tuvo alguna influencia en dicha decisión.

JOSÉ GARCÍA MONTALVO

Introducción

José García Montalvo

Catedrático de Economía, Universidad Pompeu Fabra

Profesor investigador (Ivie)

En los últimos años se ha producido una evidente sensibilización de la opinión pública y los políticos respecto a la cuestión de la pobreza en muchos países del mundo. Sin embargo, los buenos deseos y las declaraciones grandilocuentes son ineficaces para conseguir un complejo objetivo que se ha mostrado muy esquivo. La comunidad internacional ha destinado más de 2,4 billones de dólares en forma de ayuda al desarrollo desde los años sesenta, sin que se puedan mostrar grandes avances, en particular en África subsahariana.

El renovado entusiasmo por las posibilidades de la ayuda al desarrollo para propiciar el avance de las sociedades menos desarrolladas será ineficaz si los nuevos fondos no son capaces de demostrar su capacidad para conseguir progresos claros en los objetivos deseados. En las sociedades occidentales el aumento del conocimiento científico y el nivel educativo de sus ciudadanos hace cada vez más necesario que los logros en cualquier materia se midan en términos de resultados y no de recursos gastados. Además, la medición de los efectos sobre los resultados debe amoldarse a los criterios científicos para conseguir la reputación necesaria para que puedan ser transmitidos con confianza a la opinión pública.

Por desgracia, en el campo de la ayuda al desarrollo, al igual que en otros campos de la evaluación de políticas públicas, el análisis científico parece tener pocos partidarios. Cuando se plantea la comercialización de

un nuevo medicamento, se considera normal que se realicen experimentos aleatorizados para comprobar su eficacia en relación con los productos ya existentes en el mercado y la ausencia de efectos secundarios significativos. Sin embargo, cuando se plantea el monto y distribución de la ayuda al desarrollo no se adopta la misma perspectiva. Los teóricos de la economía política señalan que el motivo es la escasa importancia relativa de estas cuestiones en las elecciones políticas, dado que los fondos favorecerán a ciudadanos que no pueden votar en las elecciones de los países donantes.

Por este motivo, los debates, mediatizados por la situación política de cada país, siguen centrados en el volumen de fondos que se destinan a la ayuda al desarrollo (el famoso 0,7%) en lugar del análisis de los efectos de la ayuda que actualmente se proporciona desde una perspectiva científicamente aceptable. La aproximación basada únicamente en el debate sobre el monto total de fondos destinados al desarrollo presupone que la mejora en las condiciones de vida de los destinatarios está correlacionada positivamente con el volumen de fondos destinados. Sin embargo, esta visión es ingenua pues no tiene en cuenta el coste de oportunidad de los fondos (no todos los programas de ayuda al desarrollo tienen la misma efectividad en la consecución de sus objetivos) y, además, no considera la posibilidad de que se produzcan efectos secundarios negativos. Estudios recientes muestran que los países receptores de mayor proporción de ayuda al desarrollo sufren un síndrome similar al de la *maldición de los recursos naturales*: sus niveles de democratización disminuyen a medida que aumenta la ayuda, y la búsqueda de rentas, derivada de los intentos de controlar la misma, pueden derivar en conflictos armados como en el caso de Somalia. En otras ocasiones, puede no tener ningún efecto: si la ayuda al desarrollo se transforma en redes para protegerse del mosquito de la malaria pero los ciudadanos que las reciben las utilizan para hacerse vestidos o pescar, es evidente que su capacidad de cumplir el objetivo final será muy limitada.

Finalmente, en los últimos años las grandilocuentes declaraciones de muchos jefes de gobierno de países desarrollados sobre el aumento de los fondos dirigidos a la ayuda al desarrollo no se han visto acompañadas por un desembolso efectivo en los volúmenes comprometidos.

Por tanto, al igual que en el caso de los medicamentos, es necesario realizar una evaluación científica y creíble de programas y proyectos para asegurar que son más eficientes que otros posibles usos y que no tienen efectos secundarios negativos. Asimismo, es preciso establecer un sistema de incentivos que favorezcan la consecución de los objetivos deseados. Por análisis científico de los resultados de los programas de ayuda al desarrollo se entiende la evaluación que consiste, por analogía con las ciencias experimentales, en la comparación de los resultados de un grupo tratado y un grupo de control, donde los individuos han sido asignados a dichos grupos de manera totalmente aleatoria. Por ejemplo, para estudiar el efecto de suministrar medicinas que maten los parásitos intestinales sobre el rendimiento escolar de los alumnos, se puede asignar aleatoriamente un conjunto de escuelas a dos grupos: los alumnos de las escuelas del grupo de control no reciben inicialmente las medicinas, mientras que los alumnos de las escuelas tratadas reciben dichas medicinas. De esta manera, se puede evaluar, sin sesgos de selección ni factores compuestos, el efecto real del programa de eliminación de los gusanos intestinales sobre la asistencia a las clases. Este tipo de experimentos, que es el estándar en las ciencias experimentales, comienza a considerarse también como el método más creíble en las ciencias sociales y, en particular, en la ciencia económica y el análisis de la ayuda al desarrollo.

En general, la efectividad de la ayuda al desarrollo se ve condicionada en la actualidad por tres factores fundamentales. En primer lugar, la coordinación de los donantes y sus objetivos. Los objetivos de muchos donantes están dominados por razones políticas o estratégicas donde la mejora de las condiciones de vida del país receptor de la ayuda no es prioritaria. En estas condiciones no es raro encontrar que la ayuda tiene poco efecto. En otras ocasiones, el donante sólo está interesado en el hecho de *donar* y no tanto en sus resultados. Es lo que se conoce como el efecto *warm glow*. Por este motivo, se puede entender que en muchas ocasiones los donantes estén más preocupados por cuánto dinero se desembolsa (*inputs*) que por sus resultados (*outcomes*). Asimismo, el aumento del número de países donantes requiere un esfuerzo importante de coordinación. Trabajos recientes confirman que cuanto mayor es el número de donantes en un país menor es la efectividad de la ayuda. Los

problemas de coordinación también afectan a la ayuda humanitaria. La resolución de este tipo de situaciones requiere un cambio en los incentivos (pasar de sumar *inputs* a evaluar resultados) y una mayor coordinación entre los donantes.

Un segundo factor importante es la corrupción en los países receptores. Esta cuestión, que se ha convertido en algo políticamente incorrecto, tiene una importancia fundamental en la percepción de los ciudadanos sobre la utilidad de la ayuda al desarrollo. En la última encuesta de PIPA-Knowledge Network los ciudadanos de Estados Unidos señalan que piensan que un 60% de la ayuda acaba en manos de funcionarios corruptos.

El tercer aspecto es la necesidad de proporcionar a la opinión pública, y la ciudadanía, una evaluación creíble y precisa de los efectos de los fondos destinados para la ayuda al desarrollo. En la actualidad, se puede decir que se está viviendo un nuevo resurgir del apoyo a la ayuda al desarrollo. Las encuestas de opinión señalan que sólo un 12% de los ciudadanos están por reducir o eliminar la ayuda a África, frente al 43% de 1995 y un 33% que desearían que aumentara (frente a sólo el 13% en 1995). Este renovado optimismo sobre la ayuda al desarrollo podría verse mermado por la incapacidad para mostrar de manera creíble que el creciente volumen de fondos tiene un impacto significativo en las condiciones de vida de los ciudadanos de los países del tercer mundo. Los nuevos procedimientos de evaluación experimental permiten conseguir ambos objetivos: una evaluación creíble y precisa de los efectos (tanto los directos como los efectos secundarios); y son fácilmente trasladables a la opinión pública, pues se basan en una metodología análoga a la utilizada en las ciencias experimentales: la comparación de un grupo tratado y un grupo de control.

Mediante la aplicación de estas técnicas se han conseguido avances significativos en la evaluación de programas de ayuda al desarrollo, avanzando incluso en cuestiones metodológicas que no se presentan normalmente en las ciencias puramente experimentales. Algunos ejemplos, que aparecen comentados extensamente en los capítulos siguientes, pueden servir para ilustrar los resultados, algunos enormemente originales, obtenidos mediante estos procedimientos.

En conclusión, los procedimientos de análisis del impacto de proyectos de desarrollo basados en experimentos aleatorizados ofrecen una pers-

pectiva nueva y muy potente para evaluar de manera transparente el impacto de los mismos. De esta manera, podemos saber qué funciona en ayuda al desarrollo, y qué no funciona.

La estructura del libro

El objetivo de esta obra es sintetizar las investigaciones recientes sobre el efecto de la ayuda al desarrollo. El elemento unificador de los capítulos es la utilización, como metodología fundamental, del método experimental. El capítulo 1 plantea una visión general sobre el método experimental y sus aplicaciones al campo de la economía y otras ciencias. Cada uno de los siguientes capítulos traza los descubrimientos sobre el efecto de diferentes programas de ayuda sobre los factores determinantes del desarrollo económico: la sanidad, la educación, el sistema financiero y las decisiones públicas. En todos ellos el nexo de unión es el mismo: la utilización de evidencia experimental para evaluar el impacto de los proyectos y programas analizados.

En el capítulo 2 el profesor Kremer presenta algunos resultados relativos a cambios en los incentivos entre los individuos y la evaluación experimental de sistemas sanitarios. Respecto al cambio de incentivos, Kremer analiza los efectos de la financiación de un programa de desparasitación en Kenia, la educación sanitaria, los subsidios condicionados a la mejora de los hábitos sanitarios y los efectos de las campañas educativas respecto a los efectos del sida y la vacunación.

En la última parte de su capítulo, el profesor Kremer expone los resultados de un nuevo estudio experimental sobre el sistema sanitario de Camboya, en el que se compara un sistema estrictamente público con uno de gestión privada. Esta experiencia es la primera evaluación experimental de sistemas sanitarios completos que se ha realizado.

En el capítulo 3 la profesora Duflo analiza el efecto de diversas intervenciones educativas con la utilización del método experimental. Su estudio distingue entre intervenciones para mejorar la participación en el sistema educativo, intervenciones para mejorar la calidad y estrategias globales de reforma (como la utilización de cheques escolares o la descen-

tralización). Esther Duflo reúne y analiza multitud de estudios experimentales sobre mecanismos para aumentar la participación de los estudiantes en el sistema educativo: subsidios a la familia, reparto de uniformes gratuitamente, distribución de comidas gratuitas, becas para niñas, etc.

Otro problema fundamental en los países en vías de desarrollo es el absentismo del personal educativo y sanitario. Por ejemplo, en el conjunto de la India, el 24% de los profesores no asiste a las clases. Para analizar cómo responden estos profesionales a los incentivos, el capítulo escrito por Esther Duflo analiza la efectividad de diferentes mecanismos, como, por ejemplo, la distribución de cámaras fotográficas, la utilización del director del centro para controlar la asistencia o el pago de incentivos en función del número de estudiantes.

El capítulo 4 se centra en el efecto de intervenciones en el sistema financiero. La profesora Pande expone como motivación la existencia de restricciones al crédito en los países menos desarrollados y el menor acceso de los pobres a los créditos en el sector formal. Como formas de abordar este problema, Rohini Pande analiza dos experiencias: el efecto sobre la pobreza de los bancos rurales en la India y el impacto de los microcréditos.

En el capítulo 5, el investigador del Banco Mundial Vijayendra Rao analiza la utilidad de realizar evaluaciones interdisciplinarias, tomando como ejemplo un programa de desarrollo basado en la comunidad, y propone utilizar simultáneamente métodos cuantitativos y cualitativos. El autor del capítulo se muestra crítico con los procedimientos que se basan exclusivamente en un análisis experimental cuantitativo y no tienen en cuenta la información que difícilmente puede articularse en forma cuantitativa.

La ayuda al desarrollo: su eficacia y métodos experimentales para su evaluación

José García Montalvo
Catedrático de Economía, Universidad Pompeu Fabra
Profesor investigador (Ivie)

1.1. Introducción

En este capítulo se expone una visión general sobre el método experimental, así como aplicaciones de dicho método a distintos campos científicos como la evaluación de la formación, la efectividad de programas sanitarios, etc. Al final, el capítulo incluye una descripción breve sobre cómo se ha aplicado con éxito el método experimental a la evaluación de programas de desarrollo económico en países en vías de desarrollo.

Una cuestión muy importante en todo el tema de la ayuda al desarrollo son las buenas intenciones. Todos tenemos buenas intenciones, todos queremos que las cosas funcionen, todos queremos que los países en vías de desarrollo mejoren. Pero el conocimiento es fundamental para conseguir que esas buenas intenciones se transformen en buenos resultados. En muchas ocasiones hay situaciones en las cuales las buenas intenciones se transforman en resultados perversos. Por lo tanto, el deseo no es suficiente para que un programa funcione. Como señalaba Albert Camus en *La plaga*: «El diablo que hay en el mundo proviene de la ignorancia, y las bue-

nas intenciones pueden hacer tanto daño como las malas intenciones, si falta el conocimiento». El objetivo del método experimental es proporcionar ese conocimiento sobre la efectividad de los proyectos y programas de ayuda al desarrollo.

El capítulo está dividido en cinco epígrafes. El primero es esta introducción. En el segundo se exponen las ideas básicas sobre evaluación y causalidad. En el tercero se describe el método experimental, en general, y sus aspectos más importantes. En cuarto lugar, se consideran los aspectos generales sobre la ayuda al desarrollo y sus mecanismos de evaluación. Finalmente, el quinto epígrafe contiene las conclusiones.

1.2. Evaluación y causalidad

Evaluar supone medir un contrafactual. Evaluar supone básicamente medir qué habría pasado si no hubiéramos hecho algo que hemos hecho. Por lo tanto, el problema fundamental de la evaluación es precisamente cómo calculamos que habría pasado si no hubiéramos hecho lo que hemos hecho. Como esta situación contrafactual ni se ha producido ni se producirá, es muy difícil medir lo que habría sucedido. Imaginemos que financiamos que unos niños acudan a unas escuelas en África. La pregunta sería: ¿qué habría pasado si estos niños no hubieran asistido a la escuela? Pero estos niños ya han asistido a la escuela. ¿Cómo construimos el contrafactual? ¿Cómo medimos qué habría sucedido de no haber realizado una intervención que se ha realizado, como subvencionar la asistencia a la escuela, proporcionar comida en el comedor escolar, dotar las clases con *flip charts*¹ o la intervención que sea? ¿Cómo podemos saber si tiene o no tiene efecto y cuantificar dicho efecto?

Por lo tanto, la evaluación de un contrafactual es el problema básico en la estimación de cualquier intervención. Este problema básico se resume

¹ Un *flip chart* es una ayuda visual que se utiliza en las escuelas cuando no existen suficientes libros. Contiene los fundamentos básicos de una materia (matemáticas, ciencias, etc.) y, generalmente, también incluye mapas. Se cuelga en las paredes de las clases.

en la existencia de datos faltantes, una información que nunca vamos a poder tener.

Imaginemos que estamos probando una nueva medicina. Queremos saber si tomarla tiene efectos positivos sobre el asma. Para comprobarlo le damos a un individuo dicha medicina. Pero cuando el individuo recibe el tratamiento ya no podemos saber qué habría pasado si no la hubiera recibido. Pero éste es el principal efecto que deseamos medir.

Evaluar supone también ser capaces de asignar causas a efectos. Por ejemplo, hablemos por un momento de la reciente aplicación del carné por puntos. La pregunta clave es: ¿ha supuesto una disminución de la siniestralidad la implantación del carné por puntos? Pues en principio parece que sí, pues las tasas de siniestralidad y mortalidad se han reducido. Pero ¿cómo podemos saber si de verdad se ha reducido la siniestralidad y cuánto se debe al carné por puntos y cuánto se debe a otras cosas? ¿Por qué? Porque conjuntamente con el carné por puntos ¿qué ha sucedido? La siniestralidad ya se venía reduciendo antes de la introducción del carné. Con todo, el carné por puntos ha podido tener un efecto adicional de reducción. Pero es que ya había una tendencia a la reducción de la siniestralidad, que seguramente estaba relacionada con la mejora de las carreteras, con la renovación del parque automovilístico propiciada por el Plan Prever, etc. Durante algunos meses, desde la entrada en funcionamiento del carné por puntos, ha habido también más controles de alcoholemia. Entonces, ¿es el carné por puntos o son los controles de alcoholemia la causa de la reducción de la siniestralidad que se observa en la actualidad, o al menos parte de esta reducción? Y a la vez también hay una instalación continuada de radares. Por lo tanto, ¿son los nuevos radares o es el carné por puntos el que reduce la siniestralidad? Éste es el tipo de problemas de causalidad con los que se enfrenta una evaluación. ¿Qué parte tiene que ver con la intervención y qué parte habría sucedido incluso sin dicha intervención? Si quiero evaluar el efecto del carné por puntos, quiero saber cuál es su impacto al margen de todas las otras medidas que se han adoptado. Evaluar significa identificar el efecto específico de una medida cuando se pueden producir simultáneamente diversas intervenciones. Por lo tanto hay que buscar un grupo de control. Hay que intentar de alguna manera construir un contrafactual *razonable*. En el caso del carné por puntos,

el tratamiento se produce a partir de julio de 2006. El grupo de control sería, por ejemplo, el conjunto de meses correspondientes al año anterior (2005), a partir de julio, para que tuvieran más o menos los mismos días de vacaciones, las mismas fiestas, etc., tanto en el grupo tratado como en el grupo de control. El problema de la existencia de varios efectos simultáneos (más controles de alcoholemia, más radares, etc.) no es fácil de resolver.

Por tanto, hay dos tipos de problemas fundamentales para realizar una evaluación del impacto de cualquier medida. El primero es lo que los economistas denominamos endogeneidad. Hay varias causas de este problema: la omisión de variables explicativas correlacionadas con variables incluidas en la especificación; la existencia de errores de medida o la simultaneidad en la determinación de las variables. Por ejemplo, en el caso del análisis del impacto del carné por puntos, los radares se localizan en áreas donde se ha constatado en el pasado que existe un elevado grado de accidentalidad. Por tanto, su localización es endógena y, en principio, tendrá un impacto sobre la siniestralidad mayor. La segunda cuestión es el problema de la selección muestral. Muchas veces nuestras muestras no van a ser aleatorias en el sentido que nosotros queríamos. Por ejemplo, analicemos el efecto de los cursos de formación. Imaginemos que queremos comparar el efecto que han tenido unos cursos de formación sobre la empleabilidad de un trabajador que estaba desempleado, o sobre sus salarios en el futuro. Con ese objetivo cogemos trabajadores desempleados que han asistido al curso de formación, y luego los comparamos con otros desempleados que no realizaron el curso de formación. Podemos tener la tentación de decir: para ver el efecto del curso de formación comparamos estos dos grupos y vemos si los desempleados que han asistido al curso de formación están empleados en mayor proporción que el otro grupo seis meses después. Pero estos dos grupos no son comparables. Este sencillo cálculo del contrafactual («los trabajadores que fueron al curso estarían en la misma situación laboral que los que no fueron si hubieran decidido no ir») será erróneo. ¿Por qué? Porque los desempleados que han asistido a un curso de formación tienen una serie de características no observables, como motivación, ganas de hacer cosas, tendencia a sobreponerse de las adversidades, etc., que quizás no caractericen a los individuos

del grupo de comparación. Por lo tanto, comparar estos dos grupos es como comparar peras y manzanas, aunque en muchas evaluaciones es lo que se ha hecho durante mucho tiempo. Éste es el típico problema de selección muestral: el grupo de control no se parece al grupo tratado y, por tanto, su comparación produce resultados erróneos en la evaluación de los efectos de los cursos de formación para desempleados.

¿Cuáles son las soluciones tradicionales para separar los efectos de diversas medidas que se producen simultáneamente? Por ejemplo, ¿cómo saber qué parte de la reducción de la siniestralidad se debe al carné por puntos y qué parte a las otras medidas? Durante mucho tiempo las posibilidades que se han utilizado han sido las que aparecen en la segunda columna del cuadro 1.1.² Un primer procedimiento es básicamente imponer restricciones y supuestos: formas funcionales, restricciones de exclusión, etc. La utilidad del grupo de control obtenido al utilizar estas hipótesis depende de la validez de dichas restricciones. Por ejemplo, la identificación de un modelo de ecuaciones simultáneas implica restricciones de omisión de variables de ecuaciones, restricciones lineales entre los parámetros y/o restricciones en la matriz de va-

CUADRO 1.1: **Mecanismos de identificación e imputación de efectos**

Problemas/Métodos	Tradicionales	Recientes (en ciencias sociales)
Endogeneidad	Ecuaciones simultáneas VAR Variables instrumentales	Calibración/simulación Experimentos — Aleatorizados — Naturales RDD: diseño discontinuo
Selección muestral	Corrección de Heckman	Emparejamiento (<i>matching</i>) Experimentos

² El lector debe notar que la terminología utilizada en este epígrafe proviene del campo de la economía. Estadísticos, psicólogos, sociólogos cuantitativos, etc., usan los mismos mecanismos de identificación, pero les asignan nombres que pueden ser diferentes.

rianzas y covarianzas. Para las correcciones más habituales de la selección muestral, es preciso que la especificación del proceso de selección muestral tenga la distribución supuesta, por ejemplo, en la corrección de Heckman. El problema es que estas soluciones, poco a poco, se han ido abandonando. ¿Por qué? Porque son muy rígidas y, en la mayoría de los casos, los supuestos que hay que hacer son poco creíbles o son muy estrictos.

En la actualidad, los economistas se debaten entre dos metodologías: la simulación de economías en un ordenador y los experimentos aleatorizados o naturales. La primera posibilidad implica considerar que los datos siempre son sucios. Los datos no nos van a decir lo que queremos porque en economía todos los procesos se producen de forma simultánea. La conclusión es clara: no toquemos los datos. La crítica de Lucas nos dice que cuando realizamos intervenciones de política económica no podemos saber, utilizando los datos resultantes, cuál es su efecto. Por tanto, que no es posible encontrar una fuente exógena de variación para identificar los efectos de una intervención. El motivo es que los consumidores y los productores reaccionan a esas medidas y lo que nosotros observamos está contaminado por la reacción de los individuos a los nuevos incentivos. Por tanto, observando directamente los datos no podemos aspirar a saber cuál es el efecto de una política. Esta metodología propone generar un *mundo paralelo*: creamos unos individuos representados por funciones objetivo y restricciones matemáticas; los hacemos interactuar en mercados *ficticios*; les fijamos unos parámetros para las funciones objetivo (en un proceso denominado calibración) que sí creemos que los tenemos bien capturados y no se ven afectados por la intervención; y entonces, a partir de la interacción de estos individuos *ficticios* en un ordenador, simulamos los resultados sobre el consumo, la inversión, etc., y vemos cuál es el efecto de una medida de política económica. Ésa sería la medida del impacto que tiene una determinada política. Esta metodología, basada en la calibración y la simulación, se ha extendido rápidamente en el campo de la macroeconomía.

La otra posibilidad para hacer evaluaciones, y en la que nos vamos a centrar nosotros, es la construcción de experimentos aleatorizados o la búsqueda de experimentos naturales o pseudoexperimentos.

1.3. El método experimental

El método experimental es el estándar de las ciencias que llamamos duras: la medicina, la biología, la genética, la farmacología, etc. Cualquier nuevo medicamento que quiera ser comercializado en los Estados Unidos tiene que ser aprobado por la Food and Drug Administration (FDA), que es la entidad de Estados Unidos que tiene la potestad de evaluar los nuevos medicamentos y controlar los que ya están comercializados. En todos los países existe una agencia similar. En Europa, además de las agencias nacionales, existe la Agencia Europea de Evaluación de Medicamentos.

Las agencias de evaluación, y especialmente la FDA, requieren necesariamente que se realice una evaluación del efecto del nuevo compuesto utilizando un experimento aleatorizado. No pasará la aprobación sin un informe basado en un experimento aleatorizado.³

¿En qué se basa un experimento aleatorizado? Se basa en crear dos grupos aleatoriamente: un grupo de control y un grupo tratado. Al grupo de control se le da placebo (una pastilla que no contiene ningún ingrediente activo) y al grupo tratado se le da la nueva medicina basada en la molécula que se está evaluando. Normalmente el experimento se realiza con el sistema denominado doble ciego: los pacientes que están tomando la pastilla no saben cuál de las dos pastillas están tomando y el doctor que la está administrando tampoco sabe qué pastilla está administrando (placebo o nueva medicina). La técnica estadística para evaluar si la nueva molécula tiene efecto o no es sencillísima. Cualquier estudiante de bachillerato lo podría hacer.⁴ Es un simple contraste de diferencia de medias. Por el con-

³ En Estados Unidos sólo se ha producido un caso donde un medicamento fue aprobado sin finalizar todas las fases del correspondiente experimento aleatorizado. El motivo fue que era un medicamento para una enfermedad muy rara y parecía que en las primeras pruebas había funcionado. En este caso la FDA aceptó una evaluación utilizando un procedimiento de *matching* con el uso de una muestra observacional.

⁴ Los procedimientos discutidos con anterioridad (métodos de ecuaciones simultáneas, correcciones por selección muestral, estimación por variables instrumentales, simulación y calibración, etc.) requieren la intervención de un especialista.

trario, los métodos que hemos comentado anteriormente de imponer restricciones, etc., derivan en procedimientos matemáticos complejos donde necesitase un experto para realizar la evaluación. Por tanto, lo importante es el diseño del experimento. En la evaluación experimental aleatorizada el diseño del experimento es complejo pero la evaluación es muy sencilla. En la evaluación basada en datos observacionales no es necesario en general preocuparse por el diseño pero los métodos para corregir los problemas inherentes a la naturaleza observacional de los datos son extremadamente complejos.

Si el procedimiento experimental es más creíble y exige menos supuestos en la obtención de una evaluación del impacto de una intervención, ¿por qué no se realizan todas las evaluaciones utilizando experimentos aleatorizados? Existen varias razones. En algunos casos puede ser muy caro. Además, dependiendo del objeto de estudio, puede ser que sea imposible realizar un experimento aleatorizado. Por ejemplo, si queremos analizar el impacto sobre el consumo de la reducción del tipo marginal del impuesto sobre la renta no podremos construir un grupo de control que siga pagando el mismo tipo marginal que en el pasado. Finalmente, en algunas ocasiones existen problemas de tipo ético. Pero la gran ventaja que el método experimental tiene cuando se puede utilizar, y cuando no hay problemas de tipo ético, es que proporciona unos resultados que son creíbles y difícilmente manipulables si se han seguido correctamente los protocolos. Estas propiedades son muy importantes. Si queremos realizar una evaluación que sea creíble, tenemos que buscar métodos que no sean fácilmente manipulables y cuyos resultados se puedan replicar con facilidad.

¿Por qué las agencias de medicamentos requieren que haya un experimento aleatorizado? Lo requieren para comprobar los tres principios básicos que exigen. El primero es que la nueva molécula o droga tenga efecto realmente sobre el problema que pretende resolver. Por lo tanto, que reduzca el tamaño de un tumor, por ejemplo, o reduzca el recuento vírico. Segundo, que el procedimiento sea más efectivo que los procedimientos existentes. Por tanto, que la nueva molécula supone una mejora frente a los medicamentos disponibles con anterioridad. Y en tercer lugar, que no tenga efectos secundarios perjudiciales o que los mismos sean limita-

dos. Estos tres principios son importantes no sólo en el campo de la medicina, sino en todos los campos del conocimiento.

¿Se puede aplicar el procedimiento experimental a las ciencias sociales? La respuesta es afirmativa. Por supuesto que se puede hacer. Se ha utilizado, y cada vez se utiliza más, para evaluaciones en muchos campos de las ciencias sociales. Tenemos experimentos, por ejemplo, sobre el efecto del tamaño de las clases sobre el aprendizaje; tenemos experimentos sobre el efecto de los cheques escolares en muchos sitios, como en Colombia, el Programa de Ampliación de la Cobertura de la Educación Secundaria (PACES) y en Nueva York (School Choice Scholarship Foundation Program); tenemos experimentos sobre el impacto de intervenciones para la reducción de la pobreza, como el programa Progresá en México; tenemos experimentos sobre el efecto del programa Moving to Opportunity para familias de barrios degradados de cinco ciudades de Estados Unidos (Baltimore, Boston, Chicago, Los Ángeles y Nueva York). Y hay muchísimos más. Éstos son simplemente algunos ejemplos de experimentos en el sentido que se ha expuesto en los párrafos anteriores: existencia de un grupo tratado y de un grupo de control, asignación aleatoria de los grupos, etc. Por tanto, en las ciencias sociales también se pueden hacer experimentos aleatorizados.

Antes de comentar la aplicación del procedimiento experimental al proyecto de desarrollo, es interesante considerar un caso bien conocido de aplicación en ciencias sociales. Un ejemplo clásico de este tipo de experimentos es el proyecto STAR (Student Teacher Achievement Ratio). El STAR es un experimento que se hizo en Tennessee y que duró cuatro años. Fue diseñado para evaluar el efecto del tamaño de las clases sobre el aprendizaje y costó doce millones de dólares.⁵ El coste viene justificado por el enorme número de estudiantes que participaron en el experimento. Se diseñaron tres tipos de tratamientos. El grupo de control era una clase de tamaño normal, entendida como 22-25 estudiantes; el primer tratamiento eran clases pequeñas de 13-17, y el segundo tratamiento eran

⁵ Como comprobaremos con posterioridad sólo experimentos muy masivos tienen un coste tan elevado.

clases de tamaño normal con un profesor de apoyo. El protocolo establecía que los estudiantes, cuando entraban en la guardería, eran asignados aleatoriamente a uno de estos tres grupos y continuaban en el mismo grupo durante los cuatro años hasta llegar a tercer grado. En el primer año participaron 6.400 estudiantes en 108 clases pequeñas, 101 clases normales y 99 clases con profesor de apoyo. En total, durante los cuatro años participaron 11.200 estudiantes.

En algunos momentos el protocolo no se siguió exactamente, y éste es potencialmente un problema en el diseño de los experimentos. Por ejemplo, en este caso los padres de los niños que iban a las clases más grandes querían que sus hijos fueran a las clases pequeñas. Como en este caso el tratamiento era observable, no hay posibilidad de aplicar un procedimiento de doble ciego: todos los participantes saben cuál es la clase placebo. Obviamente, dado este diseño, había padres que querían que su hijo fuera a la clase pequeña. Para resolver este problema al final del primer año se sortearon otra vez los niños que iban a las clases de tamaño normal. Pero el resultado del sorteo era el mantenimiento en una clase de tamaño normal o el cambio a una con profesor de apoyo. No existía la opción de ser asignado a una clase pequeña. Además, un 10% de los alumnos fueron cambiados de clase fundamentalmente por problemas de comportamiento o incompatibilidad con otros estudiantes del grupo.

La medida de resultados utilizada es una medida de aprendizaje. En concreto, se utilizó la puntuación de los alumnos en el test combinado de matemáticas y lectura de Stanford, lo que se conoce como Stanford Achievement Test. El impacto de las diferentes intervenciones se puede medir a partir de una simple regresión:

$$Y_i = \beta_0 + \beta_1 \text{Clasepequeña}_i + \beta_2 \text{Papoyo}_i + u_i.$$

Donde Y es el resultado del test, Clasepequeña es una variable dicotómica que toma valor 1 si el estudiante estaba asignado a una clase pequeña; y Papoyo es una dicotómica que toma valor 1 si el estudiante estaba en una clase de tamaño regular con un profesor de apoyo.

Los resultados del cuadro 1.2 indican que tener un profesor de apoyo, excepto en el primer grado, no tenía ningún efecto. Los alumnos que

CUADRO 1.2: Resultados básicos del proyecto STAR

Intervención/curso	G	1.º	2.º	3.º
Clase pequeña	13,90* (2,45)	29,78* (2,83)	19,39* (2,71)	15,59* (2,40)
Normal con profesor de apoyo	0,31 (2,27)	11,96* (2,65)	3,48 (2,54)	- 0,29 (2,27)
Constante	918* (1,63)	1.039* (1,78)	1.157 (1,82)	1.228* (1,68)
N	5.786	6.379	6.049	5.967

Nota: Entre paréntesis las desviaciones estándar.

* Indica significativo al 5%.

iban a una clase pequeña tenían un nivel de conocimiento superior aunque el efecto era muy pequeño: en torno a 20 puntos sobre una media de 1.100 puntos. El efecto, tanto en términos de porcentajes sobre las medias como en términos de desviación estándar, es muy pequeño. El coste era muy grande pero el beneficio era muy pequeño. Además, la mayoría del efecto de las clases pequeñas se concentraba en el primer curso. Continuar en una clase pequeña en los cursos superiores no generaba ganancias adicionales.

El STAR es un caso de experimento en ciencias sociales. Pero ¿podemos confiar en los resultados que no sean experimentales? Analicemos dos ejemplos de resultados obtenidos con datos observacionales, que quiere decir que no se ha hecho un experimento. Por ejemplo, durante mucho tiempo se ha creído que el chequeo anual es un mecanismo que puede alargarnos la vida. El fundamento de esta creencia era un conjunto de estudios realizados con datos observacionales durante los años cincuenta y sesenta. Por lo tanto, los datos tenían los problemas que he comentado antes, sobre todo de selección muestral. Recientemente se ha hecho un experimento aleatorizado para evaluar el efecto del chequeo anual. Los resultados experimentales muestran que el chequeo anual no tiene ningún efecto sobre la extensión de la vida. Vayan ustedes todos los años una vez al médico pero sepan que no afecta en nada a su esperanza

de vida. Pero durante muchos años se ha creído que sí que tenía efecto. ¿Cómo ha podido pasar esto? Porque los datos observacionales tenían un problema importante de selección muestral ¿Quién es la gente que se va a hacer el chequeo anual? La gente que ya de entrada se cuida mejor, come mejor, vive mejor. Entonces, si hacemos un estudio observacional donde comparamos los individuos que se hacen un chequeo anual y los que no se lo hacen lógicamente encontramos que los que van al médico al menos una vez al año tienen mayor esperanza de vida. Pero no viven más por el chequeo anual. Viven más porque comen mejor, hacen deporte y, en general, se cuidan más. Dentro de esos cuidados se incluye hacerse un chequeo anual. Por lo tanto, los estudios observacionales nos pueden dar resultados incorrectos. Hay que tener cuidado.

Otro ejemplo es el efecto de la terapia hormonal en mujeres con menopausia. Los estudios observacionales mostraban una mejora, tanto física como psicológica, en las mujeres que seguían la terapia de sustitución hormonal. Hace pocos años se hizo el primer estudio experimental sobre este tipo de terapia.⁶ Los resultados no permitían concluir que la terapia de sustitución hormonal tuviera ningún efecto en los trastornos causados por la menopausia.

Por lo tanto, ¿cómo podemos saber si cuando evaluamos algo observacionalmente, o sea, con datos no generados experimentalmente, obtendremos los mismos resultados que cuando evaluamos experimentalmente? Al economista Robert LaLonde se le ocurrió coger un experimento aleatorizado y utilizar el siguiente argumento: tengo datos buenos de tipo experimental. ¿Qué pasaría si, olvidando que tengo estos datos, intentara hacer la evaluación por los procedimientos habituales basados en datos observacionales? Por tanto, el grupo de comparación en lugar de ser obtenido a partir del diseño experimental se obtiene con datos observacionales. En este caso, como disponía de los datos correctos sobre el contrafactual, podía comparar el sesgo y la precisión de la estimación con datos no experimentales. Si el resultado fuera el mismo entonces hacer experimentos no es tan interesante dado que la estimación con el uso de datos

⁶ El coste fue de decenas de millones de dólares.

CUADRO 1.3: El procedimiento experimental frente a los resultados observacionales: la crítica de LaLonde

	Control	Hombres	Mujeres
Experimental	NSW	886 (476)	851 (317)
Observacional: estimación en una etapa	PSID	-1.228 (896)	2.097 (491)
	CPS	-805 (484)	1.041 (505)
Observacional: estimación en dos etapas (Heckman)	PSID	-1.333 (820)	1.129 (385)
	CPS	-22 (584)	1.102 (323)

Fuente: LaLonde (1986).

observacionales proporcionaría la misma respuesta. LaLonde (1986) utilizó la evaluación de los programas de formación del National Supported World Demonstration de Estados Unidos.

Los resultados experimentales aparecen en la primera fila del cuadro 1.3. Se refieren a los salarios de la gente que hizo los cursos de formación frente al grupo de control.⁷ Si hacemos la comparación experimental resulta que tanto en hombres como en mujeres los cursos de formación tuvieron un efecto positivo en sus salarios en el futuro, en su productividad, etc. Si empleamos los procedimientos habituales formaríamos grupos de control a partir de datos observacionales de individuos *comparables*, ya sean del Panel Study of Income Dynamics (PSID) o del Current Population Survey (CPS). Utilizando diferentes procedimientos, como regresiones en un paso, bietápico con corrección muestral, etc., los resultados observacionales y experimentales proporcionan conclusiones muy diferentes, que pueden implicar incluso diferencias negativas: los individuos que no hicieron los cursos cobrarían más que los que hicieron los cursos. El rango de variación va desde -1.228 dólares hasta -22 dólares en el caso de los hombres, por ejemplo. El procedimiento experimental, *la verdad* en este caso, proporciona diferencias positivas de 886 y 851 dólares para hombres y mu-

⁷ También se podría ver el efecto en términos del nivel de desempleo seis meses después de finalizados los cursos de formación.

jeros respectivamente. Utilizando datos observacionales y las técnicas habituales los resultados son muy sensibles al grupo de comparación elegido y al tipo de estimación.

¿Qué aprendemos de esto? En primer lugar, desde la crítica de Lalonde (1986) sabemos que los métodos no experimentales producen resultados muy diferentes dependiendo del grupo de control utilizado y de la especificación econométrica, o los supuestos que estamos imponiendo en el análisis. En segundo lugar, los resultados de los métodos experimentales pueden ser muy diferentes de los resultados de los métodos no experimentales. Finalmente, incluso si los procedimientos no experimentales pasan todos los contrastes de hipótesis y, por tanto, aparentemente parece que son correctos, son incapaces de replicar los resultados experimentales.

Por lo tanto, si no podemos hacer un experimento aleatorizado y los resultados de los datos observacionales no son creíbles (son muy inestables y dependen de la configuración del grupo de control), ¿qué podemos hacer? Una posibilidad es la utilización de nuevas técnicas científicas basadas en experimentos naturales (la naturaleza proporciona la aleatorización o una fuente exógena de variación); también podemos construir cuasiexperimentos o procedimientos de emparejamiento a la *Rubin*; intentar buscar alguna variable instrumental; o utilizar un *Regression Discontinuity Design*, etc. Estos procedimientos resuelven los problemas de endogeneidad y/o selección muestral sin imponer excesiva estructura (supuestos).

Evidentemente, la evaluación experimental no es una panacea. Existen casos en los que es posible, y conveniente, aplicarla, mientras que en otras situaciones es imposible aplicarla. Por ejemplo, cuando existen externalidades importantes o el ámbito de aplicación es muy grande (por ejemplo, sistemas complejos o países enteros) resulta difícil realizar evaluaciones experimentales. También sería muy difícil analizar cuasiexperimentalmente las causas de guerras civiles (Montalvo y Reynal-Querol 2005) o genocidios (Montalvo y Reynal-Querol 2008). Además, el procedimiento experimental tiene limitaciones metodológicas que hacen conveniente la participación de un buen profesional experto en su aplicación. Finalmente, y a diferencia de muchas otras metodologías empleadas habitualmente en el campo de la economía, el análisis experimental requiere un trabajo muy cuidadoso de diseño, planificación y aplicación.

Como se ha señalado anteriormente el método experimental tiene sus propias limitaciones, aunque suelen ser menos importantes que las existentes en otras metodologías. Los principales problemas del método experimental se pueden agrupar en dos bloques: amenazas a la validez interna de los resultados y amenazas a la validez externa. Las amenazas a la validez interna de los resultados más relevantes son las siguientes:

- a)* Fallos en la randomización. Si el tratamiento no es asignado aleatoriamente, los resultados de la comparación del grupo tratado y el grupo de control no son relevantes.
- b)* Fallos en el seguimiento del protocolo. Incluso si los individuos han sido asignados aleatoriamente a los grupos de tratamiento y control podría suceder que algunos individuos en el grupo de control recibieran el tratamiento o algunos del grupo tratado no lo fueran efectivamente.
- c)* Abandono de la muestra. Después de la asignación a los grupos de tratamiento o control, algunos individuos pueden decidir abandonar el experimento. Si dicho abandono está relacionado con el tratamiento, entonces los resultados del experimento serán sesgados.
- d)* Efectos experimentales. Los individuos, sabiendo que son observados y analizados, pueden actuar de manera diferente a la que actuarían en una situación normal.

Las amenazas más importantes a la validez externa de los resultados experimentales son los siguientes:

- a)* La utilización de una muestra de individuos que no sea representativa, o de un programa que no sea representativo (no sea similar a los programas estudiados).
- b)* El estudio de un programa que no es similar a los que se intentan evaluar.
- c)* Los efectos de equilibrio general. La generalización de un experimento del que se tiene resultados para un ámbito reducido puede ser complicada por estos efectos. Por ejemplo, un determinado fármaco contra la malaria puede ser muy efectivo en una aldea, pero

si se extiende al resto de un país el parásito de la malaria puede mutar y el fármaco podría perder su efectividad.

1.4. La evaluación de la ayuda al desarrollo

En los epígrafes anteriores se han expuesto los principios básicos de una evaluación científica en general. En éste se particularizan dichos métodos al caso de la ayuda al desarrollo.

La pregunta básica, que muchas veces se hace al revés, es la siguiente: si con un dólar al día se puede salvar un niño en el tercer mundo, ¿qué ha pasado con los 2,4 billones de dólares en ayuda al desarrollo que se han desembolsado entre 1960 y el 2005? Si con el dinero de la guerra de Irak resolvemos todo el problema del desarrollo en el mundo, entonces, ¿cómo es que hasta ahora, que se ha gastado en ayuda al desarrollo bastante más que en la guerra de Irak, el problema de la pobreza no se ha resuelto? Algo está pasando. La ecuación no es tan sencilla.

1.4.1. La evidencia macroeconómica

¿Qué es lo que nos dicen los estudios macroeconómicos? La mayoría de los estudios macroeconómicos, incluidos los más recientes denominados de *tercera generación*, nos dicen que cuando estudiamos las variables agregadas, como puede ser el producto per cápita, y las intentamos relacionar con la ayuda al desarrollo que reciben los países, encontramos que no hay efecto. Pero recordemos que los datos macroeconómicos no son experimentales. El estudio de Boone (1996), por ejemplo, es uno de los primeros que inicia esta tercera generación de trabajos sobre el impacto de la ayuda al desarrollo. Dicho estudio señala que la ayuda al desarrollo no aumenta significativamente ni la inversión ni los indicadores de desarrollo humano, sólo aumenta el tamaño del gobierno.

Posteriormente, el estudio de Burnside y Dollar (2000) señala que la ayuda al desarrollo sólo funciona cuando se combina con buenas políticas monetarias, fiscales y comerciales. El problema es que unos pocos años después, Easterly, Levine y Roodman (2003) muestran que estos resultados no son robustos a cambiar el tamaño de la muestra, a extender la

muestra a más años, a incluir más países, etc. Por tanto, los resultados de Burnside y Dollar (2000) hay que ponerlos en cuarentena.

La misma conclusión se puede obtener de los comentarios de Banerjee et al. (2006, 55):

Esta línea de investigación tiene serios problemas que, en el momento actual, impiden que los resultados puedan ser percibidos como remotamente dignos de confianza, por más que uno quiera creerse los resultados.

A pesar de ser publicado en la *American Economic Review*, el trabajo de Burnside y Dollar no es convincente. El análisis usa un índice de política económica que combina el superávit presupuestario, la tasa de inflación y una medida de apertura comercial siendo al menos dos de ellas medidas de resultados y no políticas. También es claro de la forma en que el índice es construido que los resultados no son robustos; los intentos de trabajar con las tres medidas fallan, como también falla un análisis por componentes principales, y el índice final se construye usando un regresión de crecimiento sobre política que es arbitraria y, peor, es inconsistente con la ecuación de interés.

Rajan y Subramanian (2007) utilizan datos de panel para analizar el efecto de la ayuda al desarrollo sobre el crecimiento de los países receptores y técnicas para corregir el posible sesgo de que un crecimiento bajo pueda implicar un aumento de las contribuciones al país receptor. Incluso con esta corrección Rajan y Subramanian (2007) encuentran poca evidencia de una relación positiva (o negativa) entre la ayuda internacional recibida y el crecimiento económico. Tampoco encuentran evidencia de que la ayuda internacional sea más efectiva en países que tienen mejores políticas, mejor situación geográfica, o reciben un tipo de ayuda u otra. Rajan y Subramanian (2006) argumentan que los efectos beneficiosos de la ayuda pueden ser contrarrestados por efectos negativos en la competitividad de los países receptores, reflejada en un mejor crecimiento relativo de los sectores exportadores, intensivos en trabajo y las manufacturas en su conjunto.

En resumen, se podría decir que existe un amplio consenso que, desde una perspectiva macroeconómica, es difícil encontrar un efecto de la ayu-

da al desarrollo sobre el crecimiento económico. En vista de los anteriores comentarios existen dos motivos fundamentales que pueden justificar este resultado. Una causa puede ser la falta de control de la multitud de factores que intervienen en el proceso y, por tanto, la inadecuación de la metodología de regresiones de corte transversal (o incluso panel) para el estudio del efecto de la ayuda al desarrollo (v. Banerjee et al. 2006). Al mismo tiempo que un país en vías de desarrollo recibe ayuda internacional se producen otros muchos efectos que, en muchas ocasiones, es difícil medir por la ausencia de estadísticas adecuadas. Por tanto, toda una serie de factores no observables y efectos simultáneos hacen muy difícil mostrar a partir de estudios macroeconómicos los efectos de la ayuda al desarrollo. Además de los problemas metodológicos existen motivos teóricos que justificarían los nulos efectos de la ayuda al desarrollo encontrados en los estudios macroeconómicos. Los métodos experimentales pretenden superar los problemas metodológicos de las regresiones de corte transversal y son el objeto de estudio fundamental de este libro.

Una segunda causa podría ser la existencia de efectos negativos que compensarían los potenciales efectos positivos de la ayuda al desarrollo. Rajan y Subramanian (2006) utilizan esta línea de argumentación. Pero ¿qué otros elementos pueden estar socavando los efectos positivos de la ayuda al desarrollo? Fundamentalmente existen tres factores: la corrupción en los países receptores, los incentivos de los donantes y la coordinación-descoordinación entre los mismos.

1.4.1.1. La corrupción

Se ha convertido en un tema en parte políticamente incorrecto. No se puede decir que los países que reciben ayuda son corruptos porque no está bien visto, pero la realidad es que lo son. Tienen unas instituciones muy débiles y, por lo tanto, la tubería donde se mete un montón de ayuda al desarrollo pierde por el camino gran cantidad de los fondos, como las acequias españolas. Normalmente cuando se habla de corrupción se habla de una manera muy difusa. Por lo general, no se proporcionan cifras concretas ni evaluación científica. ¿Cómo podríamos saber el volumen de corrupción en los países que reciben ayuda al desarrollo? Reinikka y Svensson (2004) han utilizado unos instrumentos denominados PETS (Public

Expenditure Tracing System). La idea es medir cuánto dinero se envía al gobierno central de un país en desarrollo para un determinado proyecto y medir, en los pueblos y aldeas donde se tienen que realizar las intervenciones, cuánto dinero ha llegado. Reinikka y Svenson (2004) analizan un proyecto para construir escuelas en Uganda. Utilizando información contable van arriba y luego abajo del todo, comparando cuántos son los fondos que se le han dado al gobierno central y cuántos han llegado a los gobiernos locales, que son los que tienen que construir las escuelas. Los investigadores suman todos los gastos y toda la inversión en escuelas de esas aldeas. ¿Cuánto llega al final de la tubería? En el caso de Uganda, llegaba un 13%. En el caso de Tanzania, llegaba un 20% (Reinikka y Svenson 2005). Por lo tanto, la corrupción es muy significativa. En el caso de Indonesia, por ejemplo, Olken (2006) muestra que en programas antipobreza el 18% del arroz se desvió de su objetivo. Lo mismo sucedió en la construcción de carreteras, donde el 28% de los fondos no se utilizó para el fin previsto.

¿Hay soluciones para este problema? Se han propuesto diversas medidas. Por ejemplo, Reinikka y Svenson han encontrado que campañas informativas entre los padres de las escuelas que deben recibir los fondos (por ejemplo, en periódicos locales) pueden hacer que la pérdida se reduzca mucho. En concreto, de llegar sólo el 13% puede acabar llegando el 80%. Cuando le dices a la gente: «ese dinero es tuyo; te tienen que construir una escuela ahí y tanto tiene que llegar» el efecto es muy importante. Olken (2006) propone un sistema más tradicional, tipo *top-down*, basado en vigilar que los fondos no se desvíen y hacer un seguimiento para que esos fondos no desaparezcan.

1.4.1.2. Los incentivos de los donantes y la política económica de la ayuda internacional

El segundo problema es la cuestión de los incentivos de los donantes.⁸ Sabemos que se han gastado en ayuda al desarrollo 2,4 billones de dólares desde 1960. Pero esos 2,4 billones de dólares ¿a qué se han dedicado, o

⁸ Montalvo (2007) presenta un análisis detallado de la economía política en torno a la ayuda al desarrollo.

qué incentivos tienen los donantes en la asignación de estos fondos? Hay toda una serie de problemas. El primer problema es que los impuestos que pagan los ciudadanos de los países que son donantes están muy débilmente relacionados con los ciudadanos que se deben beneficiar de esos fondos. Cuando nosotros, con nuestros impuestos, financiamos la sanidad, nosotros mismos luego exigimos que la sanidad no tenga retrasos, que no haya listas de espera, etc. Pero aquí los que recaudan, y los donantes, transfieren esos fondos a otro país. Los ciudadanos de dicho país, que son los que se benefician, no votan en el país de los donantes. Por lo tanto, se rompe esa relación que debe existir siempre para que los fondos funcionen. No hay rendición de cuentas ni exigencia por parte de los beneficiarios, que no saben exactamente cuánto les corresponde.⁹

Luego están todas las cuestiones de los efectos mediáticos, estrategias de comunicación, etc. Con una buena estrategia de comunicación se puede conseguir que los ciudadanos piensen que se dedica a ayuda al desarrollo mucho más de lo que efectivamente se dedica.

Además, los donantes tienen objetivos múltiples y algunos de ellos son objetivos de tipo político-estratégico y, por lo tanto, ajenos al desarrollo y a la lucha contra la pobreza. ¿Por qué la ayuda al desarrollo no tiene efecto? Porque en muchas ocasiones los objetivos de esos fondos no son mejorar la situación económica de los ciudadanos del país que los recibe. Por ejemplo, en 2007 el presidente Bush propuso 23.700 millones de ayuda para los países pobres. De esos 23.700 millones solamente 3.700 millones tienen como objetivo prioritario la lucha contra la pobreza. Con lo cual, si intentamos ver cuál es el efecto de esos 23.700 millones sobre la reducción de la pobreza, encontramos que éste es muy pequeño o nulo para el volumen de fondos invertido.

Asimismo, está toda la cuestión de los motivos de los donantes. La utilidad que obtienen los donantes muchas veces se deriva del hecho de dar la ayuda, no tanto de sus resultados. «Nosotros damos la ayuda y estamos muy satisfechos de que nuestra sociedad sea tan solidaria con los países en vías de desarrollo. Damos hasta tal porcentaje del PIB.» Pero éstos son los

⁹ Sobre cuestiones de economía política asociadas a la ayuda al desarrollo se puede consultar Gibson et al. (2005).

inputs. Lo importante deberían ser los resultados que se obtengan con dicha ayuda y no cuánta ayuda se concede. Por tanto, hay un claro énfasis en los *inputs* y no en los *outputs*. También está la cuestión de pensar que simplemente incrementando el volumen de la ayuda se resuelve todo. Otra vez, la visión de los *inputs*. Sabemos que por la tubería se escapan muchos fondos. La solución: metamos más en la tubería y así llegará más. Ésta es la aproximación de Jeffrey Sachs. Pero pensemos en una tubería que ya hace agua. Si metemos mucho más caudal en la tubería puede resultar que no es que pierda parte del agua, es que podría simplemente desaparecer, explotar, y no canalizar nada. Esto es lo que sucede en un país en vías de desarrollo donde los cauces institucionales son débiles. Si metemos mucho por la tubería podemos acabar debilitando todavía más las instituciones de los países receptores de la ayuda (Djankov, Montalvo y Reynal-Querol 2008).

Además, están los incentivos propios de las burocracias, donde los funcionarios intentan maximizar el tamaño de su despacho, el gasto se distribuye con mucha rigidez, la evaluación se produce en función del volumen de fondos desembolsado y no de los resultados, etc. En fin, que los incentivos de los donantes están mal diseñados para ser efectivos sobre el objetivo de reducir la pobreza.

El Plan Director de Cooperación Española señala que la «gestión basada en resultados incluye un cambio en el modo de pensar y actuar: en lugar de atender a los recursos que se van a emplear, analizando más tarde los resultados e impactos obtenidos, se fijarán claramente los resultados e impactos deseados y, en función de ellos, se identificarán los recursos y actuaciones que se necesitan». Éstos son objetivos muy loables. Pero es importante decirlo y hacerlo y que no quede en papel mojado.

1.4.1.3. La coordinación de los donantes

Otro de los problemas para la eficacia de la ayuda internacional es la coordinación de los donantes, lo que yo llamo la aproximación impulsiva frente a la aproximación reflexiva. La primera vez que estuve en Etiopía hubo una reunión en la que me llamó mucho la atención una extraña situación. Había representantes del Banco Mundial así como de otras instituciones internacionales y representantes de ONG. Los representantes de

las ONG tenían, en general, una visión impulsiva: «la gente fuera está muriendo, ¿qué hacemos aquí sentados planificando? Hay que salir, hay que actuar». En el otro lado estaban las instituciones como el Banco Mundial. Sus representantes tenían un discurso totalmente distinto: «¿Han llegado los contratos?; ¿Quiénes tienen contratos?; ¿Ha llegado el dinero del Banco Mundial?». Entre estas dos visiones hay que llegar a un entendimiento. Tiene que haber coordinación, pero tiene que haber actuación también. No hay duda de que hay problemas importantes de coordinación, en muchos casos extremos. Un ejemplo: una niña contrajo el sarampión en Riga con posterioridad al tsunami. Entonces se pensó que podía haber una epidemia de sarampión. Se llamó a varios médicos alemanes para que evaluaran la situación. Después de una serie de averiguaciones se comprobó que la niña había sido vacunada tres veces contra el sarampión por tres ONG distintas. Obviamente, tanta inoculación había provocado la aparición de los síntomas del sarampión. Un problema evidente de coordinación, combinado con la medida del éxito en función de los *inputs* (cuantas más vacunas administradas mejor quedan los indicadores del informe anual).

Hay que tener muy claro qué es lo que hace cada uno y hay que planificar. Hay que actuar pero hay que planificar también porque sin planificación las actuaciones a veces no consiguen los resultados que se desean y se desperdician recursos escasos.

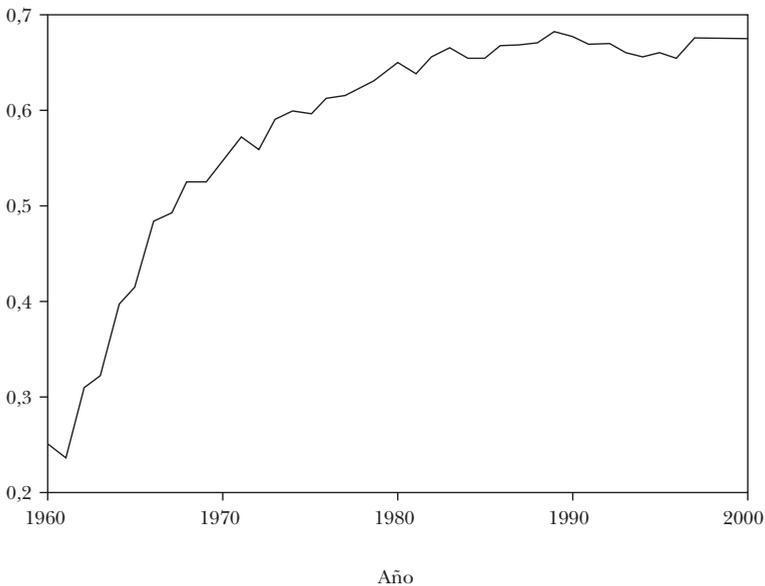
La necesidad de coordinación aumenta ante el incremento significativo de la fraccionalización, o multiplicidad, de los donantes. En la actualidad muchos países que habían sido tradicionalmente receptores de ayuda al desarrollo participan en el grupo de los donantes como por ejemplo China. El gráfico 1.1 muestra la evolución del nivel de fraccionalización de los donantes, medido a partir del índice FRAC.

$$FRAC = 1 - \sum_{i=1}^N \pi_i^2$$

Donde π es la proporción de cada donante i en el volumen total de ayuda al desarrollo.

Djankov, Montalvo y Reynal-Querol (2006a) y Djankov, Montalvo y Reynal-Querol (2006b) han mostrado que la efectividad de la ayuda al desarrollo disminuye cuanto mayor es el número de donantes por país debi-

GRÁFICO 1.1: Nivel de fraccionización de los donantes



Fuente: Djankov, Montalvo y Reynal-Querol (2006b).

do a la posible incompatibilidad de sus objetivos y la competencia entre los mismos.

1.4.2. La evidencia microeconómica y el método de experimentos aleatorizados

Cada vez son más los países y los programas destinados a la ayuda al desarrollo de países de baja renta per cápita. El motivo último es que las encuestas de opinión recientes muestran un renovado interés de los ciudadanos por apoyar a los países subdesarrollados, en concreto a África. Pero este apoyo evoluciona en ciclos. En 1995 las encuestas de opinión en Estados Unidos mostraban que el 43% de los ciudadanos eran partidarios de reducir o eliminar la ayuda al desarrollo. En la actualidad, ese porcentaje es del 12%, lo que muestra un renacimiento de una visión más optimista. Por el contrario, en la actualidad el 33% de los ciudadanos desearía aumentar la ayuda frente al 13% del año 1995. Pero este renovado optimismo volverá otra vez a entrar en un ciclo de idealismo-grandes esperan-

zas-resultados decepcionantes-cinismo si los organismos encargados de la ayuda internacional no consiguen demostrar de manera creíble, y transmisible a la opinión pública, que con la ayuda al desarrollo se están consiguiendo resultados.

La pregunta es, por tanto: si sabemos cómo evaluar correctamente desde el punto de vista técnico, ¿por qué muchas actuaciones y programas de ayuda al desarrollo no se evalúan, o no se evalúan correctamente? Hay diversas explicaciones. *Pritchett* argumenta que los promotores de las intervenciones tienen interés en mostrar que los programas tienen siempre un gran impacto. Si se hiciera una evaluación aleatorizada se revelaría el verdadero impacto, y no existe garantía de que dicho resultado sea ni positivo ni importante. Por lo tanto, una evaluación externa haría a la agencia correspondiente perder el control sobre los resultados públicos de la intervención.

Pero, cuando se decide realizar una evaluación, ¿por qué no se evalúa correctamente? Hay dos teorías contrapuestas. La primera la denominaremos la teoría del paracaídas, que justifica no realizar una evaluación aleatorizada de forma un tanto irónica. Imaginemos que queremos evaluar el efecto de los paracaídas en evitar un trauma severo cuando se salta de un avión. La teoría general nos diría que el método adecuado para evaluar el efecto del paracaídas es un experimento controlado. Cogemos a un grupo de control y a un grupo tratado. Al grupo de control no le damos paracaídas y al grupo tratado le damos paracaídas. Tiramos a los individuos de los dos grupos y comprobamos la diferencia en graves traumas. Obviamente, esto no tiene sentido. La conclusión sería que no es necesario utilizar el método experimental en todos los casos y que esto de la evaluación aleatorizada no es tan importante.

Una visión contraria a la teoría del paracaídas señala que no se hace una evaluación adecuada en muchas ocasiones porque, a diferencia de las medicinas y la seguridad alimenticia, que son consideradas muy importantes en términos políticos (una intoxicación alimentaria genera una avalancha de malas noticias en la prensa), la educación (*Angrist 2004*) y la ayuda al desarrollo no dan tanto rédito político y los aciertos y errores se ven a muy largo plazo. Por lo tanto, no hay interés político en realizar una evaluación científica y creíble de este tipo de programas. La ideología y las ideas preconcebidas se imponen al método científico.

La teoría del paracaídas aplicada al análisis y la evaluación de la ayuda al desarrollo implica supuestos muy fuertes. En primer lugar, supone que las intervenciones y la ayuda al desarrollo nunca tienen efectos secundarios contraproducentes. También supone que los procedimientos alternativos son todos igualmente eficientes o que no hay restricciones presupuestarias. Por lo tanto, hagamos cualquier cosa pues es efectiva, como si no hubiera ningún tipo de restricción presupuestaria.

La ayuda al desarrollo puede tener efectos secundarios muy negativos. Por ejemplo Maren (1997) señala que la lucha por el control de la ayuda fue una de las causas de inicio del conflicto en Somalia. Los señores de la guerra peleaban por el control de la ayuda para venderla. Otro ejemplo: cuando se proporciona ayuda al desarrollo a países en conflicto es muy importante asegurarse de que dicha ayuda no acabe en el mercado negro y pueda ser utilizada para la financiación de las facciones que están envueltas en el conflicto. Mary Anderson (1999) señala lo siguiente:

Creemos que la ayuda internacional es positiva. El mundo es un sitio mejor cuando se puede ayudar a la gente que sufre... El reto para los profesionales de la asistencia al desarrollo (en entornos conflictivos) [...] es encontrar cómo hacer el bien de manera que, sin desearlo y sin darse cuenta, no reduzcan las fortalezas locales, promuevan la dependencia y permitan que los recursos puedan ser utilizados para seguir con la guerra.

La ayuda al desarrollo también puede tener efectos negativos sobre la competitividad como muestran Rajan y Subramaniam (2006) y efectos perniciosos sobre el nivel de democracia de los países que la reciben (Djankov, Montalvo y Reynal-Querol 2008).

Además, se pueden presentar muchos efectos que son, inicialmente, impredecibles. Easterly (2006) cita un largo listado de efectos secundarios, muchos sofisticados y difíciles de prever. Éste es un motivo importante para realizar evaluaciones con técnicas adecuadas: porque hay efectos secundarios en los cuales puede ser que inicialmente no hayamos pensado que se puedan producir. Sólo después de una evaluación rigurosa se puede descubrir este tipo de efectos impredecibles. Un caso sorprenden-

te es el que se produjo en un programa de subsidios a las rentas de agricultores en China. En principio, y teniendo en cuenta el bajo nivel de renta de las familias beneficiarias se esperaba que dicha ayuda se transformara en consumo. Para sorpresa de los evaluadores una parte importante de la ayuda fue ahorrada por los agricultores (Ravallion y Chen 2005).

El segundo supuesto de la teoría del paracaídas es que todos los programas son igualmente efectivos. Antes de tirarnos del avión ya sabemos que sobreviviremos o no dependiendo de tener paracaídas o no. Sin embargo, no todos los programas son igualmente efectivos. Por ejemplo, un programa de eliminación de parásitos intestinales en niños cuesta 3,5 dólares para conseguir un año adicional de asistencia a la escuela. Un programa de provisión de uniformes y libros gratis cuesta 99 dólares por año adicional de asistencia a la escuela. Y un programa de provisión de alimentos en las escuelas cuesta 36 dólares por año adicional inducido de asistencia a la escuela. Por lo tanto no todos los programas son iguales. Hay programas que son muy efectivos, por ejemplo el primer programa de eliminación de parásitos intestinales, y hay programas que no son muy efectivos para aumentar la asistencia de niños en países subdesarrollados a las escuelas.

Por lo tanto, ninguno de los dos supuestos en los que se basa la teoría del paracaídas es correcto en el análisis de la ayuda al desarrollo: la ayuda puede tener efectos secundarios negativos, o inesperados, y no todas las actuaciones son igualmente efectivas. Por eso es tan importante realizar una evaluación precisa y basada en métodos científicos basados en experimentos. Duflo, Kremer y Glennerster (2006) ofrecen una guía práctica de cómo aplicar la metodología experimental al estudio del impacto de programas y proyectos de desarrollo económico.

¿Da lo mismo el método con el que evaluemos? No. Como ejemplo podemos utilizar un trabajo reciente sobre la efectividad de los *flip charts* en el aprendizaje de los estudiantes en Kenia (Glewwe et al. 2004). La selección de este caso no es casual: para este ejemplo disponemos de datos experimentales y de datos observacionales. Por tanto, podemos usar los datos como un experimento o interpretarlos en clave de una evaluación regular usando regresiones y supuestos estructurales.

Antes de comentar este caso, merece la pena explicar el concepto de *flip charts*. Un *flip chart* es una serie de láminas que contienen explicacio-

nes sobre materias, matemáticas por ejemplo, y que se cuelgan en las paredes de las escuelas. Normalmente son doce láminas enlazadas con una espiral. Se utilizan como ayuda visual para el aprendizaje de niños en las escuelas de países subdesarrollados donde la mayoría de los alumnos no tiene acceso a los libros.

Por lo tanto, la cuestión que queremos analizar es la siguiente: ¿mejoran los *flip charts* el aprendizaje de los estudiantes en las escuelas rurales de Kenia? En principio los *flip charts* parecen instrumentos adecuados para mejorar el aprendizaje en el contexto de las escuelas rurales de Kenia. Tres son los motivos fundamentales. En primer lugar, los libros de texto en las escuelas rurales de Kenia son muy escasos, y además están escritos en inglés que es el medio de distribución habitual. Muchos estudiantes tienen un nivel bajo de inglés y, por tanto, los libros tienen una capacidad limitada como instrumento de aprendizaje. En segundo lugar, los *flip charts* pueden promover el aprendizaje porque los estudiantes recuerdan gráficos y fotografías más a menudo que haber leído palabras. En tercer lugar, los estilos de aprendizaje varían mucho entre estudiantes y añadir ayudas visuales aumenta el espectro de estudiantes a los que se les puede enseñar efectivamente. Finalmente, en general, las clases y presentaciones se benefician de tener ayudas visuales suplementarias. Por lo tanto, en principio, usar *flip charts* es una buena idea.

Pero ¿es realmente efectivo para el aprendizaje de los estudiantes? Para contestar a esta pregunta podemos utilizar, en este caso, dos tipos de datos: observacionales y experimentales. ¿Cómo se analizan los resultados a partir de datos observacionales? El estudio se basa simplemente en los datos que recopilamos y que provienen de estadísticas accesibles y ya producidas. No existía un diseño específico para facilitar el análisis posterior de los datos. Existía información disponible sobre unos cien colegios a partir de un estudio anterior sobre el efecto de la provisión de libros de textos y becas en escuelas de Kenia. También se recogió mucha información sobre materiales docentes en las escuelas y entre ellos la disponibilidad de *flip chart* en estas cien escuelas. Se disponía de información sobre el número de *flip charts* que cada escuela tenía en ciencias, matemáticas, salud y negocios. Por lo tanto, estos datos provienen de una muestra anterior al estudio que analizamos y que no fue aleatoriza-

da, en el sentido experimental del término. Por eso, los datos son observacionales.

Utilizando estos datos observacionales de escuelas, los estudiantes en clases donde se utilizaban los *flip charts*, o tenían mayor número de ellos, aumentaban un 20% de una desviación estándar sus resultados escolares. Aumentar un 10% de una desviación estándar es como pasar del percentil 50 al percentil 54. Por lo tanto, a juzgar por estos datos observacionales, los *flip charts* eran muy efectivos. El coste por alumno de los tres *sets*¹⁰ es sólo un 10% del coste que tienen los libros en las tres asignaturas cubiertas por los *flip charts*,¹¹ pero su impacto es dos veces más grande que el efecto de proveer libros de texto. Por lo tanto, los *flip charts* serían 20 veces más efectivos, en términos de coste, que los libros de texto, según este análisis con datos observacionales.

¿Qué pasa cuando hacemos un análisis experimental? ¿Qué pasa si hacemos un grupo de control con escuelas a las que no damos *flip charts* y un grupo tratado de escuelas con *flip charts*, y las comparamos? Esto es lo que hicieron Glewwe et al. (2004) con la ayuda de ICS (International Christelijk Steunfonds), una ONG holandesa. La intervención consistía en proporcionar cuatro *flip charts* (dos sobre ciencias, uno sobre salud y uno sobre matemáticas) y un mapa del este de África para las clases de geografía. El diseño experimental consistía en la selección de 178 escuelas en Busia y Teso (dos zonas rurales de Kenia) por la ONG, que financiaba los *flip charts*. Las escuelas elegidas tenían una media de aprendizaje de los estudiantes similar a la media del distrito en su conjunto. Ochenta y nueve escuelas fueron asignadas al grupo de control y otras 89 escuelas, al grupo tratado. Después de ser ordenadas geográficamente, y alfabéticamente dentro de cada área, se utilizó un diseño muestral sistemático para asignar las escuelas a los dos grupos. Los resultados del experimento aparecen en el cuadro 1.4. La diferencia en el porcentaje de respuestas correctas entre

¹⁰ El coste de un set de *flip charts* es de 80 dólares.

¹¹ Un libro de texto cuesta aproximadamente 3,3 dólares. Si multiplicamos el coste de cada libro por el número de estudiantes medio de cada clase (80) y el número de materias (3) obtenemos un resultado de 800 dólares.

CUADRO 1.4: Media del resultado de los exámenes
(porcentajes de respuestas correctas)

Fecha	Curso	Tratados	Control	Diferencia	Desviación estándar
Julio 1997	8	45,5	46,0	-0,5	12,5
Noviembre 1997	8	48,7	49,6	-0,9	13,3
Julio 1997	8	42,7	42,9	-0,3	11,2
Noviembre 1997	8	49,5	49,5	0,0	13,0
Octubre 1998	7	37,6	37,5	0,1	11,3
Octubre 1998	6	37,3	36,9	0,4	11,4

Fuente: Glewwe et al. (2004).

los alumnos de escuelas que usaron *flip charts* y las escuelas que no tenían es pequeña y estadísticamente no significativa.

El resultado, por tanto, es el contrario al hallado al usar los datos observacionales. Esta diferencia es la análoga a la descubierta por LaLonde para el caso de los cursos de formación. Cuando utilizamos datos observacionales los *flip charts* parecen eficaces pero cuando se realiza un experimento aleatorizado el resultado es claro: no se puede encontrar un efecto estadísticamente significativo sobre el aprendizaje de los estudiantes.

Los estudios disponibles indican que hay algunos tipos de actuaciones y proyectos que funcionan. Y esto se ha probado experimentalmente. Para favorecer la presencia de alumnos en las escuelas hay varias medidas eficientes (en términos coste-beneficio): las medicinas para acabar con los parásitos de los niños, los suplementos alimenticios como el hierro y la vitamina A. En otros ámbitos también sabemos que funcionan los subsidios a fertilizantes, la vacunación de los niños, la provisión de agua potable, etc. Estas intervenciones son modestas frente a las grandes promesas y el esquema utópico,¹² pero pueden tener potencialmente una incidencia enorme sobre el bienestar de las familias pobres.

¿Qué no funciona? Pues los *flip charts* no funcionan experimentalmente. Comprar libros y distribuirlos gratuitamente tampoco mejora el rendimiento escolar. Los únicos que se benefician y mejoran el aprendizaje son

¹² Por ejemplo, duplicar la ayuda al desarrollo en línea con las propuestas de J. Sachs.

aquellos que ya saben mucho, la parte alta de la distribución. El resto no se beneficiaba, quizás por los problemas señalados anteriormente (el idioma no es el que hablan o dominan muchos de ellos, etc.). Y los incentivos económicos para los profesores tampoco funcionan, porque al final los docentes, para conseguir los incentivos, hacen que los estudiantes aprendan a hacer test. Como al final los incentivos son en función de los resultados de los estudiantes en test de respuestas múltiples, los profesores tienden a preparar mucho la realización del test. Esto no significa necesariamente que los estudiantes aprendan más.

1.5. Conclusiones

En resumen, lo que llamamos evaluación científica (experimental) de la ayuda al desarrollo implica ser capaces de arrebatarse a la política el protagonismo en ámbitos donde no debería tenerlo, en ámbitos donde las decisiones técnicas deben predominar sobre los criterios políticos. Los programas y actuaciones para reducir la pobreza en los países en vías de desarrollo pueden funcionar o fallar. Pero incluso si fallan se puede aprender de su análisis. También se puede aprender de un programa que no sea coste-efectivo. Intentar tapar los fallos o evitar la evaluación científica de los resultados, utilizando procedimientos que no son adecuados, hace un mal servicio a la comunidad científica y a la comunidad de la cooperación internacional en general.

En este capítulo se ha insistido en que la evaluación experimental puede dar resultados muy diferentes a los obtenidos por los procedimientos tradicionales y que la evaluación aleatorizada es factible para analizar la ayuda al desarrollo. También se ha señalado que este tipo de evaluación precisa recursos humanos y económicos que no son muy superiores a los necesarios para realizar otros tipos de evaluación de estilo observacional. La colaboración entre diversos tipos de especialistas es fundamental para la realización satisfactoria y eficiente del proceso. Un esquema que ha funcionado en muchos casos es el siguiente: los investigadores diseñan el experimento y los protocolos de aplicación; las ONG colaboran en la recogida de información y el trabajo de campo (en ocasiones también podrían

colaborar en la financiación). Por último, los organismos internacionales financian la evaluación, dado que, al final, se está produciendo un bien público internacional. Muchos de los experimentos recogidos en este capítulo han sido posibles gracias a la cooperación con organizaciones no gubernamentales.

Las agencias internacionales y las ONG deben experimentar y buscar intervenciones que funcionen, verificando su eficacia por medio de las mejores técnicas de evaluación científicas disponibles y usando evaluadores externos para evitar los intereses particulares de los gestores de los proyectos.

Bibliografía

- ANDERSON, M. *Do no Harm: How Aid Can Support Peace – or War*. Londres: Rienner, 1999.
- ANGRIST, J. «American education research changes tack». *Oxford Review of Economic Policy* 20 (2), 2004.
- BANERJEE, A., A. DEATON, N. LUSTIG, y K. ROGOFF. *An evaluation of the World Bank Research: 1998-2005*. Washington: World Bank, 2006 [mimeo.].
- BOONE, Peter. «Politics and the Effectiveness of Foreign Aid». *European Economic Review* 40, 1996: 289-329.
- BURNSIDE C., y D. DOLLAR. «Aid, Policies and Growth». *American Economic Review*, 90 (4) septiembre 2000: 847-88.
- DJANKOV, S., J. G. MONTALVO, y M. REYNAL-QUEROL. «Does Foreign Aid Help?». *Cato Journal* 26 (1), 2006a: 1-28.
- . *Aid with multiple personalities*, 2006b [mimeo.].
- . «The curse of aid». *Journal of Economic Growth* 13 (3), septiembre 2008: 169-194.
- DUFLO, E., M. KREMER, y R. GLENNERSTER. «Using Randomization in Development Economics Research: a Toolkit». NBER Technical Working Paper. *Handbook of Development Economics*, vol. 4, Editors T. Paul Shultz, 2006.
- EASTERLY, William. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done so Much ill and so Little Good*. Nueva York: The Penguin Press, 2006.
- EASTERLY, William, ROSS LEVINE, y David ROODMAN. «Aid, Policies and Growth: a Comment». *American Economic Review* 94 (3), 2003: 774-780.
- GIBSON, Clark, KRISTER ANDERSSON, Elinor OSTROM, y Sujai SHIVAKUMAR. *The Samaritan's Dilemma: the Political Economy of Development Aid*. Oxford: Oxford University Press, 2005.

- GLEWWE, P., M. KREMER, S. MOULIN, y E. ZITZEWITZ. «Retrospective Versus Prospective Analysis of School Inputs: the Case of Flip Charts in Kenya». *Journal of Development Economics* 74, 2004: 251-268.
- LALONDE, R. «Evaluating the Econometric Evaluation of Training Programs with Experimental Data». *American Economic Review* 76 (4), 1986: 604-620.
- MAREN, M. *The Road to Hell: the Ravaging Effects of Foreign Aid and International Charity*. Nueva York: Free Press, 1997.
- MONTALVO, J. G. «The Political Economy of International Aid». De próxima aparición en la *Princeton Encyclopedia of the World Economy*.
- MONTALVO, J. G., y M. REYNAL-QUEROL. «Ethic Polarization, Potential Conflict and Civil War». *American Economic Review* 95 (3) (2005): 796-816.
- . «Discrete Polarization with an Application to the Determinants of Genocides». *Economic Journal*, noviembre 2008: 1-33.
- OLKEN, B. «Corruption and the Costs of Redistribution: Micro Evidence from Indonesia». *Journal of Public Economics* 90 (4-5), 2006: 853-870.
- RAJAN, R., y A. SUBRAMANIAN. «Aid, Dutch Disease and Manufacturing Growth». IMF Working Paper, 2006.
- . «Aid and Growth: What Does the Cross Country Evidence Really Show?». IMF Documento de Trabajo 127, 2007.
- RAVALLION, M., y S. CHEN. «Hidden Impact: Household Savings in Response to a Poor-area Development Project». *Journal of Public Economics* 89 (11-12), 2005: 2182-2204.
- REINIKKA, Ritva, y Jakob SVENSSON. «Local Capture: Evidence From a Central Government Transfer Program in Uganda». *Quarterly Journal of Economics* 119 (2), 2004: 679-705.
- . «Fighting Corruption to Improve Schooling: Evidence from a newspaper Campaign in Uganda». *Journal of the European Economic Association* 3, 2005: 259-267.

Cambios en los hábitos y los sistemas sanitarios: algunas evidencias a partir de evaluaciones aleatorizadas

Michael Kremer

Catedrático de Economía, Universidad de Harvard

2.1. Introducción

En este capítulo se analizan las intervenciones sanitarias en países en vías de desarrollo y su evaluación. La efectividad de dichas intervenciones depende de los comportamientos de los individuos, así como de la organización del sistema sanitario en su conjunto. Por este motivo en el capítulo se analizan ambas perspectivas.

El comportamiento médico y sanitario es muy importante en los países en vías de desarrollo. Si examinamos cómo han mejorado los patrones de comportamiento sanitario en los países ricos, el inmenso aumento de la esperanza de vida en los países desarrollados tiene que ver con las mejoras de los sistemas sanitarios, de la renta, de la nutrición, etc. Pero hoy día en los países pobres la situación es muy distinta. Debido al desarrollo de los antibióticos y otros avances médicos, en la actualidad puede ocurrir que un país sea muy pobre y, sin embargo, sus habitantes tengan una esperanza de vida bastante alta. Incluso mucho más alta que la esperanza de vida

de 1900 en los Estados Unidos. Un ejemplo sería Vietnam. La renta per cápita en Vietnam es actualmente de una décima parte de la renta de los habitantes de Estados Unidos en 1900. En el año 1900, la esperanza de vida en los Estados Unidos era de 47 años. Hoy, en Vietnam, la esperanza de vida alcanza los 69 años. Esto indica que es posible tener un nivel de sanidad bastante alto incluso en un país muy pobre.

Desgraciadamente, no todos los países pobres o en vías de desarrollo muestran las estadísticas de esperanza de vida de Vietnam. En los países desarrollados la esperanza de vida es de 79 años. En el conjunto del sur de Asia es de 63 años. Sin embargo, en el África subsahariana la esperanza de vida sólo alcanza los 46 años. Esto supone que hay una gran disparidad en los niveles de esperanza de vida entre países. Esta disparidad tiene que ver con el comportamiento de los individuos frente a enfermedades como el sida, que es prevenible pero que tiene un efecto epidémico en muchos países. Todo depende de las prácticas sexuales de la gente. Otras intervenciones, como la vacunación generalizada o la terapia de rehidratación, también pueden tener un impacto muy importante sobre las tasas de mortalidad. Es cierto que hay muchas campañas de vacunación en marcha en países del tercer mundo, pero todavía existen grandes segmentos de la población de estos países que no son vacunados, a pesar del enorme impacto que tienen dichas intervenciones sobre la esperanza de vida de los ciudadanos.

Otra parte importante del efecto del comportamiento individual sobre la salud y la esperanza de vida de la gente está en los sistemas sanitarios. Los sistemas sanitarios en muchos países pobres funcionan mal tanto si son públicos como si son privados. En los sistemas públicos hay problemas de incentivos muy importantes. Se realizó hace algún tiempo un estudio que examinó las clínicas de asistencia primaria. El objetivo era comprobar simplemente si el personal sanitario acudía a sus puestos de trabajo. Los países considerados eran la India, Indonesia, Uganda, Bangladesh y los de Hispanoamérica. Los resultados del estudio fueron demoledores: en promedio, un 35% del personal sanitario que tenía que estar en las clínicas no estaba presente siquiera en su puesto de trabajo. Y entre los médicos el porcentaje de absentismo era todavía mayor. A veces estaban nominalmente en la clínica pero no se sabía dónde. Y esto es aplicable a muchas

clínicas privadas también. Muchas personas en estos países usan sistemas privados porque el sistema público es catastrófico, pero los sistemas privados también funcionan muy mal. En este caso no sólo es que el personal sanitario no tiene los incentivos correctamente establecidos sino que, en muchas ocasiones, no persiguen siquiera los intereses de los pacientes. Por ejemplo, es muy común dar vacunas o poner inyecciones cuando un paciente acude a estas clínicas. Un estudio en la India demostró que al 70% de los pacientes se le ponían inyecciones pero no porque las necesitaran sino porque se consideraba que era lo que el paciente-cliente esperaba. Otro ejemplo sería la aplicación de goteros de glucosa intravenosa. La administración intravenosa de glucosa produce un efecto inmediato de bienestar, especialmente en pacientes que tienen déficits nutricionales, pero en el largo plazo no sirve para nada. Lo mismo sucede con las inyecciones de esteroides. De hecho, la administración continuada de esteroides puede tener efectos contraproducentes a largo plazo.

Este tipo de problemas se presenta tanto en sistemas públicos como privados. Un estudio reciente muestra que entre el 30 y el 50% de las recetas prescritas en la India son incorrectas para el diagnóstico realizado. El problema es que una parte del personal sanitario de los países pobres no cuida los intereses de los pacientes, ni sirve realmente a la comunidad. En los países pobres hay muchas enfermedades infecciosas. La administración de vacunas beneficiaría no sólo a los pacientes individuales, sino también la comunidad por los efectos externos de las mismas. Sin embargo, la administración de vacunas no se produce en muchas ocasiones por el absentismo de los trabajadores sanitarios. Las madres, que tienen que recorrer en ocasiones largas distancias para acudir a un centro sanitario donde vacunen a sus hijos pequeños, en muchas ocasiones desisten del viaje ante la elevada probabilidad de que una vez hayan llegado al centro de atención primaria no exista personal sanitario para administrar las vacunas. Otras veces los fármacos, especialmente los antibióticos, se administran en dosis que no son correctas, de forma que se genera un problema de resistencia. Por tanto, una actuación que debería resolver el problema de salud de un paciente acaba generando una externalidad negativa (resistencia a los antibióticos) en el conjunto de la ciudadanía. En consecuencia, para mejorar la efectividad de los sistemas sanitarios en los países

pobres es necesario cambiar el comportamiento individual de los ciudadanos respecto a muchas actividades con riesgo para la salud, pero también es necesario mejorar el funcionamiento del sistema sanitario en general. Por desgracia en estas cuestiones la gente tiene opiniones muy contundentes, que muchas veces tienen un importante sustrato ideológico.

Como se señala en el primer capítulo, las propuestas para mejorar las condiciones de vida en los países en vías de desarrollo evolucionan en ciclos. Al principio las nuevas ideas generan mucho entusiasmo. Este entusiasmo conduce a la inversión de grandes esfuerzos. Pero si no se ven resultados comienza un ciclo de desilusión. Y así sucesivamente. Yo creo que eso en parte tiene que ver con que las ideas inicialmente se basan en la evidencia anecdótica. No tienen realmente un respaldo científico. Por este motivo es necesario realizar ensayos aleatorizados. La medicina hace años era muy similar a la economía del desarrollo actual. Había muchas ideas, muchos médicos que afirmaban que tenían nuevas curas para enfermedades. La gente se creía muchas de las cosas que se decían, pero luego, a la hora de la verdad, había una gran desilusión, pues los avances no eran tales. Quién no recuerda los famosos vendedores de crecepelo del antiguo oeste americano, que vendían unas pócimas que lo curaban todo. La medicina progresó gracias al avance científico y, evidentemente, gran parte de dicho avance científico consistió en establecer un soporte metodológico adecuado para determinar qué funcionaba y qué no funcionaba. La aplicación del método experimental aleatorizado fue fundamental para el avance de la medicina.

Este tipo de ensayo experimental podría desempeñar un papel similar en la economía del desarrollo actual al papel que desempeñó en la medicina hace ya muchos años. Facilitaría la generación de evidencia empírica, y aumentaría la transparencia y la credibilidad de los resultados. El resto del capítulo se estructura en torno a ejemplos de evaluación randomizada para el análisis de actuaciones para cambiar los incentivos y, por tanto, el comportamiento individual y el funcionamiento del sistema en su conjunto. El primer epígrafe trata sobre los incentivos para mejorar los hábitos sanitarios de los individuos. Se incluyen varios estudios. En primer lugar, se considera el efecto de suministrar medicinas antiparásitos a los niños en edad escolar. El objetivo de ese estudio era evaluar su impacto no sólo en

términos sanitarios, efectividad del fármaco, sino también sobre la asistencia a la escuela de los niños que reciben dicho medicamento, y su nivel de aprendizaje. En segundo lugar se analiza el programa Progresía de México en su vertiente de transferencias monetarias a familias pobres, condicionadas a realizar visitas al médico para la prevención sanitaria. En tercer lugar, se consideran los programas para la incrementar las tasas de vacunación. En cuarto lugar, se presentan resultados de la formación sobre la transmisión del sida en las tasas de contagio. El tercer epígrafe del capítulo analiza los sistemas sanitarios en general, y el papel que pueden tener las condiciones de los contratos y los sistemas externalizados en este aspecto. Finalmente, el cuarto epígrafe resume las conclusiones del capítulo.

2.2. El cambio de los hábitos sanitarios y sexuales

Uno de los elementos básicos de la mejora de la salud de los ciudadanos en los países pobres consiste en la modificación de los hábitos sanitarios de sus habitantes. Este cambio debe ser inducido a partir de sistemas de incentivos eficientes.

2.2.1. Los efectos de la desparasitación infantil sobre la asistencia a la escuela

En el mundo hay más de mil millones de personas que tienen lombrices intestinales. Hay sistemas muy baratos y muy seguros para tratar este tipo de problemas pero, a pesar de ello, el número de infectados no disminuye. Tenemos la tecnología, el coste del tratamiento es pequeño y, sin embargo, no estamos avanzando en un problema tan importante. ¿Por qué? Una de las dificultades es que aunque a veces algunas de estas enfermedades, como tener lombrices, pueden tener consecuencias graves para la salud, por lo general sólo generan un poco de anemia. Ésta no es una condición realmente aguda, sino más bien crónica. Produce una falta de energía crónica y sensación de debilidad. Pero no es algo que sea realmente muy grave. En la mayoría de los casos la gente no va al médico, sobre todo teniendo en cuenta que es muy probable que el médico esté ausente después de hacer un largo viaje hasta el centro de atención prima-

ria. El problema de la distribución de drogas antiparasitarias basada en clínicas está asociado con la distancia existente hasta el centro sanitario más cercano, la probable ausencia del personal sanitario del mismo y el coste del diagnóstico que, aunque pequeño, no puede ser asumido por muchas familias con rentas bajas. Por tanto, hay que pensar en instituciones alternativas para la distribución de estas drogas antiparasitarias.

Una de estas formas alternativas de distribuir los medicamentos contra las lombrices intestinales es la administración masiva de los mismos en las escuelas. Hay zonas donde el problema de las lombrices intestinales es endémico y áreas donde el 90% de los niños está afectado. Lo que está diciendo la Organización Mundial de la Salud es que si más de la mitad de los niños están afectados, lo mejor es tratar a todos los niños en el colegio con independencia de que estén o no afectados. Con mi colaborador Edward Miguel (Kremer y Miguel 2004), y la ayuda de una ONG, evaluamos una actuación en Kenia donde se introdujo un programa de desparasitación en 75 colegios. Además de estudiar el impacto sobre el nivel de infección parasitaria en los niños, también se consideró su efecto sobre la educación de los estudiantes. La aleatorización se realizó utilizando el orden alfabético. Los resultados mostraron una reducción significativa de la incidencia de parásitos. Los niños en el grupo tratado crecían más, eran más altos, y la autoevaluación de la salud por parte de los mismos era más positiva. Pero quizás el impacto más importante era inicialmente inesperado: el absentismo escolar se redujo un 25% en el grupo tratado frente al grupo de control. Este resultado es realmente impactante. Al iniciar el experimento faltaban muchos niños al colegio con regularidad. En el grupo tratado la desparasitación condujo a que los alumnos empezaron a asistir al colegio con regularidad. El coste por año extra de escolarización de esta intervención era de 3,5 dólares. Comparando con la administración gratuita de comidas escolares la desparasitación es mucho más eficiente.

Yo he trabajado con asiduidad en educación en países del tercer mundo. En Kenia para ir a la escuela se pedía un informe, que era costoso para unas familias con un nivel de renta bajo. Las ONG y yo mismo habíamos propuesto comprar uniformes y distribuirlos gratuitamente, facilitar comidas gratis en el comedor para los estudiantes, etc., con el objetivo de aumentar la asistencia a clase. Sin embargo, ninguna de estas medidas, todas

razonables, tenía la efectividad que ha demostrado el programa de desparasitación. Y eso que era un tratamiento contra parásitos y no directamente una intervención educativa. No era lo que nadie habría incluido como el primer elemento en una lista de medidas para reducir el absentismo escolar en los países pobres. Otro factor muy importante, que vale la pena considerar, es la relación coste-efectividad de la medida. Es difícil pensar que nadie hubiera apostado por que la desparasitación tuviera una relación coste-efectividad tan elevada, mucho más que todos los otros procedimientos utilizados para incentivar la asistencia al colegio.

Visto el éxito inicial, la actuación se amplió de uno a cuatro distritos. Empresarios japoneses patrocinaron la extensión del programa. El siguiente paso fue ampliar su aplicación durante cinco años para intentar extenderlo a una escala incluso superior, más general. Pero, en todas estas intervenciones, una pregunta clave es la siguiente: lo que ha funcionado a pequeña escala ¿funcionará si lo generalizamos y lo hacemos a gran escala? Este problema es conocido en la literatura como la cuestión del *scaling up*. La decisión de asistir a la escuela es compleja, pues depende de muchos factores. No siempre los mismos factores afectan de la misma forma y en todas partes. Es importante, por tanto, analizar el contexto en el que se producen los resultados. En este caso, el mismo tipo de intervención (suministro de medicinas antiparasitarias) se realizó en zonas urbanas de la India. Los resultados fueron muy similares a los encontrados en el caso de Kenia. También hay datos del sur de Estados Unidos, aunque en este caso no se trata de una evaluación experimental pues se refiere a la situación de hace más de cincuenta años. En los años veinte en el sur de Estados Unidos todavía existía un problema importante de infección de parásitos entre los niños. La Fundación Rockefeller decidió financiar programas para acabar con el problema. Bleakley (2007) muestra que el impacto de esta intervención en los estados del sur de Estados Unidos fue muy semejante en cuanto a la asistencia de los niños y la contribución a la educación, al impacto observado en Kenia y la India. Tres lugares diferentes, un impacto semejante. No es el caso en todos los programas pero merece la pena destacar que para esta intervención se observa una gran similitud entre todas las estimaciones de los efectos. Cuando varias evaluaciones en diferentes países y condiciones

señalan en la misma dirección, esta evidencia apunta a que se trata de un programa que funciona.

Hay un elemento adicional que debemos considerar en este programa. El gran impacto del programa de desparasitación no sólo se produjo por su efecto sobre los niños que recibieron el fármaco sino que, debido a la reducción de la tasa de transmisión de la enfermedad, bajó también la tasa de infección en las escuelas vecinas que no había entrado en el grupo tratado. Por tanto, se produjeron importantes externalidades positivas.

2.2.2. Educación sanitaria y copago

Pero vayamos más allá del aspecto estrictamente médico. Hay mucho debate en cuanto a la relación entre desarrollo y salud. La cuestión es cómo podemos hacer que los programas sean duraderos y sostenibles, y que el dinero se utilice de forma eficaz y efectiva. ¿Hay alguna forma de conseguir estos objetivos? La educación podría tener un efecto importante: enseñarle a la gente sobre los parásitos, cómo no infectarse, la importancia de lavarse las manos, de utilizar calzado adecuado, de no nadar en aguas contaminadas. Por lo tanto, se podría utilizar la educación para informar sobre las medidas de prevención como alternativa a usar medicinas, para aumentar la asistencia de los estudiantes a las clases.

Sin embargo, a pesar de los mensajes los niños no se lavaban las manos e iban a nadar a ríos y lagos infectados. La información no cambiaba su comportamiento. Por otra parte, el calzado es costoso y estamos hablando de zonas muy pobres. Por tanto, tampoco se consiguió un efecto sobre la proporción de niños que usaban zapatos. En fin, la educación sanitaria resultó muy poco efectiva.

¿Hay alguna otra forma de hacer un programa educativo sostenible a largo plazo? Un programa se considera sostenible si cubre costes. Kremer y Miguel (2007) muestran que el programa de desparasitación no se podía mantener sin una financiación externa durante todo el tiempo. La educación para la salud no era sostenible: cuando hay que compartir costes no se encuentra diferencia en familias con diferentes niveles. Cuando el usuario tiene que financiar una parte del programa se observa una clara reducción en la participación. A pesar de ser una cantidad muy pequeña, el copago produce una enorme reducción en la participación. Por este motivo el artículo

de Kremer y Miguel (2007) se titula «La ilusión de la sostenibilidad». En concreto, el trabajo amplía el estudio sobre desparasitación en Kenia para analizar si sería sostenible en el largo plazo, basado en la acción de la comunidad y no de la ayuda internacional. La desparasitación es un bien público, dado que, como se señaló con anterioridad, gran parte de su beneficio social se produce a partir de la reducción en la tasa de transmisión de la infección. Por tanto, los individuos eran menos propensos a tomar la medicación si sus contactos sociales estaban tomándola. Los esfuerzos para reemplazar los subsidios con medidas sostenibles de control de la infección parasitaria fracasaron: el programa de copago (coste-recuperación) redujo la toma de la medicación en un 80%; la educación sanitaria no afectó en nada al comportamiento y las intervenciones de movilización social fracasaron igualmente. Una vez se estableció un pequeño pago por las medicinas las variaciones del mismo tuvieron muy poco efecto en la toma de los medicamentos. Además, debido a la enorme caída en la participación, los ingresos obtenidos fueron muy pequeños. Por tanto, parece claro que es poco realista esperar que una intervención puntual pueda generar una provisión voluntaria de bienes públicos locales que sea sostenible en el largo plazo.

2.2.3. Los subsidios para mejorar la salud: el programa Progresas mexicano

Si un programa puntual no es sostenible, ¿pueden los subsidios servir para mejorar la participación en programas sanitarios en el largo plazo? El gobierno mexicano puso en marcha en 1998 el programa Progresas.¹ Este programa de lucha contra la pobreza combinaba un sistema tradicional de transferencia condicionada de efectivo con incentivos financieros para que las familias invirtieran en capital humano (educación, salud y nutrición). Se trataba de transferir dinero en efectivo a las madres de familias pobres en zonas rurales (sólo las familias por debajo un determinado umbral de renta podían ser beneficiarias). Pero para recibir ese dinero tenían que hacer dos cosas: primero, los niños tenían que ir al colegio² y, en

¹ Posteriormente el programa ha pasado a llamarse Oportunidades.

² Los profesores deberían confirmar que los niños habían acudido a clase al menos el 85% de los días lectivos.

segundo lugar, tenían que participar en las medidas de promoción de la salud: cada miembro de la familia debía aceptar servicios sanitarios preventivos; los niños de entre 0 y 5 años y las madres que daban el pecho debían acudir a clínicas donde se hacía un seguimiento de su crecimiento, obtenían suplementos nutricionales y recibían educación sobre nutrición e higiene; y, finalmente, las mujeres embarazadas debían visitar los centros de atención primaria para recibir atención médica durante el embarazo, así como suplementos nutricionales y educación sanitaria. La gran ventaja del Progresas, al igual que su sucesor Oportunidades, es que su diseño se realizó de manera que, con posterioridad, pudiera ser analizado como un experimento aleatorizado. El programa se puso en marcha inicialmente en un grupo de 495 localidades seleccionadas aleatoriamente de entre las 50.000 localidades Progresas. De entre éstas, 314 fueron seleccionadas aleatoriamente para recibir las actividades del programa durante los dos primeros años y las restantes 181 sirvieron de control y recibieron el programa en el tercer año (2000). El gobierno mexicano decidió utilizar expertos internacionales para evaluar el Progresas, en parte porque los que iniciaron el programa comprendían la situación política mexicana: cada vez que se elegía un nuevo presidente se cerraban programas y se abrían nuevos para atribuirse su creación y señalar que no querían continuar con lo anterior. La continuidad de los programas es crucial para sus efectos a largo plazo y para una evaluación adecuada. Una evaluación realizada por expertos internacionales podría garantizar, como así sucedió, la continuidad del programa.

¿Cuál fue el impacto del Progresas? Dadas las condiciones para el acceso a las subvenciones, es lógico esperar dos tipos de efectos: sobre el nivel educativo de los niños y sobre la salud, en general, de las familias. Gertler (2004) muestra que en el grupo tratado (los que recibían la subvención) la tasa de enfermedad de los niños era un 39,5% inferior a la del grupo de control. Además, los niños tratados crecieron un centímetro más que los del grupo de control, condicionado a la edad. Finalmente, las transferencias condicionadas reducían las tasas de anemia un 25,5%. Schultz (2004) muestra cómo el Progresas incrementó, en 0,66 años el número de años de educación del grupo tratado (sobre una media de 6,8 años).

En gran parte como resultado de estas evaluaciones, cuando el liderazgo político cambió en México, en lugar de cerrar el programa le cambiaron el nombre (Oportunidades). No obstante, el nuevo programa mantiene las características básicas del Progreso. Además, el éxito de la primera fase favoreció la ampliación del programa desde zonas rurales hacia zonas urbanas de México. La gran influencia de los resultados de la evaluación del Progreso hizo que programas similares se extendieran a otros países latinoamericanos, como por ejemplo Brasil, y otros países centroamericanos lo han adaptado. Y lo han hecho gobiernos nacionales, pero también lo han adoptado organizaciones que trabajan con dinero a partir de donantes internacionales. Éste es uno de los grandes beneficios de realizar una evaluación científica y creíble de los programas sociales: que va más allá de un proyecto en particular, y puede ser un beneficio para el país en su conjunto e incluso para el conjunto del mundo. De esta forma las evaluaciones aleatorizadas se convierten en un bien público global. Esto vuelve a poner sobre la mesa el tema de la financiación de las subvenciones. ¿Debemos pedirle a las ONG, al país o a los organismos internacionales que paguen por la evaluación? Si se produce un bien público global, sería lógico que la ayuda extranjera y de las organizaciones internacionales participaran decisivamente en la financiación y la realización de evaluaciones independientes.

2.2.4. El efecto de las campañas masivas de vacunación

Otro ejemplo de subvención es el que puede asociar como condición la vacunación. La vacunación salva tres millones de vidas al año pero podrían salvarse seis millones. Unos tres millones de muertes podrían evitarse si más personas fuesen vacunadas. Un experimento interesante sobre el efecto de los subsidios en la vacunación se está realizando en Rajastán (la India). A menudo cuando van los ciudadanos al centro de atención primaria lo encuentran cerrado, lo que implica que no se pueden vacunar. El experimento consiste en dos tipos de tratamiento. El primer tratamiento es garantizar un *stock* mensual regular y sistemático de acceso a vacunas, con información a la población sobre su estabilidad y regularidad. El segundo tratamiento es hacer lo mismo pero, cuando el paciente va a la clínica, se le da además un kilo de lentejas. Los resultados preliminares del

experimento son muy esperanzadores: el primer tratamiento producía un aumento del 22% en la vacunación frente al grupo de control. En el grupo sujeto al segundo tratamiento (un incentivo en términos de un kilo de lentejas) la vacunación era un 40% superior al objetivo.

2.2.5. Campañas educativas, comportamiento sexual y sida

Cuando pensamos en vacunaciones y tratamientos, incluso en medicina prenatal, se puede pensar en fármacos subvencionados y no tanto en cambios de las conductas. Sin embargo, cuando se trata del comportamiento sexual el problema es más complejo pues es difícil recompensar el sexo seguro. A menos que pusiéramos cámaras en todas las casas sería imposible su observación. Por tanto, en cuanto al sida, obtener buenos resultados es más difícil. Una de las razones de la escasa eficacia de estos programas es que es relativamente fácil tratar los parásitos intestinales, mientras que el tratamiento de la malaria y del sida es más complejo y costoso (los retrovirales son más caros y su administración requiere constancia en el tiempo). Pero, en el caso del sida, la educación es una buena opción para detener su expansión, pues los medios de contagio son bien conocidos.

Los niños se ven como una ventana de oportunidad en la lucha contra el VIH y el sida y la escuela es una forma de contactarlas. Podemos ir a las escuelas y transmitirles la información. Sin embargo, la eficacia de la escuela como mecanismo para proporcionar educación sobre el sida es desconocida aunque, a priori, es una buena oportunidad de llegar a los niños que no están todavía infectados. También hay una gran controversia en cuanto a los mensajes que se deben emitir. Quizás no tanto en España o en Europa, pero en Estados Unidos hay un debate importante sobre la abstinencia sexual frente al uso del preservativo. En Kenia éste es un tema de gran controversia pues hay una oposición religiosa al uso de los preservativos. En los siguientes párrafos se analiza el impacto de cuatro enfoques diferentes utilizados en Kenia para reducir las tasas de transmisión del sida siguiendo los resultados de Duflo et al. (2006). Los programas estudiados son la educación de profesores en temas relativos al sida y su transmisión; la utilización de debates y escritura de ensayos por parte de los niños; la información sobre las tasas de infección por edades y la reducción del coste de la educación en general.

En primer lugar, analizaremos el efecto de los programas de preparación del profesorado en cuestiones relacionadas con la transmisión del sida. El gobierno de Kenia, con la ayuda de UNICEF, comenzó un programa consultivo de larga duración con las organizaciones de la sociedad civil en Kenia y la Iglesia, desarrollando un currículo específico de educación sobre la transmisión del virus del sida. Aunque algunos profesores explicaban esta asignatura otros no lo incluían en el programa de estudios por no estar formados en la materia, o por temor a la reacción de determinados segmentos de la sociedad. El primer programa fue la formación de profesores de la escuela pública de Kenia sobre la transmisión del sida. Estamos hablando de miles de profesores y esto lleva tiempo. ¿Qué efecto tiene que el gobierno de Kenia empiece a formar a los profesores con un programa de estudios sobre el sida? En general esta primera intervención tuvo poco impacto en los conocimientos, actitud y comportamiento de los estudiantes con respecto al sida. Este programa también tuvo un efecto muy limitado sobre las tasas de embarazo de las estudiantes. La única noticia positiva de esta intervención era que, después de dos años, las niñas de los colegios que tenían profesores que habían participado en los programas de formación sobre el sida y habían tenido hijos se habían casado y, probablemente, tenían una relación más estable.

Pero también podía ser que las jóvenes tuvieran relaciones con personas mayores, que en ocasiones se casaban con chicas en edad escolar. Estos hombres tenían más probabilidad de estar ya infectados por el virus del sida. Este hombre de mayor edad que se casa, o seduce, a una joven es lo que se conoce en la literatura en inglés por *suggar daddy*. Además, la formación de docentes en este programa de estudios sobre el sida no les facilitaba responder a preguntas de las niñas sobre el uso de preservativos pues, en muchos casos, se impedía que se facilitase información sobre este tema.

Una segunda intervención era alentar a los estudiantes a hablar sobre el uso de preservativos y alentarles a escribir ensayos y participar en debates sobre la forma de protegerse del virus del sida. Los profesores, por lo general, no podían hablar de los preservativos pero sí podían permitir que los estudiantes debatieran sobre los mismos. En cierta forma esta situación es un buen compromiso pues se habla del tema en un debate abierto con

la implicación de los estudiantes. No es el profesor quien les dice lo que tienen que hacer. Normalmente se hace que unos tomen la posición pro preservativo y otros la posición en contra del preservativo. En estas circunstancias es aceptable hablar sobre el tema y los partidos políticos lo aceptaron, pues no se alentaba a nadie a usar el preservativo: era un simple debate. El resultado de esta segunda intervención fue una mayor información en cuanto a su uso. No se observaba un aumento de la actividad sexual previamente informada por los alumnos. No obstante, sí se observaba un mayor nivel de reconocimiento del uso de preservativos.

La tercera intervención consistía en pagar el coste de los uniformes, que son obligatorios, para reducir el precio de la educación y mantener a las niñas más tiempo en el sistema educativo. El objetivo era utilizar este mecanismo para que las niñas permanecieran más tiempo en la escuela. En esa sociedad, en la práctica las niñas que quedan embarazadas tienen que dejar la escuela. Por tanto, si quieren permanecer en la escuela se cuidan más a la hora de tener relaciones sexuales, y así tienen mayores posibilidades (por costarles menos) de seguir en el colegio. El resultado de esta intervención fue una reducción en las tasas de abandono de la escuela, en el matrimonio de jóvenes adolescentes y en sus tasas de embarazo. Ésta es la única medida de apoyo que tuvo un cierto efecto comparado con las otras tres intervenciones.

Finalmente la última intervención, que aparece recogida en Dupas (2006), es un programa que proporcionaba información a los alumnos sobre el riesgo del sexo entre personas de diferentes generaciones. El gran debate político es sobre el uso del preservativo o no. Pero un problema fundamental en la transmisión del sida es la gran diferencia que existe en la tasa de infección entre los jóvenes varones y los de mayor edad (que tienen mayor probabilidad de estar infectados). Esta cuarta intervención consistía precisamente en proporcionar información sobre la tasa de infección por diferentes edades y cuán alta era la tasa de infección en los varones adultos. Las niñas ya saben cómo se transmite el sida. La cuestión es, por tanto, informar sobre la probabilidad de contagio en actividades de riesgo con adultos de mayor edad pues ésta es una información más sofisticada que puede no estar en el conjunto de conocimientos comunes de las niñas en edad escolar. En el grupo de chicas informadas que tenían relaciones con adultos mayo-

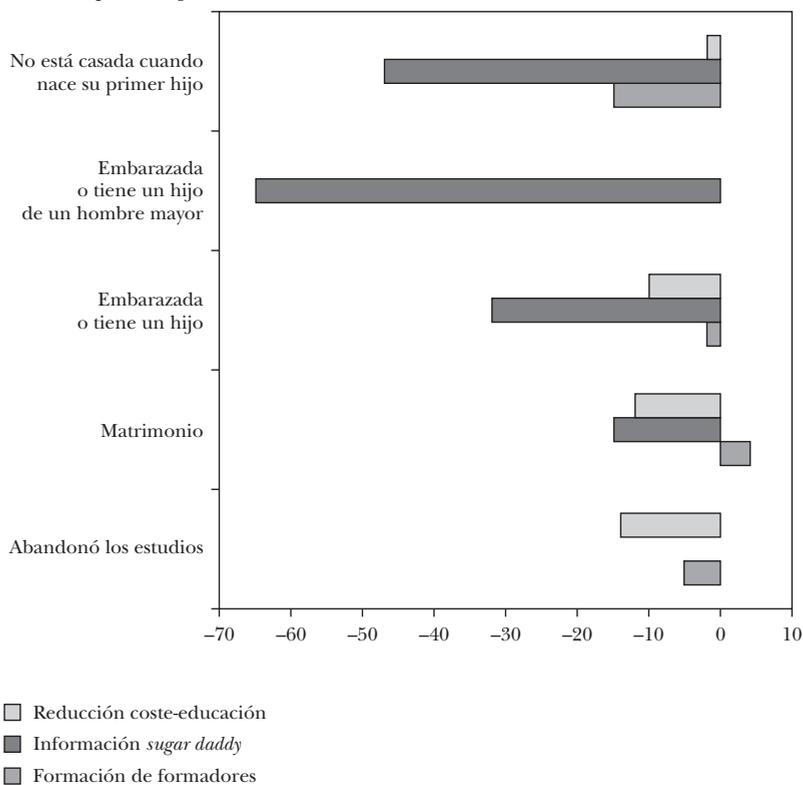
CUADRO 2.1: Efecto de las campañas informativas sobre el comportamiento sexual de las niñas y los niños

Programa	Actuaciones	Impacto	Coste por estudiante (dólares)	Coste embarazo evitado (dólares)
Formación de formadores	Formación de profesores de primaria para que puedan enseñar el currículo del gobierno sobre el tema del sida	— Aumento de la tolerancia — No efecto sobre tasa de embarazos — Aumento de tasas de matrimonio entre las jóvenes embarazadas — Aumento del conocimiento entre los estudiantes	2	Sin efecto
Debates y ensayos	Promocionar la organización de debates sobre condones en los colegios y concurso de ensayos sobre la transmisión del sida	— Aumento de la probabilidad de que los estudiantes usen condones — No efecto sobre el número de encuentros sexuales	1,10	En curso
Charlas sobre <i>Sugar daddies</i>	— Enseñar vídeo sobre los peligros de relaciones con hombres mayores — Información sobre el riesgo de contraer el sida por sexo y edad	Reducción de la tasa de embarazo por un hombre más viejo, debido a la reducción del número de niñas que practica sexo con hombres mayores	0,80	91
Reducir el coste de la educación	Proporcionar uniformes gratis a estudiantes de primaria para que las niñas estén más tiempo en la escuela	— Reducción de las tasas de abandono de los estudios — Reducción de las tasas de matrimonio y embarazo	10,8 (hombres) 12 (mujeres)	750

res hubo una disminución del 65% del número de gestaciones en chicas. Por el contrario, no se observó ningún cambio en las tasas de embarazo de chicas que tenían parejas de su misma edad (adolescentes).

El cuadro 2.1 presenta un resumen de los cuatro programas comentados en este epígrafe así como el coste por estudiantes y su efectividad

GRÁFICO 2.1: **Eficacia relativa de cada programa sobre los objetivos básicos** (porcentajes)



en términos de evitar embarazos no deseados. Estos programas no sólo son efectivos, en su mayoría, sino que también son baratos. Proporcionar dos uniformes en el plazo de tres años cuesta 12 dólares por niña, e informar a los adolescentes sobre los riesgos de una relación *sugar daddy* cuesta menos de un dólar por estudiante. Para los programas para los que existe información la relación coste-efectividad es muy favorable.

El gráfico 2.1 muestra el impacto de cada uno de los cuatro programas analizados sobre el comportamiento de las chicas adolescentes, en porcentaje relativo a las niñas de su grupo de comparación. La reducción del coste de la educación tiene efecto significativo (en torno a una reducción del 10%) sobre el abandono escolar, el matrimonio adolescente y el embarazo. Por su parte, la información sobre los peligros de una relación *su-*

gar daddy obtiene excelentes resultados en la reducción de las tasas de embarazo, las tasas de matrimonio de adolescentes y, especialmente, las tasas de embarazo por un hombre mayor.

El análisis de los programas anteriores muestra que los métodos son importantes, pero que hay que tener en cuenta las características de la sociedad concreta sobre la cual se van a aplicar. Si la cuestión de la educación sexual, o relacionada con temas sexuales, no está bien aceptada en una sociedad, hay formas alternativas de informar a los interesados de los peligros asociados a determinadas actividades de riesgo. Y para ello hay que tener en cuenta una serie de perspectivas diferentes. Todos los enfoques aportan nuevas perspectivas con las que mejorar las evaluaciones para realizar un trabajo cuantitativo de evaluación capaz de establecer cuál es el impacto de actuaciones y programas.

Es un sistema complicado que evoluciona con muchos factores que, por ejemplo, pueden modificar el comportamiento sexual. En los párrafos anteriores hemos visto el impacto de diferentes componentes o intervenciones. Quizás los efectos sean diferentes en otros entornos, por ejemplo, en Sudáfrica. Habría que realizar las mismas intervenciones y ver qué ocurre.

Hay prioridades para el futuro trabajo en la evaluación de programas destinados a cambiar los comportamientos sanitarios y sexuales. Tenemos ensayos médicos recientes que sugieren que la circuncisión puede ser un mecanismo muy eficaz para reducir las tasas de infección. Ya se ha hecho un estudio en el este de África que indica que la probabilidad de infección baja un 50% entre los varones circuncidados. Pero ¿cómo podemos incentivar la práctica de la circuncisión? Y si el varón se ha circuncidado, ¿cómo evitar que el hecho de sentirse más seguro frente al sida no produzca mayor actividad sexual o el abandono del sexo seguro?

En resumen, podemos concluir que hemos aprendido de los experimentos destinados a cambiar la conducta sexual y los hábitos sanitarios. En primer lugar, el coste es muy importante. Reducir el coste es fundamental. Las subvenciones pueden reforzar la participación en estos programas. Pero hay que buscar procedimientos alternativos a la clínica como mecanismo para contactar con la población objetivo de la intervención. Una posibilidad que se ha demostrado efectiva es utilizar las escuelas, por

ejemplo. Hay que identificar las oportunidades y los efectos que puedan servir de palanca para cambiar los hábitos sanitarios.

2.3. La mejora de los sistemas sanitarios

Después de analizar las posibilidades de cambiar los hábitos sanitarios y los comportamientos de los individuos, existe un segundo reto importante para la mejora de la salud de los habitantes en países del tercer mundo. Este segundo reto es la mejora del sistema sanitario desde el lado de la oferta. El análisis experimental de un sistema sanitario en su conjunto es muy complejo. No obstante, existe un estudio reciente de tipo experimental sobre la transformación del sistema sanitario en Camboya. El sistema de salud de Camboya era uno de los peores del mundo. Tenía un sistema público que no funcionaba (con incentivos muy débiles) y un sistema privado donde había enormes problemas (los incentivos estaban desalineados con respecto a los intereses generales de los pacientes). El experimento en cuestión evalúa una modificación en el sistema de gestión sanitaria utilizando 12 distritos de Camboya. Se puso en marcha un concurso para gestionar los servicios sanitarios de ocho distritos elegidos aleatoriamente. Se aumentaron los fondos públicos asignados a los ocho distritos tratados (reduciendo el gasto privado de los pacientes). Diversas ONG se adjudicaron la prestación de los servicios. El rendimiento de los adjudicatarios en cada distrito se midió utilizando ocho objetivos en los cuales se habían fijado metas específicas: alcanzar un determinado nivel de vacunaciones, administración de vitamina A, cuidados durante el embarazo, formación para el parto, partos en clínica, uso de anticonceptivos, conocimiento de procedimientos anticonceptivos y utilización de centros de atención públicos. La adjudicación de los servicios sanitarios en los distritos permite, al menos potencialmente, vincular los incentivos a objetivos sanitarios públicos; permite la comparación con un sistema competitivo de referencia; y agrega los riesgos del distrito en una situación donde puede existir selección adversa, aunque limitada (en zonas rurales). Por tanto, evita que un adjudicatario particular tenga que hacer frente a un gran número de enfermos por la situación concreta de la localidad que le ha

tocado servir. El único problema es que se desvía el esfuerzo de actividades importantes que no aparecen recogidas en los ocho objetivos específicos del programa.

La historia de Camboya en los últimos cuarenta años ha sido muy convulsa. Vale la pena hablar un poco del contexto histórico. El régimen de los Khmer Rouge propició un genocidio de enormes proporciones entre 1975 y 1979. Cuando Camboya se estaba empezando a recuperar del genocidio de los Khmer Rouge, se produjo la invasión vietnamita, que duró hasta 1993. En todo este proceso el sistema sanitario quedó totalmente destruido. La población era tratada por una serie de personas que vendían fármacos y que no necesariamente tenían formación médica alguna. En 1993 hubo unas elecciones y se adoptó una economía de mercado. La lucha finalizó definitivamente en 1998.

Desde 1993, el sistema ha mejorado mucho, aunque en parte la mejora es consecuencia de la desastrosa situación inicial. El sistema sanitario sufre una gran politización y tiene problemas de gestión. El salario de los trabajadores sanitarios públicos es un 85% de la renta per cápita. A partir de 1993 se produce una explosión de la sanidad privada ante el absentismo generalizado en el sector sanitario público. Los vendedores de fármacos siguen recibiendo el 33% de las visitas curativas. En general, la población sigue teniendo una visión tradicional de las enfermedades. El presupuesto sanitario es elevado pero la cobertura de los servicios sanitarios es pobre.

El proyecto que analizaremos en este epígrafe cubre el período 1999-2003. Durante este período se ha producido una mejora enorme de la salud de la población de Camboya, en parte gracias a la construcción de multitud de centros sanitarios.

El objetivo de la evaluación es medir el efecto del cambio del sistema de gestión sanitaria. Pero, obviamente, hay que separar el resultado de este programa de los efectos globales que se están produciendo en el sistema sanitario de Camboya. Por tanto, la metodología del análisis no puede ser una comparación de antes y después. Por ese motivo, se utiliza la comparación de los distritos donde estaba el programa frente a otros distritos (grupo de control) donde no se implementaba o se llevaba a la práctica el programa. El programa cubría al 11% de la población.

Se trata de un proyecto que incluía un paquete mínimo de actividades. Los objetivos de mejora se centraban en la cobertura de servicios pediátricos y sanidad materno-infantil, con una orientación fundamentalmente preventiva.

El tratamiento tenía dos variantes. El primer sistema, *contracting in* (CI), supone que el adjudicatario tiene la autoridad en gestión pero no puede contratar ni despedir, y las compras de material sanitario se producen fuera del sector gubernamental, aunque el presupuesto operativo lo transfiera el gobierno. El segundo sistema, *contracting out* (CO), implica un control total de las contrataciones y los despidos del personal sanitario, un control absoluto de las decisiones sobre los proveedores de material sanitario y un presupuesto operativo a través de préstamos del Banco Mundial y del Banco Africano de Desarrollo. Hay tres provincias con tres, cuatro y cinco distritos elegibles, respectivamente. Finalmente, hay tres distritos de la capital. Los distritos tratados fueron asignados aleatoriamente al sistema de CI, al sistema de CO o al grupo de control. Cada distrito tenía igual probabilidad de ser asignado a uno u otro grupo. El proceso de adjudicación se produjo a partir de ofrecer un precio fijo per cápita. Siguiendo los criterios técnicos y de precio, tres distritos recibieron ofertas técnicamente inaceptables.

El método de evaluación se basa en los ocho indicadores objetivo comentados con anterioridad. En las regresiones que se se presentan a continuación la variable Y^K mide el indicador k

$$Y_{jip}^K = \beta_0 + \beta_1 CI_i + \beta_2 CO_i + \beta_3 CI_i DO3_t + \beta_4 CO_i DO3_t + \beta_5 DO3_t + \varepsilon_{jip} + \rho_{pt}$$

donde $CI_i = 1$ si el distrito i tiene un sistema tipo *CI*, $CO_i = 1$ si el distrito i tiene un sistema tipo *CO* y $DO3 = 1$ si el año es posterior al 2003 (año de inicio del experimento). Los subíndices indican lo siguiente: i distrito; j individuo; t tiempo; y p provincia.

El cuadro 2.2 nos permite comparar el efecto sobre los objetivos de los diferentes tipos de sistemas frente al grupo de control. Algunos tuvieron efecto, mientras que en otros casos no está tan claro. Para hacernos una idea de la importancia del efecto, podemos centrarnos en la tercera co-

CUADRO 2.2: Efecto sobre los diferentes indicadores de servicios sanitarios

	Vac	VitA	Pre	Par1	Par2	Ant1	Ant2	Pub	Media
CI	-0,09 (0,08)	-0,02 (0,03)	-0,006 (0,03)	0,02 (0,12)	0,02 (0,03)	0,00 (0,04)	0,04 (0,08)	-0,00 (0,01)	
CO	-0,10 (0,14)	-0,14** (0,06)	0,03 (0,10)	0,13 (0,17)	0,01 (0,03)	0,11 (0,12)	-0,07 (0,12)	-0,00 (0,03)	
CI*D03	0,14 (0,08)	0,09 (0,06)	0,36*** (0,08)	0,05 (0,04)	0,11 (0,07)	0,07 (0,06)	-0,02 (0,07)	0,17*** (0,04)	0,50*** (0,09)
CO*D03	0,15 (0,12)	0,42*** (0,09)	0,26 (0,16)	-0,12 (0,11)	0,07 (0,07)	-0,04 (0,09)	0,07 (0,13)	0,29*** (0,05)	0,53*** (0,12)
D03	0,29** (0,10)	0,15*** (0,04)	0,34*** (0,11)	0,20*** (0,04)	0,12 (0,07)	0,15** (0,05)	0,58*** (0,06)	0,14 (0,02)	
N	5.100	11.213	4.993	4.993	4.976	6.994	9.537	11.223	
R2	0,27	0,13	0,25	0,03	0,05	0,02	0,34	0,12	
Media 2003	0,81	0,61	0,35	0,34	0,10	0,23	0,80	0,13	
Media 1997	0,34	0,43	0,09	0,24	0,03	0,13	0,22	0,04	
P val. H0:CO = CI									0,83

Nota: Las columnas se refieren a las siguientes variables endógenas: Vac: vacunaciones. VitA: suministro de vitamina A; Pre: tratamiento prenatal; Par1: formación para el parto; Par2: partos en la clínica; Ant1: uso de métodos anticonceptivos; Ant2: información sobre métodos anticonceptivos; Pub: uso de clínicas públicas. La estimación es por variables instrumentales incluyendo efectos fijos por provincia y año. El efecto medio es el incremento diferencial causado por el tratamiento en desviaciones estándar respecto al grupo de control. Los errores estándar que se presentan están corregidos por correlación intragrupo en los distritos. Las estrellas indican significatividad estadística: * al 10%; ** al 5%; *** al 1%.

lunna del cuadro 2.2. El sistema CI produjo un aumento del 36% en el número de mujeres que recibía tratamiento prenatal. Esto es una mejora muy considerable. Si se toman conjuntamente las medias de estos ocho indicadores diferentes comprobamos que sí existe un efecto conjunto (última columna).

CUADRO 2.3: Efecto medio por familia de indicadores

	ResCon	ResNCon	SalF	VerMed	Gestión	Percepción
CI-tratados	0,50***	0,28	-0,04	0,66***	0,60***	-0,12
Valor p	0,09	0,12	0,18	0,16	0,14	0,14
CO-tratados	0,53***	0,08	0,61**	0,70***	1,12***	-0,25**
Valor p	0,12	0,26	0,56	0,33	0,38	0,14
H0:CO = CI, p	0,83	0,78	0,31	0,88	<0,01	0,38

Nota: Las columnas se refieren a familias de indicadores de resultados: ResCon: resultados en el contrato; ResNCon: resultados no incluidos en el contrato; SalF: resultados de salud finales; VerMed: probabilidad de acudir a un médico; Gestión: gestión de los centros sanitarios; Percepción: percepción de los consumidores sobre la calidad del servicio. Los errores estándar que se presentan están corregidos por correlación intragrupo en los distritos. Las estrellas indican significatividad estadística: * al 10%; ** al 5%; *** al 1%.

El cuadro 2.3 muestra los resultados agrupados por familias de indicadores. Es evidente que el efecto de la intervención fue muy significativo en los objetivos que estaban incluidos en los contratos firmados por los diferentes proveedores de servicios sanitarios, como muestra la columna 1. Sin embargo, los resultados no son estadísticamente significativos en el caso de los indicadores que no estaban incluidos en los contratos. Por tanto, los incentivos funcionan de la manera prevista. También hay evidencia de que se produce una significativa mejora en los pacientes que acuden a los centros contratados frente al grupo de control. Parte de esta mejora está relacionada con el hecho de que acuden, con mayor probabilidad que en el grupo de control, a recibir los cuidados del trabajador sanitario. La penúltima columna muestra grandes efectos sobre la gestión de los centros, que son más importantes en la modalidad CO que en la modalidad CI. Así lo muestra el rechazo de la hipótesis de igualdad de los efectos en ambos sistemas frente al grupo de control. Éste es el resultado que esperaríamos obtener dadas las mayores restricciones sobre contratación impuestas en la modalidad CI frente a la CO. La enorme mejora en la gestión de los centros estaba asociada a la probabilidad de recibir atención por un proveedor cualificado. Es decir, se reducían las visitas al personal no cualificado y se aumentaba enormemente la presencia de personal cualificado en la instalación sanitaria. La gestión de las clínicas mejoraba porque estaban abiertas y el personal estaba presente. Finalmente, no me-

jó la percepción de la calidad por parte de los pacientes. Este resultado es reflejo de la comparación con una situación (sector público o grupo de control) donde se proporcionan servicios, como por ejemplo inyecciones de glucosa con poco efecto terapéutico, que los pacientes piensan que son útiles aunque su efecto es muy discutible.

¿Qué pasó con la financiación? Si analizamos el gasto sanitario total (público y privado) se observa cómo no aumenta, e incluso disminuye, en las nuevas modalidades de contratación. ¿Por qué? Porque los proveedores contratados no gastan recursos en inyecciones de glucosa y otros placebos que no tienen efecto real en los problemas sanitarios de los pacientes.

En resumen, podemos decir que la experiencia de Camboya es muy positiva. Con anterioridad a los programas analizados existía una mezcla de sistema privado y público donde ambos proveían servicios de baja calidad. La subcontratación de ONG para la gestión de los servicios sanitarios mejora los servicios estatales y genera menos dependencia de los proveedores privados, menos gasto global y mejor cobertura en asistencia prenatal, vacunaciones, etc.

Obviamente, el caso comentado es sólo una de las posibles intervenciones. En algunos casos no es un problema de gestión sino de disponibilidad de recursos. Pero también existen problemas de corrupción, politización de la asignación de los recursos sanitarios y dificultades incluso para gastar el presupuesto asignado. Por tanto, los cambios institucionales en el sector sanitario más el aumento de los presupuestos públicos son sólo un conjunto posible de intervenciones.

Además, Camboya era un país que estaba tan mal como puede estar el Congo hoy y, sin embargo, comprobamos que incluso en un entorno tan difícil se pueden obtener mejoras. Valdría la pena probar en otros países que viven situaciones de posconflicto como el Congo, Sierra Leona, Liberia, y quizás valga la pena probarlo también en otros lugares más estables, pero donde hay problemas muy serios en los servicios sanitarios. Hay un programa en Pakistán ahora mismo que está intentando hacer algo similar. Estos mecanismos ofrecen una forma de afrontar no solamente los problemas del comportamiento del personal sanitario, sino también otros asociados a los sistemas sanitarios.

2.4. Conclusiones

Los avances médicos proporcionan tecnologías realmente baratas y aplicables masivamente (antibióticos, vacunaciones, etc.) que permiten cubrir las necesidades básicas de la gente. De esta manera, se puede conseguir aumentar significativamente la esperanza de vida incluso en países con bajos niveles de renta per cápita. Sin embargo, hay millones de personas en el mundo que no están recibiendo asistencia sanitaria, o no la adecuada. Para proporcionarles esa asistencia hay que cambiar su comportamiento individual, pero también hay que mejorar el acceso y el tratamiento del agua, por ejemplo. Para poder difundir la información necesaria para cambiar los hábitos nocivos, o arriesgados, para la salud, se pueden utilizar diferentes técnicas. La evidencia disponible actualmente sugiere que el coste marca una diferencia importante. Algunas de estas intervenciones son muy baratas pero, incluso si fueran costosas, se podría reducir dicho coste individual con la provisión de subsidios. Experimentar con varios mecanismos es importante para determinar su influencia relativa sobre el comportamiento ante riesgos para la salud.

La subcontratación de los servicios sanitarios, como en el caso comentado de Camboya, puede mejorar sensiblemente la provisión de servicios. No obstante, se necesita más evidencia científica sobre estos nuevos sistemas de provisión descentralizada de servicios sanitarios.

Este capítulo también utiliza como procedimiento de investigación la evidencia suministrada por experimentos aleatorizados. Los métodos aleatorizados nos permiten saber lo que funciona y lo que no. Además, estos procedimientos se pueden utilizar no sólo para evaluar los programas sino también para mejorarlos una vez que se han puesto en marcha, modificando su aplicación.

Este tipo de investigación también permite comparar la calidad de diferentes actuaciones en términos de coste-beneficio. Tenemos que asegurarnos de que los recursos que se destinan a mejorar la vida de los ciudadanos de países en vías de desarrollo se gastan en los programas más eficientes y efectivos.

Ya se ha comentado que el programa de desparasitación tuvo un impacto muy importante, no sólo sobre los niños tratados sino también so-

bre los del grupo de control por el efecto de las externalidades. Sin embargo, puede haber otros programas que no funcionen tan bien. Si sabemos lo que funciona y lo que no, podemos organizar mejor las cosas, y presentar las evidencias a los escépticos para convencerles de que realmente esos programas sí sirven. La ventaja de los métodos que se basan en experimentos aleatorios es que sus resultados son transparentes, fáciles de comprender y creíbles para el público en general, investigadores y agentes políticos. De esta forma, se podrán conseguir más fondos para seguir mejorando la vida de las personas pobres en países en vías de desarrollo.

Volviendo a la metáfora inicial. Si pensamos en la transformación que ha tenido lugar en la medicina en los últimos cien años, podemos comprobar que el procedimiento experimental ha sido fundamental para evaluar qué funcionaba y qué no funcionaba. El progreso médico se ha beneficiado extraordinariamente de esta metodología. Creo que deberíamos utilizar esta experiencia para generalizar la utilización del procedimiento de experimentación aleatorizada en el análisis de los programas y actuaciones en el campo de la ayuda al desarrollo.

Bibliografía

- BLEAKLEY, H. «Disease and Development: Evidence from Hookworm Eradication in the American South». *Quarterly Journal of Economics*, febrero 2007.
- DUFLO, E., P. DUPAS, M. KREMER, y S. SINEI. «Education and HIV/AIDS Prevention: Evidence from a Randomized Evaluation in Western Kenya». World Bank Policy Research Working Paper 4024, 2006.
- DUPAS, P. *Relative risks and the market for sex: teenagers, sugar daddies and HIV in Kenya*. MPRA Paper 248. University Library of Munich, 2006 [mimeo.].
- GERTLER, P. «Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment». *American Economic Review (Papers and Proceedings)* 94 (2), 2004: 336-341.
- KREMER, M., y E. MIGUEL. «Illusion of Sustainability». *Quarterly Journal of Economics*, 112 (3), 2007: 1007-1065.
- . «Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities». *Econometrica*, 72 (1), enero 2004: 159-217.
- SCHULTZ, P. «School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program». *Journal of Development Economics* 74, 2004: 199-250.

La evaluación de las intervenciones educativas: evidencia a partir de experimentos aleatorizados

Esther Duflo

Catedrática de Economía del MIT

Directora del Centro de Investigación Abdul Latif Jameel Poverty Action Lab

3.1. Introducción

El objetivo de este capítulo es analizar lo que sabemos sobre la evaluación de las intervenciones educativas en países en vías de desarrollo. El capítulo se centra en varios programas en los que he intervenido directamente y otros desarrollados por diversos colaboradores. Aunque la evidencia es generalmente dispersa, aparece de forma resumida en las publicaciones del Poverty Action Lab (Laboratorio para Acción contra la Pobreza). La mayoría de esta información está accesible a partir de la página web del centro, aunque todos los días aparecen nuevas evidencias. En cualquier caso, en el Lab tenemos un buen punto de comparación para saber hasta dónde hemos llegado en nuestro conocimiento sobre lo que funciona y lo que no, en el campo de la promoción educativa en los países en vías de desarrollo.

Hay dos dimensiones básicas, o si se quiere dos retos, en la cuestión de la educación en países del tercer mundo. Primero, hay que conseguir que

los niños vayan a las escuelas. De hecho, uno de los Objetivos de Desarrollo del Milenio de las Naciones Unidas es que todos los niños y niñas (el 100%) completen, al menos, la educación primaria. Pero, además de que los niños y niñas acudan al colegio, es necesario que cuando estén allí aprendan algo. El problema es que si examinamos la situación en muchos países en desarrollo se están haciendo bastantes avances con respecto al primer objetivo, pero no tantos en el segundo. Hay más niños que nunca en toda la historia educativa mundial que van al colegio y que terminan, por lo menos, la educación primaria. La situación para las niñas también ha mejorado, aunque sigue siendo peor en general. Con todo, está mejorando y la tendencia es positiva. Las mejoras en el nivel de educación secundaria son mucho más limitadas. Quizás por este motivo no se habla tanto de este aspecto. Además, si se intentara que todos los niños completaran la educación secundaria existiría un problema básico: no hay suficiente presupuesto ni personal.

Como se ha señalado anteriormente, la asistencia, el acceso, a la escuela primaria ha progresado mucho. Sin embargo, la calidad de muchas de estas escuelas sigue siendo un desastre. En el capítulo anterior, Michael Kremer ha descrito el problema del absentismo del personal sanitario en países en vías de desarrollo. En el sector educativo existe el mismo problema. Los maestros en muchos sitios no acuden a su lugar de trabajo. Por ejemplo, si se realizan visitas sin anunciar, el 25% de los maestros de educación primaria de la India no está en su puesto de trabajo. En Uganda esta cifra se eleva hasta el 27%. En Indonesia alcanza el 19%. En el conjunto de países (Bangladesh, Ecuador, la India, Indonesia, Perú y Uganda) analizados por Chaudhury et al. (2006), la media de la proporción de maestros que no acuden a sus puestos de trabajo es del 19%.

Entre los profesores que están efectivamente presentes en la escuela, tan sólo la mitad del tiempo la dedican a la enseñanza. El resto está bebiendo té, charlando en los pasillos o haciendo otras actividades que no están relacionadas con la docencia. Estos datos son indicativos de un problema grave en la calidad de la enseñanza que reciben los niños y niñas que consiguen ir al colegio.

No solamente hay que conseguir que los niños vayan al colegio sino que hay que mejorar la calidad del mismo. Un problema importante es lo

que aprenden los niños en la escuela. La ONG Pratham realizó un estudio sobre la calidad de la educación primaria en la India. Se envió a voluntarios a todos los distritos de la India equipados con un instrumento muy sencillo, que era un párrafo que daban a leer a los niños. El 60% de los niños entre 7 y 12 años fueron incapaces de leer este sencillo y breve párrafo, que deberían haber podido leer sin ningún problema si su evolución educativa hubiera sido la correcta. Una tercera medida, o indicación, de la baja calidad del sistema educativo en estos países es lo que los economistas denominamos *votación con los pies*, en el sentido de que los estudiantes pueden decidir abandonar el sistema educativo si no les ofrece una calidad educativa razonable. Por ejemplo, en la India y en Pakistán, decimos que los estudiantes *votan con los pies* porque se marchan del sistema público. Chile, por ejemplo, tiene un sistema totalmente privatizado basado en cheques escolares universales que se pueden utilizar para acceder a escuelas públicas o privadas. En la India, sin embargo, el sistema básico es público. Con todo, la tasa de niños escolarizados en las escuelas privadas de India es más alta que en Chile o en los Países Bajos. Es decir, que el sistema indio de facto está viéndose privatizado porque la gente está enviando a los niños a las escuelas privadas al huir del sistema público. Es una situación similar a lo que Michael Kremer comenta en el capítulo anterior sobre la sanidad en muchos países en vías de desarrollo. La gente va al curandero, donde sea, con tal de no ir al sistema público. En la educación primaria está pasando algo similar. Habría que puntualizar que hay variación en la calidad de la escuela privada de estado a estado en la India. Esta privatización, o tendencia hacia la escolarización privada, tiene una intensidad variable en función del estado de residencia, pues, en algunos casos, el sistema público es algo mejor y en otros estados es mucho peor (normalmente en los estados más pobres). Por tanto, la cuestión no es sólo que los niños vayan a la escuela. Es necesario que también aprendan algo mientras están allí.

Ante estos problemas, la tendencia en el sector educativo, como en el sanitario, es el voluntarismo unificador: «seguro que hay algo que se puede hacer y que funciona pues algo habrá que hacer». Es parte de la naturaleza humana. La tendencia natural es decir que debe haber una solución que podría resolver todos los problemas. Lo que los ingleses llaman

una solución infalible (*silver bullet*) para todos los problemas. Es humano buscar la panacea que resuelva todo. Esta panacea varía, es decir, tiene ciclos, como ha señalado Michael Kremer en su capítulo. Una de esas *silver bullets* o panaceas, o soluciones cúralo-todo, ha sido la formación de formadores, en la que se depositaron enormes cantidades de fondos. Luego las ideas cambian, se apartan de la panacea inicial, admiten que hay limitaciones, y pasan a la siguiente gran idea. En el campo educativo la siguiente panacea fue la de los comités de padres y los comités de escuela: si se dan los recursos a los padres se asegura que los mismos se impliquen, o participen, en la enseñanza y en el aprendizaje de sus hijos. Pero tampoco funcionó con generalidad. Más tarde llegaron los cheques escolares (*vouchers*). Y así sucesivamente.

Muchas de estas ideas podrían funcionar parcialmente, pero probablemente no habrá una única solución, algo que solucione todos los problemas a la vez y de un solo golpe. Habrá que aprender de la evidencia, que nos enseña, de una forma acumulativa y lenta, que algunas cosas funcionan y otras no han funcionado. Hay que dejar de buscar panaceas para los grandes problemas y concentrarnos en encontrar soluciones que funcionen aunque sean en pequeños formatos. Habría que ir más despacio y avanzar de lo particular a lo general, para saber qué se puede hacer poco a poco. Para estar en mejor posición de resolver los problemas hay que afrontarlos por partes en lugar de conjuntamente, pues muchos de los problemas de la economía del desarrollo son extraordinariamente complejos.

Pero ¿cómo sabemos qué funciona y qué no funciona? El gran problema en el ámbito de la educación es la dificultad para medir su efectividad directamente a partir de la prueba del mercado. Si queremos, por ejemplo, saber las bondades de un nuevo coche no hay ningún problema. Lo podemos probar. Además, tenemos la prueba del mercado: si nadie compra el coche, entonces no es un buen coche. La gente no está dispuesta a pagar un precio enorme por un coche poco fiable o que no funciona bien. Esto sucede con muchos otros productos. La mayor parte de las veces hacemos la prueba del mercado. Si la gente lo compra, entonces es un buen producto.

En cuanto a la entrega de servicios como la educación o como la salud, no existe prueba de mercado. No se puede medir la efectividad en térmi-

nos de ventas. La razón es que los servicios sociales precisamente son algo que queremos subvencionar porque pensamos que son necesarios, que hacen falta. Hay que ayudar a los ciudadanos y los países pobres, y el acceso a estos servicios debe estar subvencionado. Pero si se subvencionan los servicios públicos entonces no podemos medir su efectividad en términos de ventas, a menos que sea demasiado tarde, como el sistema educativo de la India donde, a pesar de las subvenciones, la gente huye al sector privado. Ni siquiera compra lo que resulta gratis.

Hay que garantizar que se entregue algún valor, que se dé algo a cambio. Por otra parte, no hay una forma automática, como sería el mercado, para saber si se entrega algo útil o no y, por tanto, no hay garantía alguna de que se gaste el dinero público en cosas útiles. Hay dos aspectos importantes que se deben considerar en este punto. En primer lugar, hay que asegurarse de que se esté haciendo lo que realmente queremos hacer. Por este motivo la evaluación del proceso resulta esencial. Es preciso que el programa haga, o financie, lo que se supone que debe hacer o financiar. En segundo lugar, es necesario asegurarse de que el programa tiene efecto. Para esta segunda tarea se realizan evaluaciones de impacto.

Obviamente, antes de pasar a esta segunda fase es necesario haber superado la primera. Si no se sabe si lo que se ha entregado es lo que se quería entregar, entonces no tiene sentido ver si hay impacto o no. Hay que ver primero qué se ha ofrecido. Esta evaluación del proceso tiene que producir la verificación de cuentas para asegurarse de que el dinero ha llegado a su destino y está financiando las actividades previstas. Es preciso fiscalizar las cuentas para saber que allí donde había que hacer compras se han hecho compras; se ha pagado a los profesores, etc. La evaluación del proceso hay que hacerla de forma cuidadosa. No hay una gran dificultad conceptual en su desarrollo, aunque en la práctica puede ser difícil (no se producen los documentos necesarios; no se mantiene una contabilidad; no se realizan inventarios, etc.) y se necesitan especialistas. Pero en el aspecto conceptual la evaluación del proceso no tiene dificultades.

Una vez que se evalúa el proceso, sabemos que estamos haciendo lo que queríamos hacer. Por ejemplo, si queremos crear comités de escuela, queremos asegurarnos de que efectivamente se han creado. No obstante, en las siguientes líneas veremos que no es tan fácil: no es tan evidente

comprobar si el gasto se ha realizado en las partidas previstas, en algunas ocasiones. Una vez se comprueba que el proceso ha sido correcto hay que realizar la evaluación del impacto.

3.2. El papel de la evaluación del impacto de las intervenciones

Incluso los programas con buenas intenciones y bien gestionados pueden no tener los efectos deseados. Algunos investigadores señalan que si empezamos a solicitar evaluaciones de impacto a las ONG podría haber un colapso. Pero es necesario conocer el impacto de las actividades. No se trata de juzgar. Cuando se realiza una evaluación de impacto se analiza qué ha pasado, y si ha sido una buena idea o no desarrollar unas determinadas actuaciones. La mayoría de las veces cuando se programa y se realiza una intervención, no comprendemos perfectamente los efectos que puede tener. Normalmente lo que hacemos, ya seamos economistas, antropólogos u otro tipo de especialistas, es tomar un mundo muy complejo y simplificarlo, construyendo un modelo que indique cómo funciona el mundo. A partir de este modelo proponemos soluciones. Pero muchas veces, por no decir casi siempre, nos equivocamos. Es normal. En la economía del desarrollo no hay otra forma: se aprende cometiendo errores. Incluso con las mejores intenciones, con la mejor dirección, un programa puede no tener ningún impacto o, incluso, consecuencias secundarias negativas e inesperadas.

Por tanto, es necesario experimentar en el diseño de programas, probar cosas. Esta necesidad nunca va a desaparecer. Hay actuaciones que pueden ser, o parecer, muy razonables, pero para comprobar si realmente tienen algún efecto es necesario realizar la evaluación de su impacto. La comparación de los proyectos evaluados, en base a criterios comparables,¹ genera una serie de buenas prácticas (*best practices*) que permite, en el fu-

¹ Los experimentos aleatorizados, por su transparencia y dificultad de manipulación, generan una comparabilidad entre experimentos que difícilmente se puede conseguir con otras metodologías.

turo, utilizar los procedimientos más efectivos para conseguir unos determinados resultados. Si no se configura un conjunto de *buenas prácticas* todo el mundo intenta justificar su idea o aproximación particular, y entramos en los ciclos mencionados anteriormente. Sin *buenas prácticas* cualquiera parece justificado para realizar cualquier tipo de actuación.

Volviendo a la cuestión de la evaluación de procesos y la evaluación de impacto, existe una diferencia fundamental entre ambas. En la evaluación de procesos se describe el mundo como es y las cosas como pasan. En la evaluación del impacto lo comparación se produce con algo que no existe («qué habría sucedido si no se hubiera puesto en marcha la actuación o programa») y, por tanto, es muy difícil generar un grupo de comparación que nos dé indicación de qué hubiese ocurrido en ausencia del programa o proyecto.

Con todo, es necesario configurar un grupo de comparación adecuado con individuos que, excepto por el hecho de no ser beneficiarios del programa, sean similares a los que han recibido el programa. En general, las personas que se benefician de un programa se seleccionan especialmente (por ser pobres, por estar más motivados, etc.) y, por tanto, no se pueden comparar simplemente con las personas que no son beneficiarias. Un grupo de comparación formado por personas que no han participado en el programa no serviría para realizar una evaluación de impacto apropiada pues, en general, no se parecerán a las personas tratadas: serán más ricas y, por tanto, muy diferentes a los individuos con baja renta beneficiarios de programas contra la pobreza; o tendrán más motivación, y, por tanto, serán muy diferentes a los desempleados que no han querido participar en un programa de formación ocupacional.

Otra posibilidad para generar el grupo de comparación es tomar los mismos individuos de antes del programa y compararlos con su situación observada después de participar en el programa.² Este procedimiento no será, en la mayoría de los casos, informativo sobre el efecto del programa pues entre el período anterior a la recepción de los beneficios del progra-

² Este procedimiento se denomina *estimación en diferencias* o, en su versión más sofisticada, *diferencias en diferencias*.

ma y el posterior han podido suceder muchas otras cosas (además de la participación en el programa que se desea evaluar). Por ejemplo, en Camboya la situación ha cambiado mucho después de la privatización. Pero no sólo la educación se ha visto afectada: muchos otros aspectos de la economía han cambiado. Por tanto, cuando analizamos la diferencia entre el antes y el después no estamos seguros de que sea el programa de reformas educativas el causante de los efectos que observamos en el sistema educativo.

Una solución al difícil problema de generar un grupo de comparación es la aproximación experimental: el programa se asigna aleatoriamente en un determinado grupo, creando un grupo tratado y un grupo de control que son totalmente comparables. En la evaluación de intervenciones educativas la aleatorización se produce en los colegios.

3.3. La evaluación aleatorizada de las intervenciones educativas

En este epígrafe se presentan los resultados más importantes que se conocen en la literatura, a partir del procedimiento de evaluación aleatorizada, sobre los efectos de intervenciones educativas en países en vías de desarrollo. Podemos dividir dichos efectos en dos grupos: el impacto de las actuaciones sobre la participación (asistencia a las escuelas) y el impacto sobre la calidad educativa. En el segundo caso, los efectos sobre la calidad, se consideran dos tipos de actuaciones: las que afectan a los *inputs* (como por ejemplo la utilización de *flip charts*, de la que se ha hablado en el primer capítulo de este libro) y las que afectan a las estrategias de reforma global, que son mucho más complejas de analizar.

3.3.1. Intervenciones para aumentar la participación en el sistema educativo

En general, habría dos formas de conseguir que los niños acudan a la escuela: aumentar los beneficios que se obtienen por un mayor nivel de educación o bajar el coste de acudir a la escuela. Para reducir el coste, hay que considerar dos elementos fundamentales para que un niño vaya a la escuela.

Uno es el coste de pagar la matrícula, los uniformes y los libros. Por otra parte está el coste de la oportunidad: cuando el niño está en la escuela no gana dinero para ayudar a mantener a su familia. Con el programa Progresá, mencionado en el segundo capítulo, se intenta reducir el coste de oportunidad compensando a los niños por el tiempo que invierten en ir a la escuela (de esta forma, el coste de oportunidad se puede convertir en negativo). Si el niño acude a la escuela una determinada proporción del tiempo, la familia obtiene una compensación económica significativa. El programa Progresá consiguió un aumento del 3,4% en el número de alumnos matriculados en primaria, con un impacto aún mayor en los niveles secundarios. El impacto fue menor en primaria porque los niños ya estaban en la escuela y, seguramente, hubieran ido de todas formas a la escuela primaria, independientemente de la compensación ofrecida por el programa Progresá.

Otro programa, que se ha mencionado en el segundo capítulo, es el de la *distribución de uniformes escolares en Kenia*. Como parte de un proyecto de lucha contra el sida se utilizaron las escuelas como centro para transmitir información sobre la enfermedad. Pero, para poder explicar a los jóvenes los mecanismos de transmisión de la enfermedad, era preciso que acudieran a clase. Sin embargo, como se ha señalado con anterioridad, en muchos países en vías de desarrollo, la participación en la escuela es baja. En el caso de Kenia la educación es gratuita salvo por el coste de los uniformes escolares que todo niño debía adquirir.³

Los uniformes no son baratos. Cuestan unos cuantos dólares en un país donde el producto interior bruto per cápita es muy pequeño. Los uniformes escolares se distribuyeron entre 10.000 estudiantes de sexto y séptimo curso en 163 escuelas elegidas aleatoriamente. El resultado de esta actuación de distribución gratuita de uniformes fue una disminución del 14% de la tasa de abandono de las niñas y un 16% en la tasa de los niños.⁴

³ De facto, ningún niño es expulsado del colegio por no llevar uniforme, pero el estigma social asociado a no llevar uniforme genera suficiente presión como para que todos los niños lo lleven.

⁴ Este resultado se obtiene al comparar el grupo de estudiantes en las escuelas seleccionadas aleatoriamente con un grupo de comparación de escuelas que no fueron seleccionadas para participar en el programa de provisión gratuita de uniformes escolares.

La *distribución de comida* entre los alumnos es otra forma de fomentar la participación aumentando los beneficios o reduciendo el coste de oportunidad de asistir a clase.⁵ En este caso, existen varios experimentos. En Kenia se ha evaluado el efecto de ofrecer desayunos gratuitos a los alumnos de preescolar. El resultado fue un aumento de la participación en la escuela del 30%. En la India se les daba cereales a los niños cuando iban a la escuela. Ésta es otra forma de reducir el coste de oportunidad de ir a la escuela. Es un incentivo más para participar en el sistema educativo.

En muchas ocasiones los niños no asisten a clase por problemas sanitarios. No se trata, estrictamente hablando, de un problema de coste de oportunidad o de beneficios no percibidos. Los niños están simplemente enfermos, o demasiado débiles, para andar la distancia que les separa de la escuela. El *absentismo escolar a causa de las enfermedades*, y en particular de los parásitos intestinales, es muy importante en los países del tercer mundo. Por tanto, en principio, otra forma de aumentar la participación en los colegios sería mejorando la salud de los estudiantes. En el capítulo anterior se comentó el efecto de un programa de administración de medicinas en las escuelas de Kenia para reducir la incidencia de los parásitos intestinales. Este programa también tuvo un efecto significativo sobre la educación: el absentismo escolar se redujo un 25%, lo que es equivalente a 0,15 años adicionales de educación para los estudiantes que participaron en el programa.

Otro tipo de programas que ha sido utilizado para aumentar la participación en la escuela y reducir las tasas de abandono son los *programas de becas para niñas*,⁶ en función de los resultados que tuvieran en exámenes de respuesta múltiple. La idea es simple: si se da una beca a una niña en función de un buen resultado escolar en una prueba que se hace a final del

⁵ O, al menos, haciendo que dichos beneficios sean más claramente percibidos y más inmediatos. Obviamente, la educación tiene unos beneficios futuros en los individuos que participan muy superiores a los beneficios inmediatos de recibir una comida al día. Sin embargo, muchos destinatarios pueden no percibir dichos beneficios como importantes en una situación de extrema necesidad.

⁶ El motivo de utilizar a las niñas como objetivo de este programa es su mayor tasa de abandono de la escuela.

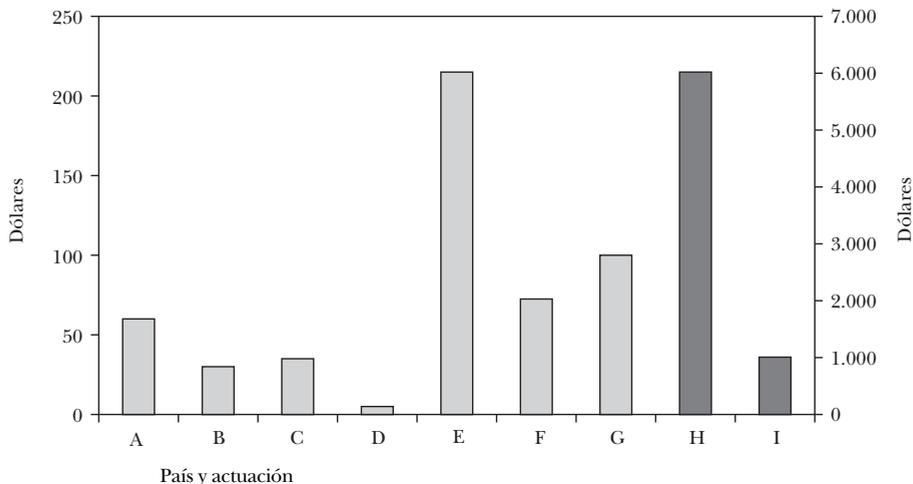
año, entonces van a la escuela con más frecuencia. Pero, además, los niños también asisten con mayor frecuencia en una especie de efecto colateral inducido.

De lo señalado en los párrafos anteriores, podemos comprobar que existen múltiples procedimientos para estimular la participación de los niños en el sistema educativo y reducir el absentismo y el abandono escolar. Pero los indicadores anteriores sobre el impacto no nos indican la efectividad relativa de los diferentes procedimientos en términos de beneficios y costes. Además, los resultados son difíciles de comparar entre países y entre actuaciones. La metodología apropiada para comparar los efectos de programas muy diferentes en distintos países precisa un indicador comparable. Normalmente, dicho indicador se basa en el coste de un año extra de educación inducido por los programas que se están comparando. Esta medida combina el coste de las actuaciones con sus beneficios (o el impacto sobre la educación). Esta magnitud es diferente del coste por niño que participa en el programa, y que depende del número de niños en el margen. El coste por alumno es un indicador irrelevante, pues pueden existir programas con un coste bajo por alumno pero que tengan un efecto pequeño sobre su participación escolar (los niños habrían ido de todas formas al colegio) o programas con un coste por alumno elevado que tienen un impacto muy grande sobre la reducción del abandono escolar. Por tanto, es necesario calcular un indicador relativo de costes y beneficios. Una forma de hacerlo es calcular el coste por año extra de educación. Dicho indicador se obtiene dividiendo el coste total del programa por el número de años/niño que se ha conseguido atraer a la escuela.⁷ Éste sería un indicador adecuado para comparaciones entre programas y entre países. Este indicador, además, puede proporcionar algunas sorpresas como hemos visto en los dos capítulos anteriores.

El gráfico 3.1 muestra el coste por año adicional de educación que se ha logrado inducir a partir de las diferentes actuaciones analizadas anteriormente. Los resultados del programa Progreso aparecen en las columnas de tono más oscuro y se miden en la escala de la derecha. El resto se

⁷ Esta segunda magnitud es la calculada en la evaluación del impacto del programa.

GRÁFICO 3.1: Coste por año adicional de educación inducido



- A: India. Profesor de apoyo.
 B: India. Proporcionar hierro y desparasitación.
 C: Kenia. Provisión de comidas escolares gratis.
 D: Kenia. Medicinas para desparasitación.
 E: Kenia. (Busia y Teso). Becas para niñas.
 F: Kenia (Busia). Becas para niñas.
 G: Kenia. Uniformes gratis.
 H: México. Subvención por asistencia de hijos a primaria (escala derecha).
 I: México. Subvención por asistencia de hijos a secundaria (escala derecha).

mide en la escala de la izquierda. El primer punto importante que hay que resaltar es que el programa Progresá, como se señaló anteriormente, fue muy exitoso, pero su efectividad en términos de coste es menos espectacular. El coste por año inducido de educación primaria alcanza los 6.000 dólares. El motivo de tan baja eficiencia es que el 96% de los niños va a cursar enseñanza primaria de todas formas, sin necesidad de ningún incentivo especial. Por tanto, es como si la subvención del programa Progresá se le diera a todas las familias, independientemente de si sus hijos van a la escuela o no. En secundaria la eficacia es mucho mayor pero aún cuesta 1.000 dólares un año adicional de enseñanza, lo que sigue siendo comparativamente muy gravoso. No obstante, es importante considerar que Progresá es un programa fundamentalmente social de transferencias. Como tal, Progresá es un excelente programa que se ha exportado a otros países latinoamericanos. Pero, como forma de incentivar la participación en la escuela, Progresá no es un mecanismo muy eficiente.

En términos de eficiencia el mejor programa es el de desparasitación: cuesta tan sólo 3,5 dólares por año extra de educación, reflejando el bajo coste de los medicamentos antiparasitarios y el enorme impacto sobre la asistencia a las escuelas que tiene la reducción de la incidencia de los parásitos intestinales. Las becas para niñas en Kenia tienen una eficiencia muy distinta en los dos distritos analizados (Busia y Teso). En todo caso, los dos indicadores muestran una eficiencia muy inferior a las medicinas contra los parásitos intestinales. La administración de hierro y medicinas antiparasitarias se ha mostrado bastante eficiente para reducir el absentismo escolar en la India, aunque su coste es superior al mostrado en Kenia. Las comidas en la escuela tienen un coste de más de 30 dólares por año de educación inducido mientras los uniformes escolares suponen 100 dólares por año adicional de educación.

3.3.2. Intervenciones para mejorar la calidad de la educación

En el epígrafe anterior se ha considerado el efecto de muy diversas actuaciones sobre el absentismo escolar. Pero, como se señaló en la introducción, una cosa es conseguir que los alumnos acudan a las escuelas y otra distinta es que consigan aprender algo mientras están allí. Se ha comentado con anterioridad la baja calidad de la enseñanza en países del tercer mundo. Por ejemplo, en Uttar Pradesh, un estado de la India, el 41% de los niños testados por la ONG Pratham no pudo leer un párrafo simple. Sin embargo, los padres piensan que, como van al colegio, son capaces de leer. En concreto sólo el 21% de los padres pensaba que sus hijos no sabían leer.

Por tanto, una vez que hemos conseguido por alguno de los procedimientos comentados con anterioridad, u otros que puedan resultar tan efectivos, que los niños acudan a la escuela, ¿aprenden algo? Es evidente que la escuela les proporciona un nivel de socialización importante para su desarrollo personal, pero si aprendiesen algo la asistencia a la escuela tendría un impacto más adecuado a sus objetivos iniciales.

Encontrar evidencia sobre el nivel de aprendizaje es más difícil que comprobar la asistencia a las aulas. Hasta el momento se han realizado una serie de evaluaciones aleatorizadas sobre actuaciones para mejorar la calidad de las escuelas. Aunque muchos de estos resultados preli-

minares han sido desalentadores, hemos aprendido bastantes cosas sobre el diseño de nuevos programas y mejorado la percepción de qué podría funcionar. No solamente hay que comprender si las intervenciones funcionan o no, sino que hay que comprender por qué funciona si éste es el caso. Ha habido un gran esfuerzo investigador para intentar comprender por qué actuaciones que a priori deberían funcionar, en realidad, no están funcionando. La siguiente generación de estudios intenta utilizar estos nuevos conocimientos para realizar la selección de aquellos procedimientos efectivos y eficientes.

En el caso del análisis de la calidad educativa a partir de evaluaciones aleatorizadas, podemos comparar coste y efectividad de diferentes intervenciones en términos del presupuesto que habría que gastar para mejorar el nivel de aprendizaje de los niños, expresados en unidades constantes: dólares por desviaciones estándar respecto al rendimiento que se hubiera producido de no haber realizado la intervención. En este epígrafe analizamos dos tipos de actuaciones: la provisión de más *inputs*, o factores en el proceso educativo, y la provisión de incentivos.

3.3.2.1. *Aprendizaje: más de lo mismo no es siempre mejor*

Los resultados iniciales en el ámbito de la mejora de la calidad de las escuelas a partir de utilizar más *inputs* o aumentar los medios tradicionales son desalentadores. Glewwe, Park y Zhao (2006) muestran que tanto los programas de provisión de libros de texto como los *flip charts* fueron poco efectivos para mejorar la calidad de las escuelas en Kenia. Los libros de texto sólo tienen un efecto positivo en el aprendizaje de los mejores alumnos, pero no tienen efecto en los alumnos de nivel medio o bajo. Tampoco los *flip charts* resultaron ser un buen procedimiento para mejorar el rendimiento académico. Finalmente, en un experimento realizado en la India, Banerjee et al. (2007) muestran que los profesores de apoyo no tienen un efecto sobre la mejora del aprendizaje de los alumnos, medido a partir de los resultados de las pruebas cognitivas, en comparación con la situación habitual (sin profesor de apoyo). La conclusión general puede ser la siguiente: la estrategia del *más de lo mismo* (más *inputs* tradicionales) no funciona. ¿Funcionaría un cambio en la pedagogía o en los incentivos, o en ambos? En el resto del epígrafe se discuten nuevos enfoques pedagó-

gicos e incentivos para la mejora del rendimiento académico de los estudiantes.

El programa de Pratham en la India es uno de los pocos desarrollados de manera experimental y, por tanto, evaluable. Los resultados de sus programas son interesantes porque se basan, como estrategia, en que puedan ser reproducibles. Pratham desarrolla dos tipos de programas: uno de educación remedial o clases de apoyo y otro de aprendizaje asistido por ordenador (*Computer Assisted Learning* o CAL).

El primer programa pedagógico de educación remedial es el llamado Balsakhi (*amiga de los niños*). El programa se desarrolló en 20 ciudades simultáneamente y fue evaluado en dos de ellas (Mumbai y Vadodara). El programa consistía en dar formación a mujeres jóvenes de la comunidad, que cobraban el salario mínimo (10 dólares al mes). Son personas que de otra forma no estarían trabajando, pues en la sociedad de la India se sigue considerando que las mujeres no deben trabajar y sólo se ve aceptable que trabajen en una escuela. Por tanto, hay muchas mujeres jóvenes disponibles. Además, la estrategia del programa se sustentaba en proporcionar una atención individualizada y *no represiva* a los niños y niñas que iban retrasados en su aprendizaje. El razonamiento era que los niños se sentirían más cómodos con mujeres de su propia comunidad que con profesores que, a menudo, tienen un origen muy diferente. Asimismo, las *balsakhi* servirían de intermediarias entre los padres y los profesores.

El proceso comenzaba por ofrecer a estas mujeres jóvenes, que habían completado la enseñanza secundaria, una formación adicional de dos semanas con una meta muy sencilla y clara: pedagogía simple para enseñar a los niños en pequeñas clases de alumnos de nivel 3 o 4, pero que no tenían conocimientos del nivel 1 (básico). Es decir, ayudar a niños que no pueden leer ni saben sumar. El experimento consistía en tomar a los 10 o los 20 estudiantes con peor rendimiento, sacarlos del aula y trabajar individualmente con ellos para intentar ponerlos al día. Esto es lo que se entiende por clases remediales o de apoyo. Básicamente, se saca a los alumnos de las aulas durante dos horas al día y se les da clases de apoyo en temas relacionados con la lectura y las matemáticas.

Inicialmente, algunos plantearon preocupaciones éticas en los experimentos de Mumbai y Vadodara. Una forma de resolverlas es intentar ha-

cer los experimentos de forma global. El programa incluía todas las escuelas de *balsakhi*, pero fue aleatorizado para trabajar con niños de tercero o cuarto. Los estudiantes de tercero que no recibieron clases de apoyo serían de grupo de comparación. Entonces, todo el mundo recibe algo, pero seguíamos teniendo un grupo control de comparación y a la vez podíamos aprender algo.

A diferencia de los profesores de apoyo normales (que como se señaló antes no tiene efecto en el aprendizaje), el grupo objetivo que participó en la formación remedial del experimento Balsakhi tuvo un gran impacto sobre el aprendizaje. Si se compara la clase tratada con la clase de control, la diferencia es 0,6 desviaciones estándares entre los estudiantes que tenían peores notas antes de la actuación. Éste es un impacto enorme si tenemos en cuenta su bajo coste, comparado con otras actuaciones que sólo habían conseguido un aumento de 0,2 desviaciones estándares en los resultados de los test de conocimientos. Estamos hablando de un resultado verdaderamente impactante. En general, la iniciativa Balsakhi mejoró el aprendizaje, medido por pruebas cognitivas, en 0,14 desviaciones estándar el primer año y 0,25, el segundo año. No hubo efecto entre los estudiantes con mejores notas que estaban en una clase *balsakhi*, pero que no recibían el tratamiento. Esto indica que los efectos desbordamiento de la mejora de los alumnos necesitados de educación remedial no se trasladaron a los alumnos con mejores notas. El coste total de esta evaluación y del programa fue de 30.000 dólares al año.⁸

Ha de notarse que el efecto de este programa podría estar causado por el hecho de que al retirar 10 o 20 alumnos de las clases del grupo tratado quedaban menos alumnos en el aula y podría suceder que esa disminución del tamaño de la clase beneficiaría a los alumnos que se quedaban en la clase. Por este motivo, era importante aleatorizarlo, y luego hay que combinar el método de aleatorización y algún método econométrico. De esta forma, se podría efectivamente identificar el efecto que era atribuible

⁸ No tiene nada que ver con los 12 millones de dólares de coste del proyecto STAR comentado en el capítulo 1. Los experimentos aleatorizados en países desarrollados son mucho más caros que las evaluaciones científicas que se pueden desarrollar en países en vías de desarrollo.

a los niños que trabajaron con el profesor de apoyo (formación remedial). La diferencia entre ambos grupos después de la intervención era espectacular: el efecto observado en los niños era de 0,6 desviaciones estándar y se debe casi exclusivamente a la mejora de los alumnos que fueron a las clases de apoyo: estos alumnos mejoraron sus puntuaciones en los test en una desviación estándar, mientras que los estudiantes que se quedaron en la clase (del grupo tratado) no mostraron ninguna mejora en su nivel de aprendizaje. Por tanto, este experimento proporciona un resultado secundario interesante, compatible con otros anteriormente señalados: reducir el número de estudiantes en la clase no va a ser efectivo en el contexto de la India.

Otro programa remedial desarrollado por Pratham en la India se llama Read India. El objetivo es conseguir que en aldeas rurales de la India los estudiantes sean capaces de leer. Para enseñarlos se formaron voluntarios y otras personas de la comunidad. La evaluación del proyecto muestra que los estudiantes en las aldeas tratadas tenían una probabilidad superior de poder leer que los alumnos de las aldeas de control.

El segundo tipo de programas de Pratham se basa en el aprendizaje asistido por ordenador. Este cambio pedagógico permite que cada niño trabaje a su propio nivel. En la India es importante que cada alumno reciba la educación adecuada para su nivel, pues el sistema habitual tiende a mezclar en las escuelas públicas a estudiantes de muy diferente nivel de formación. De hecho, ningún estudiante repite curso hasta que finalmente acaba el itinerario o deja de asistir a clase. Y los profesores siguen su programa de estudios independientemente del nivel de los alumnos que tienen en su clase. Por tanto, la enseñanza individualizada con ajuste preciso entre el nivel del alumno y el nivel de la enseñanza que recibe era potencialmente una actuación que podría funcionar bien para mejorar el aprendizaje de los alumnos. El impacto de este programa se demostró bastante fuerte, suponiendo un incremento de 0,35 desviaciones estándar en la distribución de las calificaciones de los test de matemáticas en el primer año. En el segundo año la mejora fue de 0,47 desviaciones estándar.

Otro caso interesante es la distribución de gafas entre alumnos en China (Glewwe, Park y Zhao 2006). Aproximadamente el 10% de los estudiantes de primaria en países en vías de desarrollo tiene una visión deficiente, pero

muy pocos niños llevan gafas. Sin embargo, en el pasado no ha habido casi investigación sobre el impacto de la falta de visión de los niños en su rendimiento académico. Glewwe, Park y Zhao (2006) realizaron un experimento aleatorizado en China que comenzó en 2004. La muestra incluye 165 colegios de dos zonas de la provincia de Gansu. Las escuelas fueron divididas aleatoriamente en 103 colegios que recibieron gafas (para estudiantes de tercero, cuarto y quinto) y 62 colegios que sirvieron de grupo de control. Los resultados preliminares, referidos al primer año de funcionamiento del programa, indican que las gafas consiguieron aumentar los resultados de los test entre 0,15 y 0,30 desviaciones estándar de la distribución global de notas.

3.3.2.2. Actuaciones dirigidas a modificar los incentivos

Otra posibilidad para mejorar el rendimiento de los alumnos es cambiar los incentivos tanto de estudiantes como de profesores. Las actuaciones sobre los incentivos de los alumnos suponen, por lo general, un proceso lento. Si los programas tienen que ser implantados por políticos impacientes, éstos preferirán actuar sobre otros actores del sistema educativo, como profesores y padres. Actuar sobre los incentivos de los padres (para que colaboren con los niños en la realización de los trabajos escolares o les ayuden a estudiar) es complejo pues muchos de los padres no tienen los conocimientos para poder ser de utilidad en estas tareas de sus hijos, ya que ellos mismos no han asistido al colegio.

Las posibilidades para incentivar a los profesores pasan por pagas extra por reducir el absentismo o, directamente, por mejoras en los resultados de los alumnos. Por tanto, el objetivo es incentivar a los profesores a que acudan a dar sus clases y, una vez allí, a que enseñen efectivamente a sus estudiantes.

El primer programa para proporcionar *incentivos para reducir el absentismo de los profesores consistía en darles una cámara y pedirles que sacaran una foto de ellos y de los niños dos veces al día, al principio y al final del día lectivo*. El cobro de incentivos se basaba en la realización efectiva de clases. En la zona donde se realizó el experimento, el absentismo de los profesores alcanzaba el 40%. El programa no era muy sofisticado y, sorprendentemente, fue muy bien acogido por los profesores. En las entrevistas realizadas a los mismos señalaban que el programa permitía que cobraran un

mejor salario los que realmente hacían su trabajo y no que todos cobraran muy poco con independencia de su asistencia a las clases. El efecto fue inmediato y muy grande: la tasa de absentismo de los profesores bajó al 20%, una reducción que supone un 50% del valor inicial. El efecto de la asistencia de los profesores fue que, después de un año, los resultados de los estudiantes mejoraron en 0,17 desviaciones estándar.

Este programa era bastante específico porque lo hicimos con las cámaras que se les daban directamente a los profesores, y con la ayuda de un ordenador que recibía las fotografías. No había mediación directa, lo que era uno de los elementos positivos destacados por los profesores en las entrevistas. Por el contrario, en muchas situaciones reales, cuando se intenta incentivar la presencia de los maestros se hace a través de los directores de los colegios. Éstos tienen mucha presión, pues los controlados suelen ser sus propios compañeros. Por tanto, los directores suelen evitar molestar a sus profesores. En Kenia, por ejemplo, se intentó un *programa para incentivar la asistencia de los profesores a las aulas utilizando a los directores como controladores del absentismo*. Debían tomar nota de los profesores que asistían al colegio y los que no asistían cada día. Los resultados de la evaluación mostraron que los directores apuntaban a todos los profesores, hubieran asistido o no a dar su clase. Por lo tanto, el absentismo no había disminuido, aunque las faltas reportadas indicaban que la presencia en las aulas de los profesores había aumentado. Esto indica que el procedimiento para controlar el absentismo de los profesores e incentivar su reducción no es neutral.

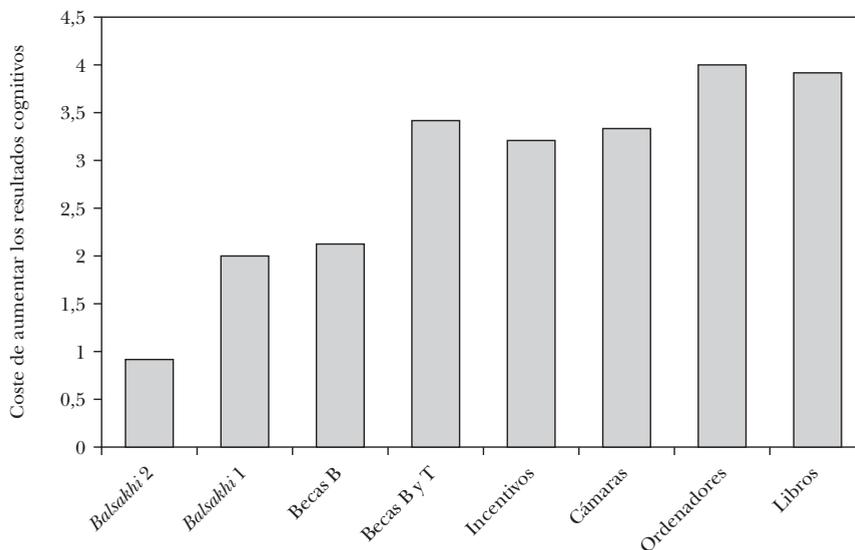
Otra posibilidad de *incentivar a los profesores consiste en pagar incentivos en función de los resultados obtenidos por los alumnos*, medidos a partir de las calificaciones en pruebas cognitivas. Los resultados de este programa en Kenia fueron positivos en el corto plazo (los resultados de las pruebas mejoraron) pero el aprendizaje en el largo plazo no mejoraba. Una posible explicación es la preparación específica para las pruebas: los profesores, en lugar de proporcionar la enseñanza que corresponde, se concentran en enseñar a los alumnos a contestar test de respuestas múltiples, que son el instrumento por el que luego se medirá el rendimiento de los estudiantes y se percibirán los incentivos. Un experimento realizado en la India, en colaboración con el gobierno de la región de Andhra Pradesh, resultó más exitoso al encontrarse una mejora de los resultados en el largo plazo.

En principio, también se pueden utilizar mecanismos para *incentivar la asistencia a clase de los estudiantes*. Éste fue el objetivo del programa de *becas para niñas de Kenia*, que fue comentado en el epígrafe anterior. Dicha actuación, basada en proporcionar becas y retribuciones monetarias a las niñas que tenían mejor rendimiento escolar, redujo el absentismo de alumnos y profesores y, sorprendentemente, mejoró las puntuaciones tanto de las niñas que tenían una elevada probabilidad de ganar la beca, como de niñas que tenían notas bajas y niños, que no eran elegibles para el programa (Kremer, Miguel y Thornotn 2005). Estos resultados indican que existen grandes externalidades en los premios por mérito, al contrario de lo que piensa la gente, más preocupada por las cuestiones de equidad, y proporciona una justificación adicional de los subsidios públicos a la educación.

Otra posibilidad para *incentivar a los jóvenes a ir a la escuela es informarles sobre las mejoras que tendrán en el mercado laboral después de finalizar sus estudios* (ya sea en términos de menor nivel de desempleo o de mayor salario). Esta información ayuda a la gente atomar una decisión informada. En la República Dominicana, Jensen (2006) muestra los resultados de un experimento aleatorizado sobre el efecto de informar a los estudiantes de secundaria de sus posibilidades laborales cuando terminaran los estudios. El efecto fue una reducción considerable del abandono escolar.

Al igual que las actuaciones comentadas en anteriores epígrafes, los efectos de las innovaciones pedagógicas y los incentivos también pueden evaluarse de manera similar, calculando costes y beneficios en términos de las desviaciones estándar de mejora en los test. En concreto, se puede calcular el coste de aumentar los resultados de las pruebas cognitivas en 0,1 desviaciones estándar. El gráfico 3.2 muestra este resultado para los programas comentados con anterioridad: el programa de profesores de apoyo locales (*balsakhi*) en el primer año y en el segundo año, el programa de becas para niñas en Busia (becas B) y en el conjunto de Busia y Teso (becas B y T), el programa de incentivos a los profesores, el programa de cámaras para comprobar la asistencia de los profesores a la escuela, el programa de enseñanza basada en ordenadores y, finalmente, la provisión de libros. De los resultados que aparecen en el gráfico 3.2 se concluye que el mecanismo más eficiente fue el *balsakhi*, o profesoras de apoyo locales, en el segundo año, seguido del mismo programa en su primer año. La siguiente posición, en tér-

GRÁFICO 3.2: Eficiencia de las actuaciones destinadas a mejorar los incentivos



minos de eficiencia, la ocupan las becas para niñas en el distrito de Busia. La eficiencia del mismo programa en Teso fue muy inferior. Las ONG han tenido una historia de dificultades en la aplicación de nuevos programas en Teso, donde, además, el grupo étnico dominante es diferente al dominante en el distrito de Busia. Estas características explican que la eficiencia del mismo programa en dos distritos diferentes de Kenia fuera tan diferente. Los programas de incentivos y de utilización de cámaras para controlar la asistencia de los profesores a las escuelas no fueron tan eficientes en la mejora del aprendizaje de los estudiantes. Por último, los programas más caros en términos de beneficios y costes fueron los basados en el aprendizaje a partir de ordenadores y la provisión de libros de texto.

3.3.3. Estrategias de reforma global del sistema

En el epígrafe anterior se han comentado procedimientos de cambio de los métodos pedagógicos (profesoras locales de apoyo, ordenadores, etc.) e incentivos para mejorar la calidad de la educación que reciben los estudiantes que acuden a las clases en determinadas localizaciones. En este último epígrafe se considera el efecto de cambios más globales en los

sistemas educativos, como pueden ser la financiación mediante cheques escolares o la descentralización.

3.3.3.1. *Los cheques escolares*

La iniciativa de Colombia llamada Programa de Ampliación de la Cobertura de la Educación Secundaria (PACES) es uno de los experimentos de provisión de cheques escolares más grandes que se han realizado hasta el momento. El objetivo de la iniciativa fue proporcionar cheques escolares a más de 125.000 estudiantes. Dichos cheques cubrían más del 50% del coste de la educación secundaria en una escuela privada. Los cheques eran renovables siempre que los estudiantes mantuvieran un buen nivel académico. El programa PACES se puso en funcionamiento a finales de 1991 como parte de un proceso de descentralización para expandir la provisión privada de servicios públicos y expandir las tasas de participación en la educación secundaria. Un problema importante en Colombia es que, a diferencia de la educación primaria donde las tasas de participación son del 89%, la educación secundaria sólo alcanza al 75% de los jóvenes en el grupo de edad relevante (en 1993). En términos distribucionales la baja participación se concentra en las familias más pobres: el 78% de los niños de las familias en el quintil más pobre de la población participa en la escuela primaria pero sólo un 55% lo hacen en la escuela secundaria.

El programa PACES estaba fundamentalmente dirigido a los niños de familias con rentas bajas, y los cheques sólo podían ser utilizados por niños que habían acudido a una escuela pública en primaria. Otra particularidad del programa PACES es que la asignación de la gran mayoría de los cheques se hizo de forma aleatoria a partir de una lotería. Este procedimiento permite la utilización de métodos de investigación cuasiexperimental para evaluar el efecto sobre el nivel de participación en la educación secundaria, y otros resultados, comparando los ganadores de la lotería con los perdedores.⁹ Angrist et al. (2002) presentan el análisis de los resultados del programa. Los

⁹ Por motivos presupuestarios, no era posible otorgar cheques escolares a todos los niños y niñas que serían susceptibles de cumplir las condiciones necesarias para recibirlos. La extensión del programa en años posteriores fue abarcando un grupo cada vez mayor de potenciales beneficiarios.

ganadores de los cheques escolares tenían una probabilidad 10% superior a los perdedores de finalizar el octavo grado, fundamentalmente porque repetían menos cursos. Los ganadores de los cheques escolares también tenían mejores notas en los test cognitivos (0,2 desviaciones estándar superiores a los perdedores). El análisis coste-beneficio muestra que estas actuaciones fueron muy eficaces: los beneficios por ganador de la lotería son superiores en 24 dólares a los costes adicionales para el gobierno de proveer cheques escolares en lugar de puestos escolares públicos.

3.3.3.2. *La descentralización*

La descentralización tiene como objetivo el que los perceptores de los beneficios tomen parte en las decisiones, ya sea en las localidades o en las regiones. El desarrollo impulsado por la comunidad (CDD, por sus siglas en inglés) es la versión más extendida de la descentralización y consiste en transferir a las organizaciones comunitarias el control sobre las decisiones de planificación e inversión de los recursos para el desarrollo, al tiempo que favorece la asociación a largo plazo de los gobiernos locales, las organizaciones no gubernamentales y otros asociados. En la India el programa Sarva Shiksha Abhiyan (SSA), que puede traducirse por Campaña para la Educación Universal, incluía la posibilidad de dar financiación adicional a los colegios que intentaran mejorar el control de la comunidad a partir de la formación de comités educativos formados por personas de la comunidad, aunque sin la participación de los profesores. Los resultados de esta medida son poco convincentes.

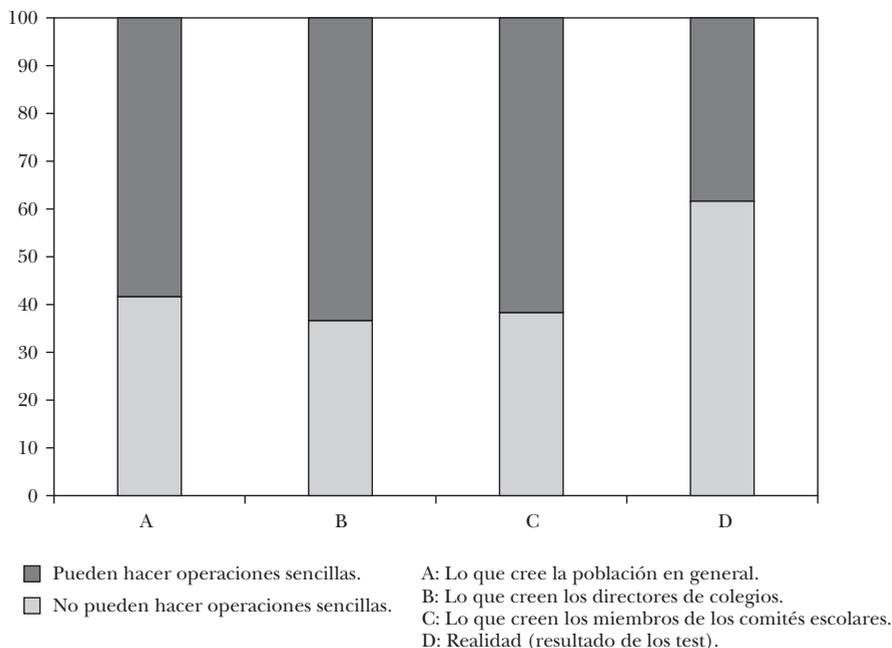
En general, las experiencias que se conocen son poco esperanzadoras. Los comités escolares en Kenia no han tenido el efecto esperado. En general, las medidas para aumentar la descentralización en la toma de decisiones han producido resultados poco alentadores. Por ejemplo, un proyecto financiado por el Banco Mundial en Indonesia combinó el CDD con análisis técnicos para evaluar si los fondos destinados a un proyecto vial en una aldea se estaban desviando a causa de la corrupción. Olken (2007) estudia, utilizando un experimento aleatorizado, el efecto sobre los fondos desviados en la construcción de carreteras en Indonesia. El procedimiento implica estimar los precios y las cantidades de materiales de construcción para las carreteras y comparar dichas cifras con las estimacio-

nes oficiales en los informes para cada carretera. Olken (2007) señala que anunciar un aumento en la probabilidad de que el gobierno realice una auditoría (desde una probabilidad base del 4%) reduce el desvío de recursos (corrupción) en un 8%. La amenaza de una auditoría era particularmente efectiva en aldeas que tenían que reelegir a sus líderes y donde los resultados de la anterior elección sugerían una victoria con un escaso margen, lo cual indica que «el seguimiento externo y la responsabilidad democrática pueden ser complementarios». Por el contrario, un aumento en la participación de la comunidad en el proceso de control de la utilización de los recursos para la efectiva construcción de carreteras, tuvo un impacto muy limitado: sólo fue efectivo cuando los problemas de *free-riding* eran reducidos y cuando la corrupción en las élites locales era limitada. Olken (2007) sugiere que el procedimiento tradicional de control (de arriba abajo) es más efectivo para reducir la corrupción que el procedimiento basado en ceder el control a las comunidades (de abajo arriba). Esto sucede incluso en ambientes con un elevado nivel de corrupción.

Por el contrario, Reinikka y Svensson (2005) encuentran que una campaña en los periódicos de Uganda sobre el volumen de fondos públicos que debería llegar a cada aldea para la construcción de escuelas de primaria tuvo un efecto positivo: redujo el desvío de fondos (corrupción), que inicialmente suponía el 80% del total de los recursos, aumentó la asistencia a la escuela y mejoró las notas de los alumnos.

Este resultado no se confirma en el estudio de Banerjee et al. (2007) sobre el efecto de los comités educativos locales (Village Education Committees o VEC) en la India. Banerjee et al. (2007) estudian el efecto de los comités educativos locales en el estado indio de Uttar Pradesh, encontrando que el 92,4% de los padres desconocen la existencia de estos comités, incluso cuando se supone que forman parte de los mismos; sólo un 5% de los padres son capaces de nombrar un solo miembro de los comités; los miembros de los comités desconocen el papel tan importante que pueden jugar en la provisión de servicios educativos; la participación pública en la mejora de la educación es muy pequeña, por lo que la educación tiene un lugar bajo en el *ranking* de prioridades de la comunidad; un gran número de niños no han adquirido ni las competen-

GRÁFICO 3.3: ¿Qué saben los padres y los comités escolares sobre el nivel de aprendizaje de los alumnos?
(porcentaje de niños que pueden realizar operaciones aritméticas sencillas)



cias mínimas en lectura, escritura y aritmética pero los padres, los profesores y los miembros de los comités escolares no tienen conciencia de la magnitud del problema. El gráfico 3.3 muestra la enorme diferencia que existe entre lo que piensan los padres y los profesores sobre el nivel de conocimiento de aritmética de los estudiantes y el conocimiento real. Este gráfico muestra que la población en general cree que los alumnos tienen muchos más conocimientos que los que realmente están adquiriendo. Los directores de colegio no son tan optimistas como la población en general pero, aun así, creen que los alumnos tienen un nivel muy superior al real: los directores piensan que el 63% de los alumnos son capaces de hacer operaciones aritméticas sencillas cuando el porcentaje real es del 38%. Lo mismo sucede con los miembros de los consejos escolares locales. Por lo tanto, el desconocimiento sobre el verdadero nivel de aprendizaje de los estudiantes coexiste con un nivel elevado de apatía pública.

3.4. Conclusiones

Se está empezando a acumular evidencia rigurosa y científica sobre lo que funciona y lo que no funciona en la ayuda al desarrollo, tanto en términos de acceso a la educación como de calidad de la educación. Los resultados de estas evaluaciones ya están influyendo las políticas. Por ejemplo, el programa Progresá se ha extendido a otros países. También se está utilizando un programa de desparasitación en Uganda, tras el éxito del mismo programa en Kenia.

A su vez, se están haciendo esfuerzos para que cada vez más países introduzcan métodos de evaluación creíbles. Muchos gobiernos están interesados en utilizar la aproximación experimental para la evaluación de los programas. Por ejemplo, la iniciativa AGEPA (Amélioration de la Gestion de l'Education dans les Pays Africains) de los países francófonos de África. El sistema ya se aplicó en formato piloto en Madagascar. Después de comprobar los resultados de Madagascar varios gobiernos decidieron participar en dicho sistema (Mauritania, Mozambique, Níger, Senegal, Burkina Faso y Guinea). La idea fundamental del AGEPA es la mejora de la gestión de la educación, especialmente la primaria, utilizando criterios que consistan en la distribución de los recursos de las escuelas (especialmente los profesores) y la transformación efectiva de los recursos tangibles de la enseñanza en resultados educativos de los estudiantes. Esto está llevando a que cada vez la gente quiera poner en marcha sus propios sistemas de evaluación con mayor frecuencia. Así que la gente se está interesando y se está generando cada vez más evidencia. Por tanto, podemos empezar a comparar la eficacia en términos de coste de todas estas actuaciones. Con todo, estamos aún al principio. Es necesario acumular mucha más evidencia del mismo tipo; probar los mismos programas en muchos sitios diferentes. Por ejemplo, el programa de desparasitación se ha probado en Kenia y en la India, pero otros programas sólo se han probado en una localización. Necesitamos más evidencia sobre diferentes actuaciones y su impacto en varios países para entender mejor lo que puede funcionar, lo que realmente funciona y lo que no funciona. Para ello, es necesario analizar una combinación de diferentes análisis de los programas existentes.

Por ejemplo, hay psicólogos que están examinando muchos de los programas médicos; especialistas en educación ayudan en el análisis económico de programas educativos; etc. El objetivo último es comprender los mecanismos que explican por qué unos programas que, en principio, deberían funcionar, no funcionan y otros sí funcionan y tienen buenos resultados.

Bibliografía

- ANGRIST, J., ERIC BETTINGER, ERIC BLOOM, y ELIZABETH KING. «Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment». *American Economic Review* 92 (5), 2002: 1535-1558.
- BANERJEE, A., S. COLE, E. DUFLO, y L. LINDEN. «Remedying Education: Evidence from Two Randomized Experiments in India». *Quarterly Journal of Economics*, 122 (3), 2007: 1235-1264.
- CHAUDHURY, N., J. HAMMER, M. KREMER, K. MURALIDHARAN, y H. ROGERS. «Missing in Action: Teacher and Health Worker Absence in Developing Countries». *The Journal of Economic Perspectives* 20 (1), 2006: 91-116.
- GLEWWE, P., A. PARK, y M. ZHAO. *The Impact of Eyeglasses on the Academic Performance of Primary School Students: Evidence from a Randomized Trial in Rural China*, 2006 [mimeo.].
- JENSEN, R. «Do the perceived return to education affect schooling decisions? Evidence from a randomized experiment». Harvard University, 2006 [mimeo.].
- KREMER, M., E. MIGUEL, y R. THORNTON. De próxima aparición en *Review of Economics and Statistics*.
- OLKEN, B. «Monitoring Corruption: Evidence from a Field Experiment in Indonesia». *Journal of Political Economy*, 115 (2), 2007: 200-249.
- REINIKKA, R., y J. SVENSSON. «Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda». *Journal of the European Economic Association* 3, 2005: 259-267.

La evaluación de las intervenciones financieras en países en vías de desarrollo: ¿pueden los bancos rurales reducir la pobreza?

Rohini Pande

Profesora de Política Pública

Kennedy School of Government, Universidad de Harvard

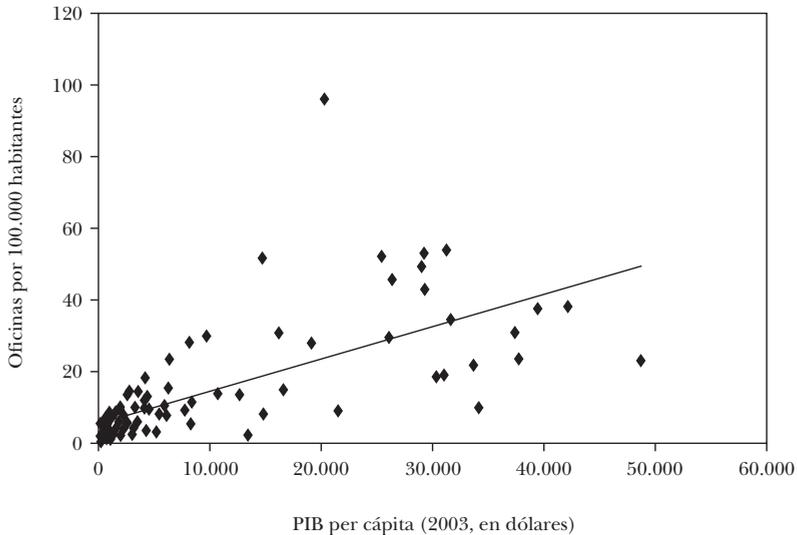
4.1. Introducción

Hay mucha literatura en la economía del desarrollo que ha tratado la importancia de los créditos en el progreso de los países con niveles bajos de renta per cápita. Los microcréditos y las microfinanzas han constituido un tema fundamental en esta literatura, y las restricciones crediticias, una limitación enorme para los países en desarrollo. Gran parte de la literatura, tanto teórica como empírica, sugiere que las restricciones crediticias limitan la capacidad de la población para salir de la pobreza. Por este motivo, muchos países han introducido programas de créditos destinados a mejorar el acceso de los pobres al sector financiero formal en los países en vías de desarrollo. Los créditos del sector informal se consideran problemáticos pues los tipos de interés son muy elevados y existe poco control.

Cuando analizamos los programas de créditos que han sido desarrollados desde los años cincuenta, para comprobar su efectividad, es muy difícil llegar a un consenso sobre cómo diseñar programas que sean coste-efectivos. Dos son las causas fundamentales de la dificultad para evaluar este tipo de programas. En primer lugar, muchos programas de crédito en países en vías de desarrollo fueron diseñados para centrarse en los excluidos por parte del sector financiero formal. Como resultado, si nos centramos en los programas de créditos para los más pobres es poco probable que podamos identificar, como resultado del programa, el efecto causal de los créditos en la economía de las personas desfavorecidas o las regiones pobres. El problema de identificación se puede caracterizar de una manera sencilla: sabemos cómo afectan los créditos a la economía de las familias más pobres pero no tenemos un punto de referencia: no sabemos el efecto que tendrían dichos créditos en la población general. Al no disponer de un grupo de control, es difícil evaluar el impacto de las políticas de promoción crediticia enfocadas hacia las familias más desfavorecidas.

El segundo problema es que los programas de créditos en países en vías de desarrollo están altamente subsidiados, normalmente por organizaciones internacionales. Para realizar un análisis coste-beneficio, es necesario conocer los componentes del programa y, en particular, el volumen de la subvención. Esto no es fácil. Por tanto, es difícil descomponer este tipo de intervenciones para saber qué elementos hay que copiar, cuáles se deberían reproducir y cuáles no.

En este capítulo se presenta una panorámica sobre cómo la evaluación científica de los programas de créditos en países en vías de desarrollo puede mejorar nuestra comprensión sobre sus efectos. En primer lugar, se analiza, desde una perspectiva comparada, el acceso al crédito. El tercer epígrafe analiza la importancia de las restricciones de crédito en el desarrollo económico. El siguiente epígrafe presenta evidencia sobre el efecto de los bancos rurales en la India. El penúltimo epígrafe contiene un resumen de los principales resultados conocidos sobre el efecto de los microcréditos; y el último presenta las conclusiones.

GRÁFICO 4.1: Penetración bancaria y PIB per cápita

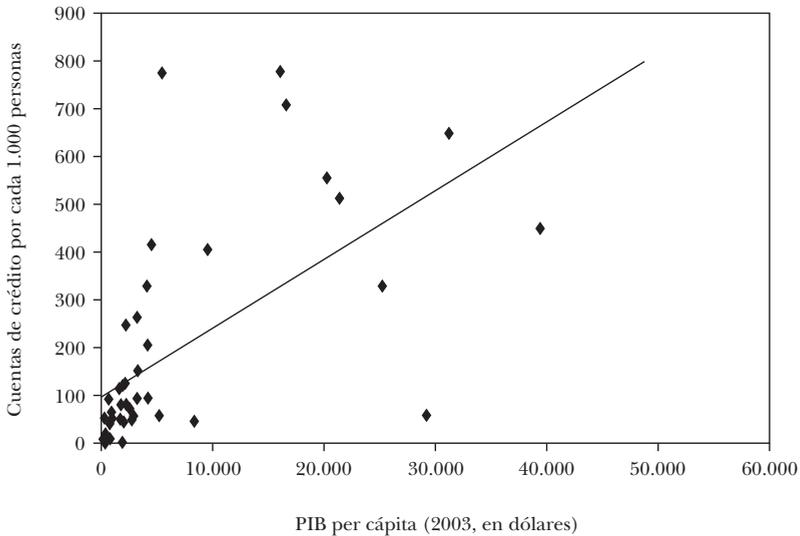
Fuente: Beck, Demirgüç-Kunt y Martínez Peria (2006).

4.2. El acceso al crédito: una comparación internacional

Los indicadores más importantes para medir la relación entre el acceso al crédito y el nivel de desarrollo son de dos tipos: los indicadores que miden el acceso a los servicios financieros, como la penetración bancaria, la distribución geográfica de los cajeros, la penetración demográfica de oficinas y cajeros, etc. Un segundo tipo de indicador es el que describe el uso real de los servicios financieros como las cuentas de crédito per cápita, los depósitos por habitante, etc.

Dentro del primer grupo, un indicador fundamental en las comparaciones internacionales de acceso al crédito es la relación entre el nivel de renta per cápita y el número de oficinas por habitante. El gráfico 4.1 contiene la relación entre el producto interior bruto (PIB) per cápita y el número de sucursales bancarias. El gráfico muestra una correlación positiva entre los niveles de renta per cápita y la densidad de oficinas bancarias. Por tanto, cuanto mayor es el nivel de desarrollo económico mayor es el número de oficinas bancarias per cápita, aunque esta relación se basa en una simple correlación y no indica una dirección particular para la causalidad.

GRÁFICO 4.2: Cuentas de crédito y PIB per cápita



Fuente: Beck, Demirguc-Kunt y Martínez Peria (2006).

Sin embargo, la existencia de oficinas bancarias no garantiza el acceso real a los servicios financieros. Un indicador del grado de utilización de dichos servicios en función del nivel de desarrollo es el número de cuentas de ahorro y los créditos concedidos.

El gráfico 4.2 muestra que los países con mayor PIB per cápita también tienen mayor densidad de cuentas de crédito y ahorro. Por tanto hay mayor acceso en los países más ricos tanto a cuentas de ahorro y más solicitudes de crédito. En los países en vías de desarrollo hay menor penetración del sector formal.

Otra prueba clara de esta relación la encontramos dentro de los países. El grupo de población con menores ingresos tiene un menor acceso al crédito: las sucursales bancarias están concentradas en zonas urbanas. Los bancos comerciales sitúan sucursales allí donde hay una mayor demanda, que es alrededor de las zonas urbanas. Por tanto, utilizar la penetración media del sector bancario en términos de oficinas sobreestima el acceso que la población tiene a los servicios bancarios. Dentro de las áreas donde hay servicios bancarios, aquellos que son más pobres tienen menos probabilidad de tener cuentas en el banco.

CUADRO 4.1: Porcentaje de familias en cada país que usan las instituciones del sector financiero formal para ahorrar

Armenia	8,9	Austria	81,4
Botswana	47	Bélgica	92,7
Brasil (urbano)	43	Dinamarca	99,1
Costa de Marfil	25	Finlandia	96,7
Guatemala	18	Francia	96,3
Guyana	14	Alemania	96,5
India (AP, UP)	48	Grecia	78,9
Jamaica	59	Irlanda	79,6
Kenia	10	Italia	70,4
Kyrgyz Rep.	1,3	Luxemburgo	94,1
Lesotho	17	Países Bajos	98,9
México (DF)	25	Portugal	81,6
Namibia	28	España	91,6
Nicaragua	4,7	Suecia	98
Pakistán	12	Reino Unido	87,7
Rumanía	23	Estados Unidos	90,9
Sudáfrica	46		
Suiza	35		
Tanzania	5		
Vietnam	8,7		

El cuadro 4.1 muestra los hogares, por países, que utilizan las instituciones bancarias formales y a las que llevan sus ahorros. En el Reino Unido, Estados Unidos, Suecia y España, el porcentaje de hogares que utilizan las instituciones formales es superior al 80%. Hay mucha variación en los niveles de formalidad bancaria en los países en vías de desarrollo. Más de la mitad de la población intenta ahorrar a través de entidades financieras formales. Pero, por otra parte, en los países desarrollados el crédito formal es más barato que el crédito no formal (no bancario). La falta de acceso al sector bancario formal hace que las familias pobres, si tiene acceso a un crédito, tengan que pagar más por el mismo crédito que podrían haber obtenido a un tipo de interés menor en el sector formal.

4.3. ¿Son importantes las restricciones crediticias para el desarrollo económico?

A partir de los resultados del epígrafe anterior, nos debemos plantear una pregunta: este escaso acceso al crédito de los países pobres ¿tiene implicaciones sobre el desarrollo económico? Que un gran segmento de la sociedad se ve excluida del sector formal de las instituciones crediticias, ¿tiene un impacto en cuanto al desarrollo? La correlación positiva entre medidas de desarrollo financiero y desarrollo económico entre países es muy sugerente, aunque difícil de interpretar. Hay mucha literatura que estudia sistemáticamente las diferentes medidas del desarrollo del sector financiero e intenta correlacionar dichas variables tanto con la pobreza como con el crecimiento económico. Las regresiones, utilizando datos de corte transversal de diferentes países, muestran que cuanto mayor es el desarrollo de las instituciones financieras mayor es el crecimiento económico y menor es el nivel de pobreza (King y Levine 1993; Honohan 2004). Sin embargo, es evidente que esta correlación no implica una relación causal: un país rico puede invertir más en actividades productivas y, a la vez, invertir más en el sector financiero. Por tanto, la evidencia macroeconómica sugiere que existe una relación entre desarrollo económico y acceso al crédito, aunque su interpretación es muy compleja por el tipo de datos utilizados. Por este motivo, es importante manejar datos microeconómicos sobre empresas y familias, a partir de encuestas o experimentos.

4.3.1. ¿Importan las restricciones crediticias?

Evidencia con el uso de datos micro

Un segundo tipo de evidencia que podemos buscar se basa en estudios sobre individuos y empresas en lugar de países. Hay evidencia de que las restricciones crediticias impiden el crecimiento de las empresas. Por ejemplo, Paulson y Townsend (2004) indican que, en una encuesta entre empresarios de Tailandia, el 54% considera que sería beneficioso expandir su negocio. De este grupo un 56% señala que no pueden expandirse debido a las restricciones crediticias. Además, las familias más ricas tienen más probabilidad de comenzar un negocio, pero aquellos que realmente lo

comenzaron tuvieron que ahorrar antes de comenzar sus actividades. Esto nos sugiere que el acceso al sector financiero, en términos de obtención de crédito, no está ampliamente desarrollado en los países menos desarrollados.

La comparación de países no puede, por sí sola, proporcionar suficiente información para diseñar políticas apropiadas para superar las restricciones crediticias. Será difícil, por tanto, tomar una correlación entre, por ejemplo, la extensión de la penetración del sector financiero y la pobreza, y hacer sugerencias en cuanto a cómo expandir estos mercados de crédito en países en vías de desarrollo, y cómo desarrollar en ellos el sector financiero. La correlación positiva entre el desarrollo financiero y el producto interior bruto per cápita puede ser el resultado de diversos factores, como, por ejemplo, las características del sistema crediticio. La investigación reciente se ha centrado en diseñar y analizar las intervenciones públicas donde es posible aislar los efectos de la expansión del crédito separadamente de los otros factores. Las técnicas basadas en experimentos naturales o en evaluaciones aleatorias se están convirtiendo en el estándar de referencia en el estudio de las microfinanzas. Los capítulos anteriores han desarrollado la metodología de los experimentos aleatorizados. En este capítulo se destaca también la posibilidad de utilizar experimentos naturales para identificar los efectos de interés en el análisis de las intervenciones financieras.

El segundo tipo de metodología de investigación en este campo son los experimentos naturales. Lo que hacemos es tomar un programa de créditos, que no ha sido diseñado de forma experimental, planteándonos que hay algunas características que son únicas del programa y que se pueden utilizar para identificar sus efectos. A menudo, en programas de gran escala, éstos son los tipos de técnicas disponibles. Puede no ser posible realizar un experimento aleatorizado en un país, como por ejemplo hacer una expansión de las sucursales bancarias, pero sí es posible utilizar algunas características sobre cómo ha ocurrido la expansión para evaluar el impacto de la presencia de una red de sucursales bancarias.

Existen dos investigaciones muy interesantes sobre la importancia de las restricciones crediticias y los microcréditos. Los métodos de análisis son los comentados en los párrafos anteriores: un experimento natural y

uno aleatorizado, respectivamente. Banerjee y Duflo (2004) analizan la financiación a empresas en la India para evaluar hasta qué punto la extensión de la inversión depende o no de la falta de créditos. En la evaluación se utiliza un experimento natural. En la India hay un programa de créditos directos, que es parte de las intervenciones sociales sobre determinados sectores, como por ejemplo las industrias de artesanía, que pueden recibir créditos subvencionados a partir de bancos comerciales. ¿Qué empresas pueden utilizar estos créditos? Los potenciales beneficiarios se deciden políticamente. Lo importante, en este caso, es que la política ha cambiado dos veces en un período de tiempo relativamente corto. Esto permite identificar, por tanto, el efecto de los créditos concedidos. La idea del trabajo es muy simple: tanto las empresas que tienen restricciones crediticias como las que no tienen, podrían absorber todo el crédito directo (formal) que puedan conseguir, dado que es más barato que el crédito obtenido en el sector informal. Sin embargo, mientras las empresas que sufren restricciones crediticias usarán el crédito para expandir la producción, las empresas que no tengan restricciones sustituirán otras fuentes de financiación más caras por este tipo de crédito. Esta idea se aplica a empresas en la India que son elegibles para recibir crédito directo después del cambio de política de 1998 y que perdieron su elegibilidad en 2000 tras la reversión de las condiciones de 1998. Usando empresas que tenían este tipo de créditos antes de 1998, y mantuvieron su elegibilidad en el año 2000, los autores muestran que el crédito directo se utiliza para realizar una expansión de actividades y no para sustituir otros créditos más costosos obtenidos en el sector informal. Por tanto, se invierte más y no se sustituye el crédito costoso por un crédito más barato. La utilización del crédito por parte de las empresas acelera la tasa de crecimiento de sus ventas y sus beneficios. Por tanto, la conclusión es que las empresas deben estar severamente restringidas en sus posibilidades de acceso al mercado de créditos. La estimación muestra que el rendimiento privado de una rupia extra de préstamo es cercano al 90%; el rendimiento social es del 84%.

El segundo estudio, con datos microeconómicos sobre el efecto de las restricciones crediticias, se puede encontrar en Karlan y Zinman (2006), que analizan una intervención aleatorizada en Sudáfrica. La expansión del acceso al crédito es un elemento destacado de las estrategias de de-

sarrollo de la mayoría de los organismos internacionales. Los donantes y los gestores de la política económica tienen un interés enorme por aumentar el acceso a los créditos y evaluar su efecto. Mientras que existe acuerdo sobre el efecto de la expansión del acceso al crédito de las microempresas y las pequeñas empresas, es menos claro el efecto que tienen los microcréditos destinados al consumo. De hecho, son muchos los analistas que desconfían de los efectos beneficiosos de los microcréditos destinados a actividades no productivas. Karlan y Zinman (2006) estiman el efecto de los créditos al consumo en países en vías de desarrollo. Para evaluar su impacto realizan un experimento aleatorizado de expansión del crédito al consumo y recogen datos posteriores sobre la situación de las familias que fueron agraciadas con el préstamo frente a las que no lo obtuvieron. En muchos mercados de créditos (como por ejemplo en Estados Unidos) se utiliza una puntuación crediticia (*credit score*) para decidir el tipo de interés que se carga a cada cliente. Es decir, se ve todo un conjunto de características del individuo (pagos de créditos en el pasado, crédito acumulado, renta, etc.) y se decide si se le concede un crédito y a qué tipo de interés. El sistema de asignación de los créditos fue el siguiente: se solicitó a una empresa crediticia de Sudáfrica que relajara sus criterios de riesgo y aconsejara a sus gestores el conceder aleatoriamente préstamos a solicitantes que estuvieran marginalmente por debajo de la puntuación crediticia necesaria para obtener la concesión de un crédito.

Los resultados de Karlan y Zinman (2006) indican que el incremento de la concesión de crédito entre los solicitantes con puntuaciones crediticias marginales aumentó el acceso general al crédito y, al menos aparentemente, resultó beneficioso para los bancos (a pesar de que en circunstancias normales no habrían concedido dichos créditos). Además, los receptores de los créditos tenían una probabilidad un 11% superior de mantener un empleo asalariado que los que no recibieron el préstamo; tenían una probabilidad un 6% menor de sufrir hambre severa y tenían un 7% menos de probabilidad de empobrecimiento respecto a los que no recibieron los créditos. Estas evidencias muestran cómo las restricciones de crédito realmente importan en los países en vías de desarrollo y su relajación puede favorecer tanto el desarrollo empresarial como el bienestar de los ciudadanos de estos países.

4.3.2. ¿Tienen menor acceso al crédito formal los ciudadanos más pobres?

En el epígrafe anterior hemos visto que el acceso al crédito puede mejorar el desarrollo empresarial y el bienestar de los ciudadanos en los países en vías de desarrollo. Desde el punto de vista de la distribución de la renta, es importante analizar si el acceso al crédito formal se distribuye uniformemente entre todos los ciudadanos.

Es importante hacerse esta pregunta porque una razón por la que podemos pensar que los pobres tienen más probabilidad de verse limitados en cuanto a créditos es por sus dificultades para acceder a créditos baratos en el sector bancario formal. Utilizando información sobre encuestas en 13 países, Banerjee y Duflo (2006) han mostrado que los pobres no necesariamente están menos endeudados, pero tienen menor probabilidad de que dichos préstamos provengan del sector formal (bancos comerciales, cooperativas, etc.). En dos entornos diferentes de la India (una zona urbana pobre y una zona rural) se observaron altos niveles de endeudamiento pero bajos niveles de endeudamiento con el sector formal. En Udaipur dos tercios de los encuestados tenían un crédito, pero sólo un 6,4% de los mismos provenía del sector formal. En Hyderabad sólo un 5% de los encuestados tenía un crédito, de un banco comercial. Por tanto, parece que los ciudadanos con pocos recursos se endeudan básicamente a partir de cauces e instituciones no bancarias.

¿Cuáles son las razones teóricas que explican esta situación de mayor informalidad crediticia entre los ciudadanos más desfavorecidos? Hay dos razones fundamentales recogidas en la literatura: la dificultad de acceso a instituciones crediticias formales (por su ausencia en muchas zonas) y la ausencia de bienes que puedan utilizarse como colateral o garantía real. Con respecto al primer punto, es evidente que en las áreas donde viven las familias más ricas existen sucursales bancarias. Además, los ciudadanos ricos están mejor informados de las posibilidades de acceso a créditos formales y pueden acceder a las instituciones financieras existentes.

En segundo lugar, los ricos tienen mayor posibilidad de utilizar colateral (o garantías reales sustentadas en activos propiedad de los solicitantes de crédito) para obtener préstamos formales, pues tienen activos propios.

Los pobres pueden tener algunos activos pero, en muchos casos, no existen derechos de propiedad claros (por ejemplo, no existen títulos de propiedad de viviendas en los barrios de chabolas de los alrededores de muchas grandes ciudades del tercer mundo) o la posibilidad de defenderlos. A falta de colateral, los pobres no tienen más remedio que solicitar créditos fundamentalmente en instituciones no formales, soportando un tipo de interés muy superior al ofrecido por bancos comerciales y cooperativas.

Por tanto, existe evidencia que sugiere que los pobres tienen relativamente menor acceso a los créditos formales y que esto afecta negativamente a sus condiciones económicas (probabilidad de tener un empleo asalariado, sufrir hambre severa o empobrecimiento). El problema del acceso al crédito de los pobres no es exclusivo de los países en vías de desarrollo. En Estados Unidos más del 35% de las familias de renta baja no tienen siquiera una cuenta corriente. Obviamente, esta proporción es mucho mayor en países en vías de desarrollo donde la media es del 50%. En Sudáfrica el 73% de las familias de bajos ingresos no tiene cuentas bancarias. En los países en vías de desarrollo son los pobres de las zonas rurales los que tienen menor acceso al sector bancario. Como los costes de bienestar de la exclusión de los pobres del sistema financiero son altos, existe un cierto acuerdo sobre la conveniencia de realizar intervenciones gubernamentales en el sector bancario de los países pobres.

Por tanto, la insistencia en la importancia de los programas de créditos para los pobres no es tan novedosa. La auténtica novedad consiste en el interés por realizar un diseño adecuado que permita la evaluación de los programas de acceso al crédito. Hasta hace diez años no se ponía ningún énfasis en intentar identificar el impacto causal de estos programas. Los créditos subvencionados para los pobres tuvieron un desarrollo inicial en los años sesenta y setenta, aunque luego sufrieron un rechazo claro en muchos países durante los años ochenta. En los años noventa se desarrollaron nuevas políticas destinadas a mejorar el acceso al crédito, como los microcréditos. Sin embargo, las decisiones adoptadas de expansión, contracción o cambio en los métodos para mejorar el acceso, no se basaron en evaluaciones científicas de sus efectos, sino más bien en cuestiones políticas.

Parte de la explicación de la ausencia de evaluación de estos programas es la dificultad de realizarla. Por ejemplo, supongamos que queremos sa-

ber cuál es el impacto de mejorar el acceso al crédito en una región sobre los resultados económicos de los individuos. Para realizar esta evaluación queremos comparar dos personas similares, una de las cuales consiguió el crédito mientras la otra no lo consiguió. La comparación de la situación antes y después del cambio en las condiciones de acceso al crédito en una región es difícil que proporcione una respuesta adecuada a la pregunta sobre el efecto de mejorar el acceso al crédito. Un motivo es que otras cosas han cambiado al mismo tiempo entre el período anterior y el posterior a la mejora del acceso al crédito y, por tanto, es difícil atribuir a la mejora del acceso al crédito los resultados económicos observables. Además, los individuos de una determinada región pueden tener unas condiciones económicas mejores o peores por cuestiones no relacionadas con la posibilidad de acceder al crédito. Por ejemplo, nos podemos hacer la siguiente pregunta: dos años después de haber llegado una entidad de microcréditos a una determinada zona, ¿ha mejorado la economía de la aldea y el bienestar de sus ciudadanos? Si se compara la situación económica antes y después de la llegada de los microcréditos, algunos de los efectos serán consecuencia de la disponibilidad de los mismos. Pero también puede haber otros elementos que estén asociados a otros cambios temporales (subsidios agrícolas, sequía, malas cosechas, etc.) en la aldea. Es muy difícil desentrañar esta madeja de posibles causas. A lo mejor llegó a la aldea una entidad de microcrédito y, al mismo tiempo, se produjo una sequía en la aldea que empeoró la situación económica de sus habitantes. Esta situación no debería interpretarse como que la disponibilidad de microcréditos tiene un efecto negativo sobre la economía de la aldea.

Una metodología alternativa a la comparación del antes y el después de la actuación (también llamada estimación en diferencias) consiste en comparar a quienes se beneficiaron del acceso al crédito con los que no se beneficiaron, siempre dentro de la misma región. Sin embargo, este método también puede tener problemas. En principio, pueden sobreestimar los efectos de los microcréditos si personas más capaces o energéticas son las que finalmente reciben los créditos. Si éste es el caso, entonces no podremos saber si sus condiciones económicas han mejorado gracias a los microcréditos o a sus propias características personales (tenacidad, constancia, capacidad organizativa, etc.). Por ejemplo, en muchas ocasiones

las entidades de microcréditos prestan a las mujeres y, a veces, sólo a unidades familiares que no tienen posesión de tierra o menos de un acre de terreno. Se podría pensar en plantear la evaluación mediante la comparación, dentro de una misma aldea de los individuos que no tienen tierra y tienen acceso a microcréditos con los que tienen tierra pero no recibieron microcréditos. Por tanto, los individuos de ambos grupos son similares en términos de posesión de tierra. Pero puede haber un problema, porque incluso dentro del grupo de poseedores de tierra va a existir un cierto sesgo de selección. ¿Quién recibe el crédito? Normalmente, dentro de los elegibles, reciben los créditos los que tienen un carácter más empresarial. Entonces, habrá una sobreestimación del efecto del programa de créditos.

Por otra parte, si nos centramos en los ciudadanos más desfavorecidos de una aldea, entonces el microcrédito puede tener efectos muy positivos. Sin embargo, los muy pobres que no tienen acceso a los cuidados de salud, por ejemplo, no pueden utilizar productivamente los créditos, y quizás no se pueda medir un impacto significativo en su nivel de vida. Esto no quiere decir que los microcréditos no tengan efecto: simplemente que se le concedió el crédito a una persona que a la vez contrajo una enfermedad que le imposibilitó utilizarlo de forma productiva. En esta situación no hay forma clara de identificar los efectos. Es muy difícil saber si el sesgo, teniendo en cuenta los dos factores comentados anteriormente, es positivo o negativo. Tal vez ésta es, en parte, la razón última por la que a lo largo de los años hemos visto programas de créditos que han sido descartados en algunos países y no ha habido acuerdo sobre si eran buenos o si no lo eran. En muchos casos se desplegaron programas que no se evaluaron.

4.4. ¿Pueden los bancos rurales reducir la pobreza?

¿Puede una expansión del crédito liderada por los estados reducir la pobreza? Ésta es la pregunta que se plantea en el trabajo de Burgess y Pande (2005), que utilizan la expansión rural de bancos en la India como caso de estudio.

4.4.1. El contexto bancario de la India

Entre la nacionalización bancaria de 1969 y la liberalización financiera de 1990-1991, más de 30.000 zonas rurales de la India pasaron de no tener ningún banco comercial a albergar al menos una sucursal bancaria. Esta expansión produjo un aumento significativo del crédito rural y de los ahorros en bancos comerciales, que pasaron de 1,5 y 3%, respectivamente, al 15%.

En 1969 el gobierno de la India nacionalizó sus 14 bancos más grandes y anunció un *programa de banca social*, con el objetivo de igualar el acceso a las sucursales bancarias en todos los estados de la India. Cualquier banco que quiere abrir una sucursal en la India precisa de una licencia que concede el banco central. En enero de 1977, el banco central de la India anunció una nueva política: las licencias «1:4». Esto quería decir que, por cada licencia concedida para abrir una oficina en una localidad que ya tenía alguna oficina de ese banco, era preciso que abriera cuatro oficinas en localidades que no disponían de ninguna. Los bancos fueron obligados a seleccionar para su expansión las localidades que no tenían oficinas a partir de una lista que confeccionó el banco central. La lista tenía todas las localidades cuya población superaba un determinado límite, pero no tenían ninguna oficina bancaria. Dicho límite era el mismo para todos los estados. Por tanto, la política implicaba una localización relativa mayor en localidades que tenían inicialmente un menor desarrollo financiero. El programa «1:4» finalizó en 1990 con la liberalización del sistema bancario indio. Entre 1977 y 1990, los estados menos desarrollados financieramente de la India vieron un mayor crecimiento de las oficinas bancarias en zonas rurales que inicialmente no tenían oficinas. Lo contrario era cierto antes de 1977 y después de 1990.

La evaluación del impacto basada en la correlación entre el número de oficinas abiertas en zonas que no tenían inicialmente oficinas y la pobreza rural puede proporcionar la respuesta incorrecta a la pregunta inicial de este epígrafe. Los estados indios con menos oficinas antes del comienzo del programa (a partir de ahora las denominaremos localidades menos desarrolladas financieramente) tenían también más localidades sin ninguna oficina. Por tanto, veremos una correlación positiva entre la pobreza y

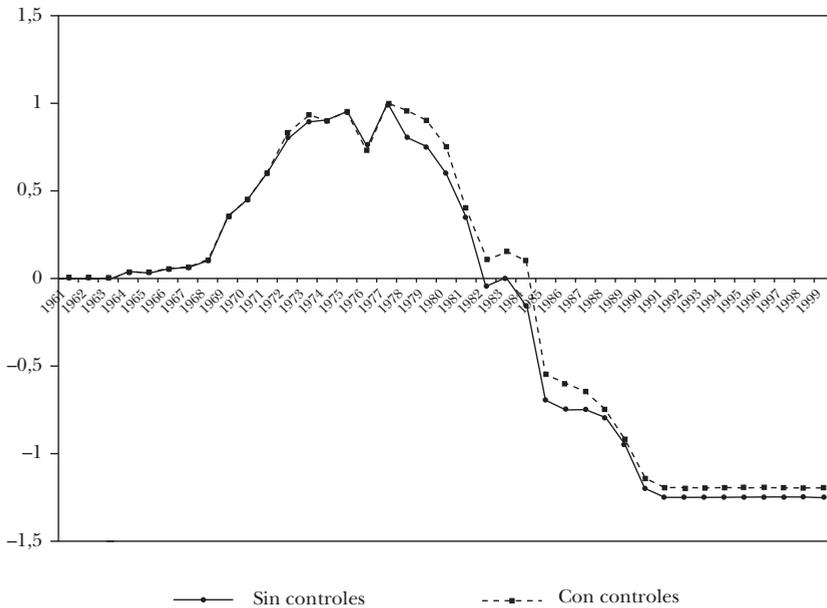
la incidencia del programa. Sin embargo, la interpretación de dicha relación no es causal: no implica que el programa de expansión de oficinas en zonas rurales aumentó la pobreza.

4.4.2. La evaluación del impacto del programa de expansión bancaria rural

Entonces, ¿cómo se puede evaluar este programa de expansión bancaria? Lo que no se puede hacer es simplemente comprobar si las áreas que desarrollaron la presencia bancaria mejoraron su rendimiento económico. El motivo es que el programa estableció como objetivos específicos los estados menos desarrollados. La India es un país federal que está dividido en estados. Un objetivo específico de la política bancaria era, como ya se ha comentado, igualar el acceso a los bancos en todos los estados de la India. Parece lógico pensar que los estados de la India que tenían menos bancos, al principio del programa, acabarían teniendo más bancos al final. Los estados mayoritariamente más pobres eran los que tenían menos bancos inicialmente. Por tanto, como se señaló con anterioridad, si se calcula simplemente la correlación entre el aumento de los bancos y los resultados en términos de pobreza, se encontrará una relación positiva muy fuerte: cuantas más sucursales bancarias se han creado, mayor será la pobreza del estado. Para evitar interpretar causalmente esta correlación Burgess y Pande (2005) explotan el programa de licencias «1:4» metodológicamente, como experimento natural para evaluar el impacto de dicha política. Entre el 1969 y 1990, más de 30.000 ubicaciones rurales en la India, normalmente de menos de 5.000 personas, pasaron de no tener ninguna oficina bancaria comercial a tener por lo menos una sucursal bancaria comercial. Esta transformación es potencialmente muy importante en la India, un país donde la población es mayoritariamente rural.

La política «1:4» significó que durante trece años, entre 1977 y 1990, los estados indios menos desarrollados vieron un mayor crecimiento de sucursales bancarias en las localidades rurales. El gráfico 4.3 contiene en el eje horizontal los años y en el eje vertical la correlación, positiva o negativa, entre el desarrollo financiero o económico inicial del estado, es decir, las sucursales bancarias antes de la puesta en marcha del programa, y

GRÁFICO 4.3: Evolución temporal de la correlación entre desarrollo bancario inicial y nuevas oficinas en zonas rurales



cuántas sucursales se abrieron durante el programa. Hasta 1977 los estados indios más desarrollados tuvieron un mayor crecimiento de sucursales bancarias. Entre el 1977 y 1990, esta relación se invirtió de forma que lo que hizo la política fue asegurar que los estados menos desarrollados tuvieran mayor crecimiento bancario.

De manera más técnica se puede señalar que el gráfico 4.3 muestra γ el coeficiente de la regresión $B_{it}^R = \alpha_i + \beta_t + \gamma B_{i,1961} + \delta_t X_{i,1961} + \varepsilon_{it}$

Donde B^R es el acumulado de oficinas abiertas en zonas rurales que no disponían de bancos, en términos per cápita, $B_{i,1961}$ es la medida de desarrollo financiero inicial, o número de oficinas per cápita en el estado i en 1961. Esta variable entra en interacción con una dicotómica anual que refleja el efecto temporal. La diferencia entre el parámetro en t y en $t + 1$ nos dice cómo el desarrollo financiero inicial afecta al crecimiento del número de oficinas entre t y $t + 1$. Finalmente, la X es un vector de condiciones iniciales del estado que incluye la renta real per cápita, la densidad de población y el número de localidades rurales per cápita, todo medido en el año 1961. Estas variables también tienen un coeficiente que es específico para cada año.

El gráfico 4.3 muestra la evolución temporal del coeficiente del desarrollo financiero inicial en zonas rurales (γ) y muestra cómo los estados más desarrollados ofrecieron a los bancos mayores oportunidades de beneficios y, por esto, se observa una creciente localización de bancos en zonas que ya tenían oficinas bancarias, inicialmente, entre 1961 y 1977 (el coeficiente es positivo y, por tanto, cuanto mayor desarrollo financiero inicial habrá mayor acumulación de nuevas oficinas habrá). Esta tendencia se revierte en 1977, precisamente cuando la política «1:4» entra en vigor. Entre 1977 y 1990, el coeficiente disminuye en el tiempo y, por tanto, los estados menos desarrollados muestran un mayor crecimiento en el número de oficinas per cápita en zonas que no tenían oficinas. Después de 1990 el coeficiente deja de disminuir, mostrando el final de la política «1:4» y el comienzo de la liberalización. El resultado es muy similar si se incluyen controles (las X) o no se incluyen, como muestra el gráfico 4.3. La misma reversión se observa si se trabaja con distritos en lugar de estados. Por tanto, parece claro que la política de licencias «1:4» favoreció que los bancos priorizaran los distritos menos desarrollados financieramente dentro de cada estado.

El gráfico 4.4 muestra la expansión bancaria en localidades que ya tenían banco. Entre 1977 y 1990 el crecimiento bancario fue muy lento (la pendiente de la recta se aplanó). El motivo es que la política «1:4» hacía más cara la expansión bancaria y, por tanto, reducía también la apertura de oficinas en áreas que ya tenían bancos. En parte, la reacción de los bancos fue reducir la cantidad de localidades en las cuales se abrió una oficina entre 1977 y 1990. Después de 1990, como se puede observar en el gráfico 4.4, los bancos volvieron exactamente a la misma tasa de crecimiento en localizaciones en las que ya había previamente bancos que tenían antes de 1977. Por tanto, la política «1:4» lo que hizo fue reducir la expansión de los bancos en localidades que ya tenían oficinas. En grandes ciudades como Bombay o Delhi se abrieron, entre 1977 y 1990, menos oficinas porque resultaba muy caro (había que abrir cuatro sucursales en zonas rurales sin oficinas por cada oficina que se abriera en estas ciudades).

En conclusión, es claro que los estados indios que tenían menos bancos al principio acabaron teniendo más sucursales bancarias. El gráfico 4.5 contiene la evolución de la relación entre desarrollo financiero inicial y el

GRÁFICO 4.4: Evolución del número acumulado de oficinas en zonas que ya tenían oficinas inicialmente

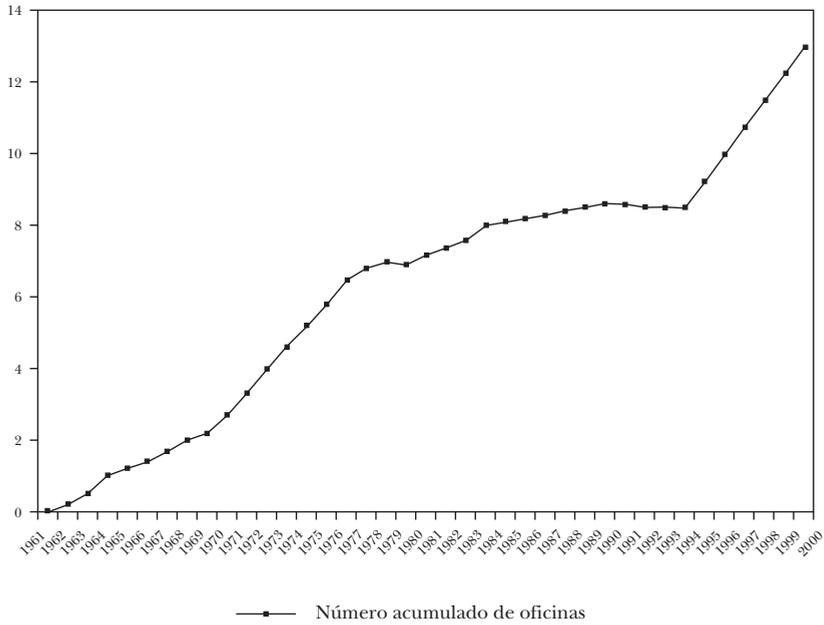
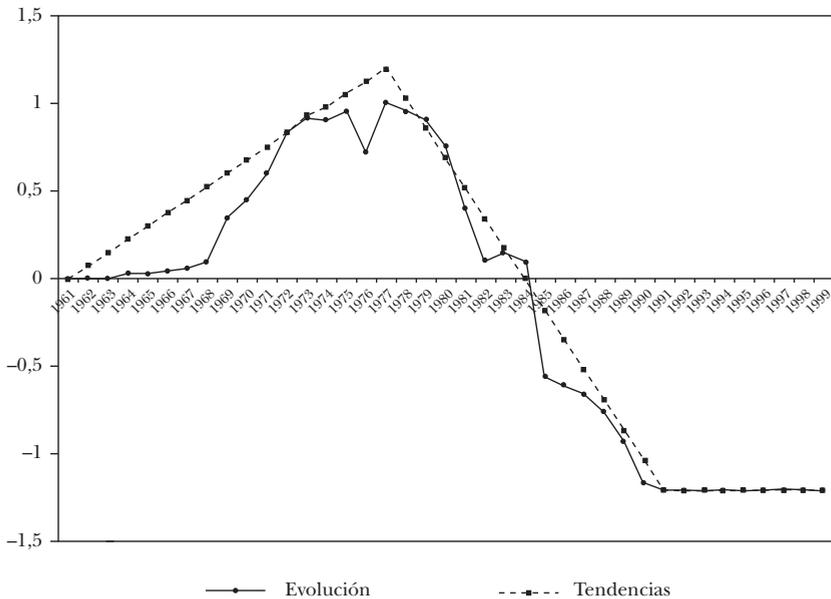


GRÁFICO 4.5: Tendencias en la relación entre desarrollo financiero inicial y apertura acumulada de oficinas



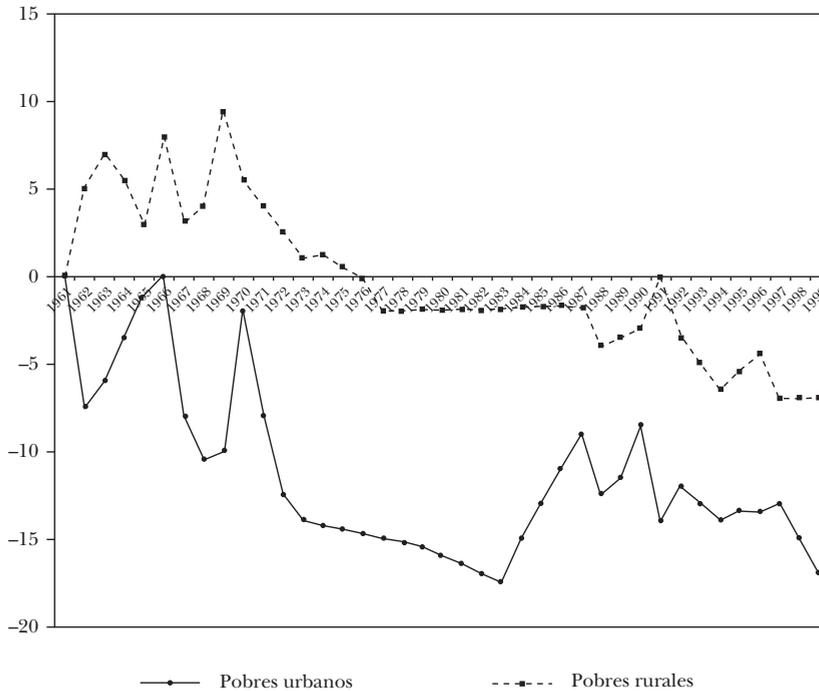
número de nuevas oficinas durante el período, resumida en los tres subperíodos relevantes: antes de 1977, entre 1977 y 1990 y después de 1990. El gráfico 4.5 muestra una clara tendencia creciente en la primera parte de la muestra, una clarísima tendencia decreciente en la segunda parte y un estancamiento en la tercera submuestra.

Pero ¿qué implican estas tendencias sobre los indicadores que nos interesan? Si examinamos el porcentaje de crédito rural, observamos que sigue la misma tendencia: hasta 1976 aproximadamente los estados más ricos tuvieron una mayor expansión crediticia pero, a partir de esa fecha, son los estados menos desarrollados financieramente los que aumentan más la proporción del crédito rural.

Resumiendo, podemos decir que no sería correcto comparar sencillamente los resultados en los diferentes estados indios y ver cuál es el crecimiento de los estados más o menos desarrollados. Pero lo que sí se puede explotar es el hecho de que hubo un cambio continuo entre el 1977 y 1990, y preguntarnos si este cambio provocó algún resultado en las variables que nos importaban, como por ejemplo la pobreza, el crecimiento industrial, etc. Por tanto, preguntarnos si los cambios provocados por el aumento de las oficinas, en zonas que inicialmente no tenían oficinas, mejoraron el desarrollo de los estados utilizando el período entre 1977 y 1990 como un experimento natural (un cambio en la política de concesión de licencias de apertura de establecimientos bancarios). ¿Es correcto señalar que entre 1977 y 1990 los estados más pobres tuvieron un desarrollo económico mayor como consecuencia de tener más oficinas bancarias? Si la respuesta es afirmativa, entonces la política bancaria que invirtió el sistema de organización de los bancos tuvo como consecuencia una convergencia de resultados económicos en los estados. Uniformó, en definitiva, el desarrollo económico de los estados de la India. El gráfico 4.6 muestra los puntos de inflexión o cambios que tuvieron lugar en las tendencias de la pobreza en el período.

El gráfico 4.6 muestra la relación entre el cambio en la política y la evolución de la pobreza rural. Entre 1961 y 1977 la pobreza se redujo más rápidamente en los estados que estaban financieramente más desarrollados en 1961. Pero a partir de 1977 y hasta 1990 se observa la inver-

GRÁFICO 4.6: Evolución del número de pobres en zonas rurales y urbanas



sión de la tendencia anterior: una relación negativa entre el desarrollo financiero inicial de un estado y la reducción de la pobreza: los estados que estaban recibiendo más oficinas bajo el programa «1:4» vieron reducir su pobreza relativa más rápido que los estados que recibieron menos oficinas. Por tanto, antes de 1977 y después de 1990 la reducción de la pobreza rural fue más rápida en estados que estaban más desarrollados financieramente en 1961, es decir, al inicio del período muestra. Entre 1977 y 1990, esta tendencia se invirtió. Esto sucede en términos de la pobreza rural. Sin embargo, no hay relación con los resultados en la pobreza urbana.

Hay que tener en cuenta que los cálculos más ingenuos que examinan la correlación entre la expansión de las sucursales bancarias rurales y la pobreza rural muestran el resultado opuesto: los bancos rurales parecen incrementar la pobreza. Esta relación se debe al tipo de política desarrollada (de expansión en los estados más pobres) e indica muy claramente por qué los estudios observacionales, que examinan sola-

mente la correlación, proporcionan un resultado que no es correcto. De hecho, la reducción en la pobreza rural está asociada a la movilización de los ahorros y la provisión de créditos en áreas rurales. Los resultados sugieren que la política del banco central de la India de favorecer el desarrollo de una red extensa de oficinas permitió la acumulación de capital de las familias rurales y la obtención de créditos a largo plazo para inversiones productivas. Empezando en un nivel muy bajo, el número de cuentas de ahorro rurales y de créditos aumentaron en 126 y 25 millones, respectivamente. Además, los tipos de interés son atractivos en términos de los disponibles en los mercados informales. Evaluado en su media, la expansión de oficinas bancarias explica entre el 14 y el 17% de la caída del número de pobres (aproximadamente la mitad de toda la reducción durante el período). En concreto, cada sucursal bancaria adicional por cien mil habitantes reduce la pobreza rural en un 4,7%.

Por tanto, aquí tenemos un ejemplo de programa de expansión crediticia que podría tener efecto sobre la pobreza. El programa de expansión de oficinas rurales en la India produjo una reducción significativa de la pobreza rural e incremento el crecimiento en el sector no primario. El análisis de Burgess y Pande (2005) también demuestra que las evaluaciones simplistas pueden proporcionar resultados erróneos. Sin embargo, estos resultados positivos del análisis no responden a la cuestión de si el programa fue coste-efectivo. Las tasas de impago en los bancos indios en los años ochenta fueron en torno al 40%. Una razón importante para la finalización de la política de concesión de nuevas licencias tipo «1:4» fue que era muy poco rentable para los bancos. Al mismo tiempo, es posible que los subsidios que en el corto plazo resultan gravosos y negativos para la rentabilidad puedan tener efectos positivos sobre el crecimiento económico en el largo plazo. Quizás la tasa de subsidios de este programa fue demasiado alta, pero no se puede decir que como era poco rentable se trataba de un mal programa. Esto nos vuelve a llevar a la cuestión de que necesitamos evaluar toda una serie de programas crediticios, de diferentes tipos, en diferentes países pobres, para poder empezar a compararlos y también para empezar a hacernos una idea de cuál es el nivel mínimo de subsidio que hace falta para ayudar a los pobres.

4.5. El impacto de los microcréditos: ¿cuáles son las características importantes para su efectividad?

Este programa de expansión de bancos rurales en la India, que fue descrito en el epígrafe anterior, alcanzó a los pobres, y por eso se redujo la pobreza. Pero la evidencia sugiere que no llegó a las familias muy pobres. O sea, que uno de los objetivos de este programa (ayudar a los muy pobres) no se consiguió. ¿Qué significa eso? Pues que quizás existan programas alternativos que se podrían diseñar para ayudar a los muy pobres. Entre estos programas alternativos de subsidios, destacan las microfinanzas, que se centran en préstamos a mujeres muy pobres, y utilizan formas no monetarias de colateral. Estos programas tienen una tasa de amortización muy alta y una tasa de fracaso de menos del 10%. Entre estos programas el caso del banco Grameen de Bangladesh es quizás el más conocido ejemplo de créditos a pequeña escala, o microcréditos, para los pobres, siendo el 90% de los clientes mujeres. Pitt y Khandker (1998) estudiaron el caso de Bangladesh, que es uno de los países más microfinanciados, y explotaron la elegibilidad en estos programas y sus efectos. Específicamente, analizan el efecto del mayor acceso a los créditos sobre la educación de los niños, el consumo y la oferta laboral de las mujeres. Como en muchos casos anteriormente mencionados, la evaluación de este programa se ve complicada por la existencia de problemas de endogeneidad: hay atributos o características de la localidad y el hogar difíciles de medir, que afectan tanto a la demanda de crédito como a la renta del hogar. Por tanto, es difícil determinar la dirección de la causalidad entre crédito y gasto. Pitt y Khandker (1998) señalan que los efectos del programa dependen, decididamente, del género del participante: el gasto en consumo de las familias aumenta un 18% para préstamos a mujeres y sólo un 11% si el préstamo fue solicitado por un hombre. Además, el estudio de Pitt y Khandker (1998) indica que el programa de microcréditos ayuda a estabilizar y aumentar el consumo, especialmente no alimenticio, y a acumular activos diferentes de la propiedad de tierra. También promueve la educación y la formación de capital humano y contribuye al conocimiento sobre procedimientos

anticonceptivos. Finalmente, el impacto de los microcréditos se extiende a los pobres que no participan en el programa, lo que indica que los programas de microfinanzas ayudan a los pobres más allá de la redistribución de la renta, con una contribución real al crecimiento de la renta local. Los programas, por tanto, tienen efectos externos y ayudan a reducir la pobreza extrema más que la pobreza moderada al menos localmente.

La metodología de Pitt y Khandker (1998) se basa en un cuasiexperimento que toma como variación exógena las condiciones de elegibilidad en la identificación de los efectos del programa. Obviamente, si las condiciones de elegibilidad no se cumplen estrictamente (y existen algunas sospechas sobre la implementación de los criterios de elegibilidad por parte de los gestores del programa), entonces el procedimiento de identificación de estos efectos es imperfecto, lo que socava las bases de la evaluación.

Se ha señalado con anterioridad que un procedimiento alternativo para la evaluación es el uso de experimentos utilizando como base la aleatorización. Ésta es la aproximación en un estudio de campo, todavía en fase de realización, de Banerjee y Duflo. El trabajo se está haciendo en el marco de zonas urbanas de la India e intenta aleatorizar entre zonas deprimidas de la gran ciudad con acceso a microcréditos y otras que no tienen acceso a microcréditos para ver cómo esto afecta a los resultados económicos.

El segundo tipo de evaluación que se está haciendo tiene relación con el diseño de productos. Para pensar cómo expandir las microfinanzas hay que preguntarse qué características de su diseño son responsables de su éxito. Una característica importante de las microfinanzas es que los microcréditos muchas veces dependen de la responsabilidad del grupo. Por tanto, si esta característica fuera determinante y pudiéramos generar una situación similar para todos los créditos (si un individuo no pagara lo que debe, el resto del grupo se haría cargo de su deuda) conseguiríamos mejorar la eficacia de los créditos. La presión del grupo suplanta la ausencia de colateral: una vez que el banco se asegura de que si el individuo no pagara, el grupo se haría cargo, entonces presta. Sin embargo, la presión del grupo genera tensiones: los malos clientes viven de los buenos pagadores, pero el banco consigue la devolución del crédito. Gine y Karlan (2006)

han estudiado el uso de otras formas de responsabilidad solidaria para analizar su impacto. En un experimento basado en la aleatorización, encuentran que los clientes de microcréditos de Filipinas aceptan la responsabilidad conjunta del grupo porque es el único procedimiento que tienen para acceder a los créditos. Pero si pudieran elegir preferirían pasar a una responsabilidad individual en lugar de la colectiva. Por su parte, si los bancos pudieran resolver el problema de falta de colateral de los clientes también preferirían contratos individuales con responsabilidad individual, pues la configuración y formación del grupo supone un coste extra. El modelo de presión por parte del grupo es sólo un posible sustituto de la ausencia de colateral. En algunos países de Asia los pobres son capaces de pedir prestadas pequeñas cantidades de dinero utilizando documentos personales (como, por ejemplo, un certificado de nacimiento) que no tienen valor comercial para el prestamista pero que son importantes para el prestatario. Por tanto, no está claro que el mecanismo grupal que tienen muchos microcréditos sea una característica tan positiva de dichos contratos.

Otro factor común de los microcréditos es que tienen que devolver el crédito cada semana. Es decir, dan el préstamo hoy y la semana que viene tienen que empezar a devolverlo. Los expertos señalan que esto es muy importante para asegurarse de que estos individuos, muchos de los cuales no han tenido un préstamo antes, devuelvan el crédito. Field y Pande (investigación en curso) han dividido a las personas entre las que se concede un crédito en grupos de devolución semanal del préstamo y las de devolución mensual. Los resultados preliminares no muestran ninguna diferencia entre el comportamiento de ambos grupos. Por tanto, no parece que el repago semanal sea una característica determinante del impacto positivo de los microcréditos.

4.6. Conclusiones

Existe una enorme literatura que proporciona evidencia persuasiva sobre la fuerte relación entre el acceso al sistema bancario y el desarrollo económico. La evaluación de programas alternativos de crédito puede ayudar-

nos a identificar los que son más eficientes, en términos de coste, para reducir las restricciones de crédito de las empresas y las familias en países en vías de desarrollo.

Investigaciones recientes sugieren que se pueden utilizar dos tipos para este tipo de programas. El primero sería la explotación de un experimento natural, donde las condiciones de identificación provienen de modificaciones en las políticas públicas, como el programa de expansión de oficinas bancarias rurales en la India. En segundo lugar, se pueden utilizar métodos experimentales, basados en la evaluación aleatorizada, para identificar el impacto de muchas intervenciones microfinancieras. Muchas de estas evaluaciones todavía están en fase de campo, pero los resultados preliminares proporcionan grandes esperanzas respecto a su capacidad como mecanismo de evaluación.

Bibliografía

- BANERJEE, A., y E. DUFLO. «Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program». MIT, 2004 [mimeo.].
- . «Addressing Absence». *The Journal of Economic Perspectives* 20 (1), 2006: 117-132.
- BECK, T., A. DEMIRGUC-KUNT, y M. MARTÍNEZ PERIA. *Banking Services for Everyone? Barriers to Bank Access and Use around the World*. World Bank, 2006 [mimeo.].
- BURGESS, R., y R. PANDE. «Can Rural Banks Reduce Poverty? Evidence From the Indian Social Banking Experiment». *American Economic Review*, 2005.
- GINE, X., y D. KARLAN. *Group Versus Individual Liability: a Field Experiment in the Philippines*. Yale University Economic Growth Center Discussion Paper 940, 2006 [mimeo.].
- HONOHAN, P. «Financial Development, Growth and Poverty: How Close are the Links?». En Charles Goodhart, ed., *Financial Development and Economic Growth: Explaining the Links*. Londres: Palgrave, 2004: 1-37.
- KARLAN, D., y J. ZINMAN. *Expanding Credit Access: Using Randomized Supply Decisions To Estimate the Impacts*. Princeton, 2006 [mimeo.].
- KING, R. G., y R. LEVINE. «Financial and Growth: Schumpeter Might be Right». *The Quarterly Journal of Economics*, vol. 108, 1993: 717-27.

PAULSON, A., y R. M. TOWNSEND. «Entrepreneurship and Financial Constraints in Thailand». *Journal of Corporate Finance* 3, 2004.

PITT, Mark, y Shadihur KHANDKER. «The Impact of Group Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?». *Journal of Political Economy* 106 (5), 1998: 958-996.

El valor de la evaluación interdisciplinar: el análisis de programas de desarrollo basados en la comunidad

Vijayendra Rao

Lead Economist, Departamento de Investigación del Banco Mundial, Washington

5.1. Introducción

A los investigadores sobre el problema de la pobreza se les pide, en muchas ocasiones, una receta mágica para acabar con ella. Como se ha señalado en los capítulos previos, no existe una receta mágica que pueda eliminar la pobreza de un plumazo. Pero, al menos en un nivel metodológico, sí existe propuesta que nos ayudará a encontrar soluciones. Dicha propuesta radica en considerar todas las metodologías existentes sobre evaluación en todas las disciplinas eligiendo un conjunto formado por todas aquellas que resultan más apropiadas. Centrarse en un solo método, aunque sea útil, no es necesariamente la respuesta. Es admirable el trabajo de los autores de los capítulos anteriores de este volumen, pero quedan algunas dudas sobre la efectividad del método de evaluación experimental aplicado a cualquier cosa, y sin la consideración de otros procedimientos. Este capítulo centra en la evaluación de actuaciones denominadas *desarrollo basado en la comunidad* (*Community Driven Development* o CDD en

su acrónimo en inglés). Ésta es una evaluación continuada y difícil. Parte de los comentarios se centran en un CDD que tiene lugar en Indonesia y cuya evaluación no ha finalizado. Una evaluación a largo plazo de la que todavía no hay resultados ya que se encuentra pendiente de finalizar.

En este capítulo se desarrolla la metodología y los elementos base de dicha evaluación. Los epígrafes se dividen de la siguiente manera. En primer lugar, se comentan las limitaciones asociadas a las evaluaciones de impacto basadas en experimentos aleatorizados. También se comentan otros aspectos como la preocupación por el efecto de la aleatorización; las alternativas y los elementos complementarios; el valor añadido de los métodos cualitativos; el significado de un programa del tipo CDD; las evidencias disponibles sobre el impacto de los programas de desarrollo basado en la comunidad y algunas cuestiones todavía abiertas en este tipo de actuaciones. El penúltimo epígrafe describe la evidencia disponible sobre un programa de CDD, todavía pendiente de finalización, así como la metodología de la evaluación en el caso de Indonesia. Finalmente, el último epígrafe presenta las conclusiones del capítulo.

5.2. Preocupaciones generales sobre la evaluación de impacto

Al comenzar una investigación la pregunta que queremos contestar debe conducir al método que ha de utilizarse y no al revés: el método antes de la cuestión. Un segundo problema de la evaluación de impacto es que las intervenciones se comparan con la no intervención. Por tanto, comparamos los incentivos de ir a la escuela con la ausencia de cualquier intervención. Éste no es el caso de muchas evaluaciones que se realizan hoy día, pero ha sido el caso de algunas evaluaciones que se presentan en los capítulos anteriores. Sin embargo, el problema es que, al comparar los resultados con la ausencia de intervención, encontrar qué tiene efecto no es tan sorprendente. La cuestión importante debería ser: ¿se podría haber hecho algo mejor? Hay mejores pruebas, en particular cuando hablamos de países en desarrollo.

El tercer tema importante tiene que ver con la validez interna de los resultados comparados con la validez externa. Los resultados de una evalua-

ción se obtienen en un entorno específico. La pregunta es hasta qué punto se pueden extrapolar dichos resultados, y si los efectos serán los mismos a una mayor escala. Hemos visto que algunas evaluaciones hechas por gobiernos se han realizado a gran escala. Pero, para poder hacer estudios experimentales aleatorizados, se necesita un entorno cerrado y controlado por las ONG o los investigadores. Por tanto, los impactos se refieren a entornos pequeños y cerrados.

El tema de la validez interna frente a la validez externa es fundamental. Las diferencias contextuales no deben pasarse por alto: el papel de la cultura, la estructura social, la política y la geografía es crucial a la hora de determinar el impacto de una intervención concreta. Esto es algo muy difícil. Podríamos hacer experimentos aleatorios en un lugar y luego en otro, pero ciertamente no es lo mismo. El ensayo no es igual, porque se trata de entornos diferentes y en estas actuaciones, que no son intervenciones sencillas, pueden tener un impacto importante. Cuando hablamos de una intervención en el seno de la comunidad, es muy difícil copiarla a otra porque el contexto donde se lleva a cabo la intervención es fundamental y, además, se aprende a medida que se aplica la intervención. Detectamos errores, los resolvemos, seguimos adelante y entonces la intervención cambia; cambia para mejorar. La evaluación de este tipo de actuaciones es muy compleja porque está en continuo cambio, en continua mutación.

Pero esto es sólo parte del problema. El impacto va normalmente más allá de lo que inicialmente hemos pensado. Se pueden producir externalidades en el sentido de que el impacto de la intervención afecte al grupo tratado, pero puede generar efectos externos también en el grupo de control. Y esto es lo que ocurre en cualquier tipo de evaluación. Este proceso de generación de externalidades, que van más allá de lo previsto, ¿cómo se interpreta? Además, cuando el programa pasa de una dimensión pequeña a ampliarse al conjunto de un país o una región, entonces la pregunta es la siguiente: ¿el efecto observado en el pequeño grupo experimental se mantendrá cuando el programa se generalice a una gran parte de la población?

Por otra parte, en cuarto lugar, está la cuestión del análisis del impacto frente al análisis de los procesos. ¿Qué significa un proceso? Imagínense que vemos dos hermosísimos cuadros de Goya. A distancia vemos que son maravillosos. Pero cuando nos acercamos cada vez más empezamos a ver

la técnica, empezamos a ver cómo logró los efectos, las pinceladas, el juego con la luz, las diferentes capas de pintura, etc. Esto es lo que hace de un cuadro de Goya ser un Goya, lo que le hace diferente de cualquier otro pintor. Esto es el proceso. Hay que analizar el impacto, por supuesto. Cuando vemos un Goya podemos apreciar su impacto a distancia. Es un cuadro maravilloso que nos impacta. Pero luego está el proceso. El proceso son los eventos que han llevado a ese impacto, cómo llegamos a ese resultado. El proceso es la serie de eventos instigados por una intervención, que han llevado a una serie de resultados (políticas locales, interacciones entre entidades gubernamentales y funcionarios con los directores de las escuelas, etc.). Estos aspectos no se pueden anticipar con una pregunta o con un cuestionario estructurado. En la evaluación cuantitativa la herramienta estándar, o una herramienta que lo caracteriza, es el cuestionario estructurado. Es un cuestionario que tendría que predecir lo que va a suceder para poder tener preguntas en concreto. Hay que buscar diferentes técnicas para tener las preguntas adecuadas, pero es una tarea difícil, y sólo podemos hacerlo en cierta medida. Éste es uno de los problemas de cualquier evaluación de impacto. El tercer elemento en la dicotomía entre impacto y proceso es la consideración del verdadero *impacto*. Puede haber elementos positivos y negativos, y externalidades. Hace diez años se trabajó en una evaluación de un proyecto en Calcuta, en la India, que intentaba extender la idea de usar el preservativo entre las prostitutas, enseñándoles los horrores del sida para convencerlas de que sus clientes utilizarasen preservativos. Si la evaluación se hubiera hecho aleatoria entonces las prostitutas que tenían que obligar a sus clientes a ponerse el preservativo habrían perdido el 50% de los ingresos. Éste era un problema para la evaluación aleatorizada. La idea es que debido a esta intervención, que se había hecho con una base comunitaria, es decir, a partir de que unas prostitutas hablaran con otras sobre el virus del sida y sus consecuencias, unas prostitutas enseñaban a otras las ventajas de utilizar preservativos. Este mecanismo de transmisión de información, parte del proceso de aplicación del programa, tuvo un gran impacto en la sociedad. Las mujeres empezaron a pensar que tenía sentido y las prostitutas tuvieron cierta sensación de ser como el personal médico. Esto llevó a las prostitutas a iniciar una organización cuyas representantes han estado en San Francisco, han apa-

recido en el *New York Times* y en muchos otros medios. El éxito de este emergente asociacionismo entre las prostitutas ha llegado hasta el punto de iniciar un movimiento para que se legalice la prostitución en Calcuta. Teniendo en cuenta lo conservadora que es la sociedad de la India en cuanto al sexo, estas consecuencias no esperadas del proceso han sido muy importantes. Estos resultados no anticipados pueden ser mayores que el impacto concreto que se esperaba obtener (aumentar la utilización de preservativos entre las prostitutas).

5.3. Problemas adicionales con los experimentos aleatorizados

Las preocupaciones generales comentadas con anterioridad se agravan, o se centúan, en los experimentos aleatorizados.¹ Es evidente que en muchas circunstancias, como se ha señalado en los capítulos anteriores, existen problemas importantes de selección muestral. Los ensayos aleatorizados permiten resolver el problema de la selectividad muestral, al hacer totalmente aleatoria la decisión de participar en un programa o la aplicación de una actuación a determinados individuos o aldeas, y no a otras. De esta forma, se generan dos grupos comparables (tratado y de control), que es la gran ventaja de los ensayos aleatorizados. Pero la aleatorización provoca sus propios problemas. En primer lugar, la gente se comporta de forma diferente cuando se siente observada: es el llamado efecto Hawthorne. El sesgo de la aleatorización se produce por los cambios de comportamiento que puede provocar el hecho de que el individuo sepa que está siendo observado.

En segundo lugar, las intervenciones más sencillas son las que se pueden estudiar científicamente. Sin embargo, las intervenciones reales, a gran escala, son generalmente muy complejas. Entonces el dilema es la diferencia entre el *mundo* de la experimentación y el mundo real. Y esto nos lleva, en

¹ Para ver una discusión general sobre la aplicación de esta metodología a programas y proyectos de desarrollo económico en países en vías de desarrollo, se puede consultar Duflo, Glennerster y Kremer (2007).

tercer lugar, a una serie de condicionamientos políticos. No sería sensato pasar por alto los problemas políticos. Vivimos en democracia. La gente es elegida para un trabajo determinado, o para una actividad política. El enfoque científico para la evaluación de una intervención determinada no es posible en algunos casos, en particular los ensayos aleatorizados, debido a que los políticos consideran que pueden perder votos si la opinión pública se hace eco de que una intervención pública selecciona a los receptores al azar. Las ganancias políticas a corto plazo superan el interés científico por una evaluación bien hecha. En estos casos, hay que buscar alternativas, pues no es políticamente factible hacer algunas cosas. Y, en ese caso en el que no se puede aleatorizar, ¿qué se puede hacer? Éste es el problema y el reto.

Finalmente están las preocupaciones éticas. Los problemas éticos se pueden resolver en muchas ocasiones. Por ejemplo, en las actuaciones en países en vías de desarrollo se podría dar algo a todo el mundo y no sólo a los participantes a los que les ha tocado estar en el grupo tratado. Es decir, que todo el mundo tenga algún tipo de beneficio. Pero, por otra parte, uno de los problemas en los que más se invierte es en tratar de mitigar la pobreza de grandes capas de la sociedad en países en vías de desarrollo. Pero si la evaluación es aleatorizada, incluso aunque sabemos que hay gente pobre en el grupo de control, les negaremos los beneficios para poder evaluar, lo que genera un problema ético: sabemos que alguien debería recibir algo que no va a recibir por la necesidad de realizar una evaluación adecuada. Éste es un problema ético grave. El reto es qué hacer en estas situaciones.²

5.4. Soluciones imperfectas

En los dos epígrafes anteriores se habló de los problemas generales de las evaluaciones de impacto y, en particular, de las evaluaciones basadas en experimentos aleatorizados. En este epígrafe se discuten algunas soluciones que, aunque imperfectas, son bien conocidas. Por tanto, no son soluciones perfectas, pero habrá elementos sustitutivos y elementos comple-

² Véase Ravallion (2007), para un análisis detallado de diferentes métodos cuantitativos para la evaluación de programas contra la pobreza.

mentarios de otros métodos. En primer lugar, están los procedimientos no experimentales. Dentro de este grupo están los procedimientos de emparejamiento, las estimaciones con discontinuidades y la estimación por variables instrumentales. En los *procedimientos de emparejamiento* utilizando el *propensity score* es necesario tener un conjunto muy exhaustivo de indicadores observables para poder realizar una evaluación con garantías. En este caso tenemos el grupo tratado que recibe la intervención y habrá que encontrar grupos de personas de control que sean similares, respecto a su posibilidad de participación en el programa, en función de los indicadores observados. Clones estadísticos, por así llamarlos, de las personas tratadas. Esto es lo difícil del método de *matching on the propensity score*: saber si el método ha producido un grupo de control comparable al grupo tratado. Si no tengo suficiente información para hacer estos emparejamientos, es difícil que los grupos acaben siendo comparables. En ocasiones no podemos tener un experimento aleatorizado, pero sí podemos hacer buenos emparejamientos. Con una correspondencia adecuada, podremos lograr los mismos resultados que si hubiéramos hecho un estudio aleatorizado, pues los grupos (tratado y control) serán plenamente comparables.

En segundo lugar, un método muy útil, en particular en la evaluación de programas de lucha contra la pobreza, es el *diseño de discontinuidad*. Este método consiste en analizar las diferencias entre el efecto que una intervención tiene entre los individuos que son (marginamente) beneficiarios y los que no son, también marginalmente, beneficiarios. Por ejemplo, imaginemos que tomamos cualquier programa de lucha contra la pobreza. El programa puede indicar que no va a dar dinero a aquellas aldeas donde hay un tanto por ciento de personas que está por encima del umbral de la pobreza. Supongamos que las reglas de elegibilidad indican que sólo aldeas con más de un 50% de la población por debajo del umbral de la pobreza recibirán los fondos. Las aldeas que están con una proporción menor no reciben dinero, pues están por debajo del punto de corte. Pero si consideramos aldeas que tienen un 50% de pobres, un 51, o un 49 o un 48% y comparamos dos grupos (por encima y por debajo del umbral) esta diferencia del 2% entre la elegibilidad puede considerarse simplemente un efecto aleatorio, pues depende de datos imperfectos. Lo que se hace es poner en el grupo de control las aldeas que están por debajo pero muy

cerca del 50% y en el grupo tratado aquellas que están por encima pero muy cerca. Sólo podemos hacer estas comparaciones en una banda muy estrecha: por ejemplo, cuando comparamos las aldeas con un 49% de pobres con las que tienen un 52%. El diseño de discontinuidad es un método válido y se puede utilizar en programas de lucha contra la pobreza con bastante eficacia, aunque su validez se circunscribe a un entorno muy local del punto del umbral de separación.

Por último, también se puede utilizar el *procedimiento de variables instrumentales* como una alternativa no experimental a los métodos de ensayo aleatorizado. En este caso, el problema es encontrar un instrumento válido para realizar la evaluación. En general, resulta muy difícil encontrar un indicador que cumpla las características de un buen instrumento y, por tanto, puede ser un procedimiento potencialmente factible aunque realmente complicado de aplicar.

5.5. El control como evaluación

El seguimiento es una faceta importante de cualquier evaluación o puede ser, incluso, el elemento básico de la evaluación. Quizá realmente es más importante la monitorización que la evaluación de impacto. Cada proyecto debe ser monitorizado de una manera insesgada y detallada, con *feedback* continuo, e independientemente de cuán brillante pueda ser la intervención. ¿Qué significa monitorizar? Consiste en realizar un seguimiento a partir de una base de datos de gestión de información o Management Information System (MIS) (sistema de información para la gestión). El MIS es muy importante, pues estos sistemas permiten saber qué va bien, qué va mal, qué ha funcionado correctamente con anterioridad, etc. Ha habido un par de estudios de proyectos en Indonesia, por ejemplo en zonas urbanas, donde el análisis se ha realizado comparando un CDD con una monitorización de arriba abajo y con un control exclusivamente local. Una monitorización que alcanza desde las clases sociales más altas a las más bajas (los que se podría llamar una monitorización de «arriba abajo») y con un control local y, por tanto, descentralizado. Obviamente, no se ve el impacto de más monitorización en la comunidad si lo comparamos con

la auditoría normal. El impacto adicional de una monitorización es menor que el impacto normal de una auditoría. Por lo tanto no se puede inferir que el control de una comunidad funciona peor que la monitorización de arriba abajo.

Por otra parte, los facilitadores del proyecto, los que hacen la implementación, deben ser antropólogos, sociólogos, economistas, ingenieros, estadísticos, contables, etc. La aproximación multidisciplinar proporciona una enorme fuente de información. Los buenos proyectos utilizarán esta información para aprender a hacer una monitorización mejor. ¿Y cómo ayuda al proceso? Pues aprendiendo al hacerlo, incluso cuando el proyecto empieza mal, pues un buen seguimiento puede corregir errores, rectificando su evolución.

5.6. Los métodos cualitativos

Por lo general, la parte más importante de una evaluación se basa en métodos cuantitativos que miden el consumo, la pobreza, etc. Sin embargo, en muchas ocasiones algunos de los elementos importantes de la evaluación deben obtenerse a partir de procedimientos cualitativos. Hay muchas preguntas que consideramos clave y que deben analizarse a partir de métodos cualitativos. Es importante señalar que el método cualitativo es un método complementario³ y, por tanto, añade coste a la evaluación. De hecho todos los métodos que hemos discutido hasta el momento son procedimientos complementarios a los estrictamente cuantitativos.

Otro procedimiento complementario es la *discusión en grupos*. En un *focus group* la gente, seleccionada adecuadamente, proporciona mucha información sobre la localidad, de una forma poco sesgada. Se puede seleccionar a 10 personas de manera aleatoria, juntarlas en una sala y comenzar la sesión después de informarlas de los objetivos que se persiguen. Otra gran ventaja de los *focus groups* es la posibilidad de adquirir información sin necesidad de utilizar un cuestionario cerrado y de forma interactiva.

³ Rao y Woolcock (2004) presentan una propuesta de integración del análisis cualitativo y cuantitativo.

Otro método de investigación cualitativa es la *entrevista en profundidad* con personas clave de la comunidad. Seguramente esto nos daría una idea clara sobre la naturaleza del proyecto y su aplicabilidad. Por otra parte, se puede utilizar el *análisis de textos*. ¿Qué dicen los periódicos locales, en la lengua local, sobre el proyecto? ¿Cómo ven el proyecto? De esta forma, por ejemplo, en algún proyecto en la India hemos constatado que la prensa local planteaba que produciría una pérdida de rentas, que se estaba desarrollando incorrectamente o que, en su entorno, existía bastante corrupción. Todo esto se supo porque se había hecho una lectura sistemática de la prensa local.

¿Cuál es el valor añadido de lo cualitativo? Normalmente hay que pensar en forma cuantitativa, pero hay que actuar en forma cualitativa. La utilización de métodos cualitativos no significa que no haya que realizar un buen diseño de la parte cuantitativa del proyecto. No significa que no se pueda utilizar el método de experimentos aleatorizados o no se busque un experimento natural o una discontinuidad. Todos los procedimientos son complementarios en nuestro deseo de conocer mejor cómo está evolucionando el proyecto y sus efectos.

El problema de los métodos cualitativos es que la información es profunda, densa, y no se pueden hacer a gran escala. Este procedimiento sólo se puede aplicar a un subconjunto de la población y, por tanto, es un método parcial que no podría ser la única base de la evaluación. Los métodos cuantitativos nos permiten realizar encuestas a gran escala, pero sobre cuestionarios cerrados. Por tanto, para extrapolar y generalizar son más adecuados. Pero hay que complementarlos con métodos cualitativos. Los métodos cualitativos pueden, además, ayudarnos a preparar una estrategia, y a identificar una fórmula adecuada de identificación de los efectos del programa.

¿Cómo encontramos estas estrategias? Un ejemplo: los tres días en una comunidad de 2.000 personas en Calcuta en un estudio sobre trabajadoras del sexo. Los métodos cualitativos permitieron ajustar la estrategia para la formación del grupo de control del estudio cuantitativo aleatorizado que se intentaba llevar a cabo. Las personas iban por callejuelas, entraban en ocho o nueve burdeles que estaban en varios edificios. Había clientes pobres y ricos. Era un sistema social complejo y, por tanto, la

aleatorización sin la visión cualitativa hubiera seguramente fracasado. Estos métodos ayudan, por lo tanto, a mejorar los resultados de la evaluación. Sería necesario pasar mucho tiempo sobre el terreno para poder tener una idea de cómo formular preguntas y qué preguntas hay que hacer. Hay que sentarse a observar y escuchar y hablar antes de diseñar un instrumento. Y luego hacer un seguimiento del proceso, viendo los diferentes cambios que tuvieron lugar durante el transcurso de la evaluación.

Finalmente está lo que llamamos la *triangulación*. Es la comparación de conclusiones obtenidas a partir de diferentes fuentes de información. Si los resultados cuantitativos y los cualitativos proporcionan efectos similares entonces podemos tener más confianza en que nuestras conclusiones son correctas. Esto es lo que llamamos triangulación. Dos tipos diferentes de datos para llegar a una misma conclusión.

5.7. El desarrollo basado en la comunidad o CDD

El desarrollo basado en la comunidad, o *Community Driven Development*, está muy de moda en la actualidad.⁴ ¿Qué significa desarrollo basado en la comunidad? Significa que vamos a dar a la comunidad el poder de tomar las decisiones sobre temas que les conciernen. Es lo que tradicionalmente se resume en la frase «enseñar al pescador a pescar en lugar de regalarle peces». Enseñar a las comunidades a ocuparse de sí mismas en cuanto a proyectos que van a durar un determinado tiempo. Se les proporciona el poder político y el poder económico. ¿Qué es lo que realmente significa? En primer lugar que es la propia comunidad la que busca los colectivos objeto. Es la propia comunidad la que selecciona a los que deben ser los beneficiarios del programa. ¿Quién se beneficia, de qué y cuándo? ¿Quién se beneficia del proyecto? La comunidad lo decide y no viene asignado desde fuera, como sucedería en los proyectos tradicionales financiados por los organismos internacionales. En segundo lugar, la gestión del proyecto

⁴ Mansuri y Rao (2004) discuten, desde una perspectiva crítica, los proyectos basados en la comunidad.

también se basa en la comunidad. En resumen, se le proporcionan a los habitantes de una aldea o región los fondos para mejorar su situación económica. Ellos son los encargados de crear el proyecto, lo construyen y deciden los beneficiarios, lo miden, facilitan, reciben la ayuda del gobierno. Esto es fundamental: se les da el poder a los potenciales beneficiarios de la ayuda al desarrollo. A lo largo de los últimos diez años, el Banco Mundial ha aumentado muy significativamente el volumen de los fondos concedidos para desarrollo basado en la comunidad hasta alcanzar los 7.000 millones de dólares. Estamos hablando, por lo tanto, de una parte enorme de la cartera de actividades del Banco Mundial. Pero ¿qué sabemos del impacto de los CDD? La verdad es que el número de proyectos CDD ha aumentado mucho en el mundo sin que se hayan realizado estudios sobre sus efectos.

Normalmente, la utilización del CDD se basa en argumentos teóricos razonables pero a priori. En principio, se supone que los beneficiarios se eligen apropiadamente, dado que es la propia comunidad la que decide sobre los mismos. En segundo lugar, un CDD debería mejorar la oferta y calidad de los servicios públicos. En tercer lugar, a la vez que mejora los servicios públicos (por ejemplo, las escuelas), mejora la capacidad para la acción colectiva (o capital social): genera amistades entre miembros de la comunidad que tienen que trabajar juntos para gestionar el proyecto, lo que permite mejorar las relaciones y provocar un círculo virtuoso de formación de capital social. También se plantean proyectos más sostenibles. Por otra parte, se supone que, en este caso, es menos probable que el proyecto sea captado por la élite o la gente poderosa de la aldea. Finalmente, el proyecto se puede escalar o aumentar con facilidad.

¿Cuál es la evidencia hasta la fecha sobre estas supuestas ventajas de los proyectos basados en la comunidad? Hace muy poco que se ha empezado a acumular información para responder a esta pregunta. La mayoría de la evidencia proviene del análisis de los fondos sociales, que son proyectos de desarrollo financiados por el Banco Mundial. En primer lugar, sabemos que es más fácil si el objetivo son comunidades pobres que si el objetivo se define en términos de los pobres dentro de una determinada comunidad. Es relativamente fácil encontrar una comunidad pobre que debe ser beneficiaria de un programa. Es muy difícil detectar a los benefi-

ciarios últimos de proyectos, los más pobres dentro de la comunidad. Con todo, en términos generales, la selección de beneficiarios basada en las decisiones de la comunidad funciona mejor que el procedimiento basado en métodos externos.

La heterogeneidad es muy importante en este tipo de proyectos. Lo hemos visto, por ejemplo, en una comunidad de 2.000 familias vinculadas unas con otras con lazos familiares. Y éste fue el peor proyecto jamás visto porque todos se quejaban de los demás. El proyecto finalmente se colapsó. En general, en los CDD se observa una búsqueda de beneficiarios pobres basados en las preferencias. La comunidad recibe lo que necesita, no lo decide el gobierno. Lo decide la propia comunidad, porque de esta forma encaja mejor en sus preferencias y necesidades que lo que recibe del gobierno.

En cuanto a la entrega de servicios, vemos que los proyectos en los que la comunidad participa son normalmente proyectos eficaces y efectivos: mejora de la calidad de agua, saneamiento de aguas fecales, etc. No obstante, no hay pruebas de que las alternativas no participativas funcionen mejor o peor que las participativas, ni tampoco hay evidencia de que la participación sea la causa de las mejoras en el proyecto. Esto todavía es objeto de investigación.

En cuanto a la participación, hay muchas pruebas de que altos niveles de participación están positivamente correlacionados con la eficacia y efectividad del proyecto. No obstante, no existe evidencia convincente de que dicha relación sea causal. En cuanto a la creación de capital social, hay muchas pruebas de una correlación fuerte entre la formación del capital social y la efectividad del proyecto, pero no existe evidencia convincente en cuanto a la existencia de una relación causal entre ambas variables. En cuanto a la sostenibilidad la evidencia es, también, muy limitada. Hay algún trabajo antropológico que señala el papel crucial del soporte de niveles superiores de gobierno en el apoyo de los proyectos comunitarios para que dichos proyectos mejoren. En cuanto a la captura de fondos por parte de las élites, que se supone limitada, se ha realizado una serie de ensayos que muestran que la élite domina, lo que no implica necesariamente la captura del proyecto. Cuando se da dinero a una comunidad, normalmente son las personas más educadas, son las élites de la aldea, las

que se orientan a ello y toman el control de estos fondos. No significa necesariamente que se apropien de los fondos. Puede ser la única forma en que se pueda desarrollar el proyecto, pues quizás las élites son las únicas personas preparadas para entender inicialmente la actuación. La captura puede tener un elemento negativo, o un elemento de buena intención (no malintencionada) en la apropiación del proyecto. Lo importante es el papel del facilitador, la persona que hace el trabajo de desarrollo comunitario; que va a la aldea, los moviliza, los enseña a gestionar un presupuesto y a realizar la contabilidad. Los buenos facilitadores son líderes carismáticos, con dotes docentes, y con conocimientos de antropología, ingeniería, economía y contabilidad.

Finalmente, en cuanto al aumento de la dimensión del proyecto a partir de una actuación limitada existe muy poca experiencia. De hecho, la expansión de este tipo de proyectos se encuentra con multitud de dificultades. En muchas ocasiones la necesidad de un gran número de facilitadores provoca la aceptación de gente con poca experiencia. En algún caso analizado eran niños que terminan doce años de formación y no encuentran trabajo. Entran en la gestión del proyecto porque realmente no se encuentran suficientes personas adecuadas para el trabajo, dado el aumento de la dimensión del proyecto.

Lo que no parece una aproximación razonable es pensar que cualquier problema tiene solución si se plantea tomando como base la comunidad. Un ejemplo es qué hacemos en una situación de conflicto como el de Sierra Leona. En principio, podría parecer que, dados los problemas generados por el conflicto, el desarrollo basado en la comunidad no funcionaría. Sin embargo, la evaluación realizada hasta el momento parece indicar que está teniendo buenos resultados, aunque el motivo para iniciar el proyecto fuera el supuesto general de que los proyectos basados en la comunidad siempre funcionan.

El segundo ejemplo es Rusia. En este caso, el problema era que la electricidad no estaba llegando a zonas rurales y, por lo tanto, los negocios locales no podían funcionar bien sin electricidad o con un suministro eléctrico poco fiable. ¿Qué se propuso?, un programa de desarrollo comunitario.

O sea que hay una cultura de desarrollo comunitario que se está contagiando a todo el mundo. Podríamos llamarlo desarrollo propulsado por la

demanda y la oferta. En general, los proyectos de desarrollo comunitario funcionan de abajo arriba, es decir, empezando por los más desfavorecidos. Pero cuando los bancos empiezan a participar en estos proyectos no se diferencian de ninguna otra iniciativa. Los bancos no funcionan con una estructura de control de abajo arriba sino de arriba abajo. Por tanto, cuando los bancos comienzan a participar en estos proyectos los mismos empiezan a perder valor y se convierten en proyectos convencionales.

El último punto es la evaluación. ¿Cómo podemos evaluar este tipo de proyectos? Es un reto, porque el desarrollo comunitario se supone que tiene muy buenos resultados (mejora las vidas de los pobres, mejora la infraestructura de los pueblos, acumula capital social, mejora los métodos de gobierno, reduce la corrupción, aumenta el poder de la gente del pueblo, etc.), pero es necesario probar dichos efectos con evidencia empírica. Aunque, teóricamente, puede tener todos estos beneficios, en muchas ocasiones no se consiguen. Para concretar las cuestiones sobre la evaluación de proyectos basados en la comunidad, se puede tomar un ejemplo. El programa UPP2-Indonesia es un proyecto urbano. En Indonesia, el lugar donde se lleva a cabo, el concepto de zona urbana es extraño porque todas estas zonas parecen rurales. No obstante, la definición de algo urbano en Indonesia es cualquier localidad de más de 10.000 habitantes. La mayor parte de estas zonas urbanas son realmente pueblos grandes. Este proyecto cuenta con un presupuesto de 200 millones de dólares que se deben repartir en tres provincias. La unidad de recepción de fondos es el *Keurahan* y el presupuesto son unos 20.000 dólares por comunidad de 1.000 personas. Se elige un comité local para gestionar el programa y se hace una reunión en la que se decide quién es el presidente. El proceso de elección de cargos es bastante informal. El comité empieza a elegir grupos de infraestructura en el pueblo: gente que va a construir diferentes infraestructuras. Se les dan pequeños subsidios, créditos, de 1.000 dólares, por ejemplo, que utilizan para la construcción. Algunas personas donan su tiempo, fuerza de trabajo, comida o lo que puedan y así se consigue poco a poco ir reuniendo los fondos y los medios necesarios para construir la infraestructura. También hay grupos de microcréditos que trabajan en la microfinanciación. Luego hay una evaluación cuantitativa. Cuando se hizo la evaluación, el gobierno decidió ampliar el programa a todo el país y en-

tonces perdió algo de sentido la evaluación parcial, aunque se siguió con ella. En este punto, la política también desempeñó un papel importante, pues el presidente fue elegido, en gran parte, precisamente por apoyar la expansión del proyecto a la totalidad de Indonesia.

La evaluación de este tipo de proyectos es muy compleja. Antes de preparar las preguntas de los cuestionarios, los investigadores nos dividimos en diferentes provincias y nos pasamos tres semanas estudiando la primera fase del proyecto para ver cuáles podrían ser los impactos, antes de hacer los cuestionarios. Éstos eran muy complicados: incluían cuestiones relativas al proyecto, preguntas sobre otros proyectos gubernamentales y comunitarios que pudieran estar relacionados, la participación en los grupos de crédito, cuestiones socioeconómicas, etc. Los cuestionarios dirigidos a líderes y activistas incluían también una serie de preguntas sobre las cuestiones de la oligarquía, las actividades comunitarias, los facilitadores, las redes, la percepción de la corrupción, las horas dedicadas a proyectos concretos, el dinero dedicado o invertido en proyectos, y todo tipo de preguntas que surgen para medir los impactos del proyecto.

El aspecto cuantitativo se trata a partir de un diseño de discontinuidad. La muestra contiene aproximadamente doscientas observaciones en el grupo tratado y doscientas en el grupo de control. Se realizan tres fases: la inicial, la intermedia y la continuación. El aspecto cualitativo tenía que estar incluido en el estudio cuantitativo. No se puede hacer algo cualitativo en todas las comunidades, así que se eligieron sólo 12 grupos de control y 12 de tratados. Era un equipo de cuatro personas que invertían diez días de estudio en cada uno de estos grupos o pueblos. Además había una serie de reuniones de discusión y entrevistas en profundidad (unas veinte o treinta). Se cartografió la comunidad en términos de las características de cada zona.

Los resultados de la etapa inicial muestran que los bienes públicos se encuentran financiados en un 40% por las contribuciones del gobierno y en un 60% por la comunidad. Los niveles de participación cívica son altísimos. Los indonesios tienen incluso en su constitución un concepto que llaman *gotong royong*, que es la noción de participación comunitaria y cívica. Esto se aplica desde el colegio, y ya desde la educación primaria se inculca este principio ético. Por tanto, hay altos niveles de participación y

también un gran solapamiento entre las instituciones religiosas y las gubernamentales. Muchas de las reuniones iniciales se hicieron en mezquitas o en iglesias. En las mezquitas y en las iglesias, los líderes religiosos eran muchas veces los que lideraban el proceso. Por eso el islamismo en Indonesia se llama el islamismo cívico.

5.8. Conclusiones

Los métodos científicos no debemos interpretarlos demasiado literalmente. Debemos interpretarlos en el contexto político, social y cultural, para averiguar si se pueden reemplazar los programas por otras alternativas más efectivas y ver cómo se pueden mejorar los proyectos y retroalimentar el proceso. Por tanto, el análisis cuantitativo debe complementarse con el análisis cualitativo, pues incluso los procedimientos de evaluación aparentemente más precisos tienen problemas. Los experimentos aleatorizados, que son el método más pujante en la actualidad, tienen limitaciones de tipo ético y, además, existen situaciones en las cuales no se pueden aplicar. Incluso cuando se pueden aplicar no está claro si la consistencia interna de la evaluación nos dice algo sobre su aplicabilidad externa, fuera del caso estudiado.

Los métodos cualitativos son múltiples y complementan y mejoran los resultados obtenidos por medio del análisis cuantitativo. En particular, la confección de los cuestionarios que posteriormente serán utilizados para el análisis cuantitativo requiere de la consideración de técnicas cualitativas para captar efectos o condicionantes (tanto políticos como sociales) que no se habrían previsto en una evaluación estrictamente cuantitativa.

Por último, este capítulo discute la eficacia y evaluación de los llamados proyectos de desarrollo comunitario. A pesar del evidente atractivo a priori de este tipo de actuaciones, la evidencia empírica existente es parcial e incompleta. Por esto, es necesario seguir trabajando en el análisis del impacto de este tipo de programas que, en todo caso, no pueden ser la panacea para cualquier problema asociado al desarrollo de las sociedades más desfavorecidas.

Bibliografía

- DUFLO, E., R. GLENNERSTER, y M. KREMER. «Using Randomization in Development Economics: a Toolkit». CEPR Discussion Paper 6059, 2007.
- MANSURI, G., y V. RAO. «Community-based and Driven Development: a Critical Review». *World Bank Economic Observer* 19 (1), 2004: 1-39.
- RAO, V., y M. WOOLCOCK. «Integrating Qualitative and Quantitative Approaches to Program Evaluation». En Bourguignon y Pereira da Silva, eds., *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*. Oxford: World Bank and Oxford University Press, 2004.
- RAVALLION, M. «Evaluating Anti-Poverty Programs». En Schultz y Strauss, eds., *Handbook of Development Economics*. De próxima aparición.

Bibliografía adicional

Esta bibliografía adicional sobre los temas tratados en este volumen, aunque no ha sido citada en los capítulos anteriores, puede ser de utilidad para adentrarse más en el fascinante mundo de la moderna teoría del desarrollo económico y la evaluación de programas y proyectos de ayuda al desarrollo. La bibliografía se divide en varias partes: la primera tiene carácter general sobre desarrollo económico y evaluación; la segunda presenta una selección de publicaciones de Michael Kremer, autor del segundo capítulo, sobre estos temas; la tercera parte contiene una selección de las publicaciones más relevantes de Esther Duflo, autora del tercer capítulo; la cuarta parte presenta las publicaciones más relevantes de Rohini Pande. Finalmente, la última parte contiene una selección de trabajos de Vijayendra Rao.

1. Publicaciones generales

ACEMOGLU, D., J. SIMON, y A. R. JAMES. «The Colonial Origins of Comparative Development: An Empirical Investigation». *American Economic Review* 91, 2001: 1369-1401.

—. «Reversal of Fortunes: Geography and Institutions in the Making of Modern World Income Distribution». *Quarterly Journal of Economics* 117, 2002.

ALESINA, A., y D. DOLLAR. «Who Gives Foreign Aid to Whom and Why». *Journal of Economic Growth* 5, 2000: 33-63.

ANGRIST, J. «American Education Research Changes Tack». *Oxford Review of Economic Policy* 20 (2), 2004.

- ANGRIST, J., E. BETTINGER, E. BLOOM, y Elizabeth KING. «Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment». *American Economic Review* 92 (5), 2002: 1535-1558.
- BALAND, J., y P. FRANCOIS. «Rent Seeking and Resource Booms». *Journal of Economic Development* 61, 2000: 527-542.
- BANERJEE, A., y E. DUFLO. «Addressing Absence». *The Journal of Economic Perspectives* 20 (1), 2006: 117-132.
- . *Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program*. MIT, 2004 [mimeo.].
- BANERJEE, A., S. COLE, E. DUFLO, y L. LINDEN. «Remedying Education: Evidence from Two Randomized Experiments in India». De próxima aparición en *Quarterly Journal of Economics*.
- BANERJEE, A., R. BANERJI, E. DUFLO, R. GLENNERSTER, y S. KHEMANI. *Can Information Campaigns Spark Local Participation and Improve Outcomes: a Study of Primary Education in Uttar Pradesh*. World Bank Policy Research, Documento de Trabajo, 3967, 2006.
- BECK, T., R. LEVINE y N. LOAYZA. «Financial Intermediation and Growth: Causality and Causes». *Journal of Monetary Economics*, vol. 46, 2000: 31-77.
- BECK, T., A. DEMIRGUC-KUNT, y M. MARTÍNEZ PERIA. «Banking Services for Everyone? Barriers to Bank Access and Use around the World». World Bank, 2006 [mimeo.].
- BLEAKLEY, H. «Disease and Development: Evidence from Hookworm Eradication in the American South». *Quarterly Journal of Economics*, febrero 2007.
- BOONE, P. «Politics and the Effectiveness of Foreign Aid». *European Economic Review* 40, 1996: 289-329.
- BURGESS, R., y R. PANDE. «Can Rural Banks Reduce Poverty? Evidence from the Indian Social Banking Experiment». *American Economic Review*, 2005.
- BURNSIDE, C., y D. DOLLAR. «Aid, Policies and Growth». *American Economic Review*, 90 (4), septiembre 2000: 847-888.
- CHAUDHURY, N., J. HAMMER, M. KREMER, K. MURALIDHARAN, y H. ROGERS. «Missing in Action: Teacher and health Worker Absence in Developing Countries». *The Journal of Economic Perspectives* 20 (1), 2006: 91-116.
- COLLIER, P., y D. DOLLAR. «Aid Allocation and Poverty Reduction». *European Economic Review* 46, 2002: 1475-1500.
- DEATON, A. *The Analysis of Household Surveys: a Microeconometric Approach to Development Economics*. John Hopkins University Press and the World Bank, 1997.
- DICKS-MIREAUX, L., M. MECAGNI, y S. SHANDLER. «Evaluating the Effect of IMF lending to Low-income Countries». *Journal of Development Economics* 61, 2000: 495-526.

- DJANKOV, S., J. G. MONTALVO, y M. REYNAL-QUEROL. «Does Foreign Aid Help?». *Cato Journal* 26 (1), 2006: 1-28.
- DUFLO, E., P. DUPAS, M. KREMER, y S. SINEI. «Education and HIV/AIDS Prevention: Evidence from a Randomized Evaluation in Western Kenya». World Bank Policy Research Documento de Trabajo 4024, 2006.
- EASTERLY, W. *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*. MIT Press, 2002.
- . «The Cartel of Good Intentions: The problem of Bureaucracy in Foreign Aid». *Journal of Policy Reform* 5 (2), 2003: 67-91.
- . *The white Man's Burden: Why the West's Efforts to Aid the Rest Have Done so Much Ill and so Little good*. Nueva York: Penguin Press, 2006.
- EASTERLY, W., y R. LEVINE. «Africa's growth Tragedy: Policies and Ethnic Divisions». *Quarterly Journal of Economics* 112, 1997: 1203-1250.
- . «Tropics, Germs and Crops: How Endowments Influence Economic Development». *Journal of Monetary Economics* 50, 2003: 3-39.
- EASTERLY, W., R. LEVINE, y D. ROODMAN. «Aid, Policies and Growth: a Comment». *American Economic Review* 94 (3), 2003: 774-780.
- GERTLER, P. *Final Report: The Impact of PROGRESA on Health*. Washington: International Food Policy Research Institute, 2000.
- . «Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment». *American Economic Review (Papers and Proceedings)* 94 (2), 2004: 336-341.
- GIBSON, C., K. ANDERSSON, E. OSTROM, y S. SHIVAKUMAR. *The Samaritan's Dilemma: the Political Economy of Development Aid*. Oxford: Oxford University Press, 2005.
- GINE, X., y D. KARLAN. *Group Versus Individual Liability: a Field Experiment in the Philippine*. 2006 [mimeo.].
- GLEWWE, P., A. Park, y M. ZHAO. «The Impact of Eyeglasses on the Academic Performance of Primary School Students: Evidence from a Randomized Trial in Rural China». Documento de Trabajo, 2006.
- HANSEN, H., y F. TARP. «Aid and Growth Regressions». *Journal of Development Economics*, vol. 64, 2001: 547-557.
- HONOHAN, P. «Financial Development, Growth and Poverty: How Close are the Links?». En Charles Goodhart, ed., *Financial Development and Economic Growth: Explaining the Links*. Londres: Palgrave, 2004, 1-37.
- JENSEN, R. *Do the Perceived Return to Education Affect Schooling Decisions? Evidence from a Randomized Experiment*. Harvard University, 2006 [mimeo.].
- KARLAN, D., y J. ZINMAN. *Expanding Credit Access: Using Randomized Supply Decisions To Estimate the Impacts*. Princeton, 2006 [mimeo.].

- KING, R. G., y R. LEVINE. «Financial and growth: Schumpeter Might be Right». *The Quarterly Journal of Economics*, vol. 108, 1993: 717-727.
- KLEIN, M., y T. HARFORD. *The Market for Aid*. Washington: International Financial Corporation, 2005.
- KNACK, S. «Aid Dependence and the Quality of Governance: a Cross Country Empirical test». *Southern Economic Journal* 68, 2001: 310-329.
- KREMER, M., y E. MIGUEL. «Illusion of Sustainability». De próxima aparición en *Quarterly Journal of Economics*.
- . «Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities». *Econometrica* 72 (1), enero 2004: 159-217.
- KREMER, M., y R. THORNTON. *Incentives to Learn*. Harvard University, 2005 [mimeo.].
- MARTENS, B., U. MUMMERT, P. MURRELL, y P. SEABRIGHT. *The Institutional Economics of Foreign Aid*. Cambridge: Cambridge University Press, 2002.
- MONTALVO, J. G. «Regional Evolutions in Labor Markets in the Philippines: a Dynamic Approach». *Journal of Asian Economics* 17 (3), 2006: 448-477.
- MONTALVO, J. G., y M. REYNAL-QUEROL. «Religious Polarization and Economic Development». *Economic Letters*, 2003.
- . «Ethnic polarization, Potential Conflict and Civil War». *American Economic Review* 95 (3), 2005: 796-816.
- . «Ethnic Diversity and Economic Development». *Journal of Development Economics*, 2005.
- . «Fighting Against Malaria: Prevent Wars While Waiting for the Miraculous Vaccines». *Review of Economics and Statistics* 89 (1), 2007: 165-177.
- MONTALVO, J. G., S., LINDELOW, y A. BARR. «For Public service or Money: Geographical Imbalances in the Health Workforce of Ethiopia». *Health Policy and Planning* 22 (3), 2007: 128-138.
- OLKEN, B. «Monitoring Corruption: Evidence from a Field Experiment in Indonesia». *Journal of Political Economy* 115 (2), 2007: 200-248.
- . «Corruption and the Costs of Redistribution: Micro Evidence from Indonesia». *Journal of Public Economics* 90 (4-5), 2006: 853-870.
- PAULSON, A., y R. M. TOWNSEND. «Entrepreneurship and Financial Constraints in Thailand». *Journal of Corporate Finance*, 2004 (3).
- PITT, M. y S. KHANDKER. «The Impact of Group Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?». *Journal of Political Economy* 106 (5), 1998: 958-996.
- RAVALLION, M. «Growth, inequality and Poverty: Looking Beyond the Averages». *World Development* 29 (11), 2001: 1803-1815.

- RAVALLION, M. «The Mystery of Vanishing Benefits: an Introduction to Impact Evaluation». *World Bank Economic Review* 15 (1), 2001: 115-140.
- . «Measuring Aggregate Welfare in Developing Countries: How Well do National Accounts and Surveys Agree». *Review of Economics and Statistics* 85, 2003: 355-361.
- . «Evaluating Anti-poverty Programs». En Schultz, ed. *Handbook of Agricultural Economics*, vol. 4. De próxima aparición.
- RAVALLION, M., y S. CHEN. «How Did the World's Poor Fare in the 1990s?». *Review of Income and Wealth* 47 (3), 2001: 283-300.
- . «Measuring Pro-poor Growth». *Economic Letters* 78 (1), 2003: 93-99.
- . «China's Uneven Progress Against Poverty». *Journal of Development Economics*, 2006.
- RAVALLION, M., y G. DATT. «Why Has Economic Growth Being More Pro-poor in Some States of India than in Others?». *Journal of Development Economics* 68 (2002): 381-400.
- RAVALLION, M., y E. GALASSO. «Decentralized Targeting of an Anti-poverty Program». *Journal of Public Economics* 85, 2005: 705-727.
- RAVALLION, M., y E. GALAZO. «Social Protection in a Crisis: Argentina's Plan Jefes y Jefas». *World Bank Economic Review* 18 (3), 2004: 367-399.
- RAVALLION, M., y J. JALAN. «Does Piped Water Reduce Diarrhea for Children in Rural India». *Journal of Econometrics* 112, 2003: 153-173.
- . «Estimating the Benefit Incidence of an Anti-poverty Program by Propensity Score Matching». *Journal of Business and Economic Statistics* 21 (1), 2003: 19-30.
- RAVALLION, M., y M. LOKSHIN. «Testing Poverty Lines». *Income and Wealth* 52 (3), 2006: 399-421.
- RAVALLION, M., y D. VAN DE WALLE. «Creating a Market: Land Reallocation in Vietnam's Agrarian Transition». *Economic Journal*, 2006.
- REINIKKA, R., y J. SVENSSON. «Local Capture: Evidence from a Central Government Transfer Program in Uganda». *Quarterly Journal of Economics* 119 (2), 2004: 679-705.
- . «Fighting Corruption to Omprove Schooling: Evidence from a Newspaper Campaign in Uganda». *Journal of the European Economic Association* 3, 2005: 259-267.
- REYNAL-QUEROL., M. «Ethnicity, Political Systems and Civil Wars». *Journal of Conflict Resolution* 46 (1), 2002: 29-54.
- SACHS, J., y A. WARNER. «The Curse of Natural Resources». *European Economic Review* 45, 2001: 827-838.
- SCHULTZ, P. «School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program». *Journal of Development Economics* 74, 2004: 199-250.

- SVENSSON, J. «When is Foreign Aid Credible? Aid Dependence and Conditionality». *Journal of Development Economics* 61, 2000: 61-84.
- «Foreign Aid and Rent-seeking». *Journal of International Economics* 51, 2000: 437-461.
- «Why Conditional Aid Does Not Work and What Can Be Done About it?». *Journal of Development Economics* 70, 2003: 381-402.

2. Selección de publicaciones de Michael Kremer

- «Illusion of Sustainability». *Quarterly Journal of Economics* 112 (3), 2007: 1007-1065 (Con E. Miguel.).
- «Advance Market Commitments for Vaccines Against Neglected Diseases: Estimating Costs and Effectiveness». (Con Ernst Berndt, Rachel Glennerster, Jean Lee, Ruth Levine, Georg Weizsäcker, y Heidi Williams.) De próxima aparición en *Health Economics*.
- «Creating Markets for Vaccines». (Con Rachel Glennerster, y Heidi Williams.) *Innovations* 1 (1), 2006: 67-79.
- «Advance Market Commitments: A Policy to Stimulate Investment in Vaccines for Neglected Diseases». (Con Owen Barder, y Heidi Williams.) *The Economists' Voice* 3 (3), 2006.
- «Applying the Odious Debt Doctrine while Preserving Legitimate Lending». (Con Seema Jayachandran, y Jonathan Shaftter.) De próxima aparición en *Ethics & International Affairs*.
- «Asian Growth and African Development». (Con Marcos Chamon.) *AER Papers and Proceedings* 96 (2), junio 2006: 400-404.
- «Missing in Action: Teacher and Health Worker Absence in Developing Countries». (Con Nazmul Chaudhury, Jeffrey Hammer, Karthik Muralidharan, y F. Halsey Rogers.) *Journal of Economic Perspectives* 20 (1), invierno 2006: 91-116.
- «Odious Debt». (Con Seema Jayachandran.) *American Economic Review* 96 (1), marzo 2006: 82-92.
- «Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia». (Con Josh Angrist, y Eric Bettinger.) *American Economic Review* 93 (3), junio 2006: 847-862.
- «Schools, Teachers, and Education Outcomes in Developing Countries». (Con Paul Glewwe.) De próxima aparición en *Handbook on the Economics of Education*, Elsevier.

- «Teacher Absence in India: A Snapshot». (Con Nazmul Chaudhury, Jeffrey Hammer, Karthik Muralidharan, y F. Halsey Rogers.) *Journal of the European Economic Association* 3(2-3), mayo 2005: 658-667.
- «Ways to Spur Malaria Vaccine Research». (Con Rachel Glennerster.) De próxima aparición en *The Economics of Malaria*.
- «Retrospective vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya». (Con Paul Glewwe, Sylvie Moulin, y Eric Zitzewitz.) *Journal of Development Economics* 74 (1), junio 2004: 251-268.
- «Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities». (Con Edward Miguel.) *Econometrica* 72 (1), enero 2004: 159-217.
- «Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons». *American Economic Review* 93 (2), mayo 2003: 102-106.
- «Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment». (Con Josh Angrist, Eric Bettinger, Erik Bloom, y Elizabeth King.) *American Economic Review* 92 (5), diciembre 2002: 1535-1558. NBER Documento de Trabajo 8343, 2001.
- «Pharmaceuticals and the Developing World». *Journal of Economic Perspectives* 16 (4), otoño 2002.
- «Income Distribution Dynamics with Endogenous Fertility». (Con Daniel Chen.) *Journal of Economic Growth* 7 (3), septiembre 2002: 227-258. NBER Documento de Trabajo 7530, 2000.
- «The Impact of Development Assistance on Social Capital: Evidence from Kenya». (Con Mary Kay Gugerty.) En Christian Grootaert, y Thierry van Bastelaer, eds., *The Role of Social Capital in Development: An Empirical Assessment*. Cambridge University Press, abril 2002.
- «A Purchase Commitment for Vaccines». En Inge Kaul, Katell Le Goulven, y Mirjam Schnupf, eds., *Global Public Goods Financing: New Tools for New Challenges. A Policy Dialogue*, UNDP/ODS, 2002.
- «Creating Markets for New Vaccines: Part I: Rationale». En Adam B. Jaffe, Josh Lerner, y Scott Stern, eds., *Innovation Policy and the Economy*, vol. 1, MIT Press, 2001 (mayo 2000; borrador).
- «Creating Markets for New Vaccines: Part II: Design Issues». En Adam B. Jaffe, Josh Lerner, y Scott Stern, eds., *Innovation Policy and the Economy*, vol. 1, MIT Press, 2001 (mayo 2000; borrador).
- «Searching for Prosperity». *Carnegie-Rochester Conference Series on Public Policy* (55), 2001: 275-303.

- «Stimulating Industrial R&D for Neglected Infectious Diseases: Economic Perspectives». (Con David Webber.) *Bulletin of the World Health Organization* 79 (8), 2001: 693-801.
- «Education Reform». En Anne O. Krueger, ed., *Economic Policy Reform: The Second Stage*, University of Chicago Press, 2000.
- «Outside Funding of Community Organizations: Benefiting or Displacing the Poor?». (Con Mary Kay Gugerty.) NBER Documento de Trabajo 7896, septiembre 2000.
- «A World Bank Vaccine Commitment». (Con Rachel Glennerster.) Brookings Institution Policy Brief #57, mayo 2000.
- «A Better Way to Spur Medical Research and Development». (Con Rachel Glennerster.) *Regulation* 23 (2), 2000: 34-39.
- «Patent Buyouts: A Mechanism for Encouraging Innovation». *Quarterly Journal of Economics*, noviembre 1998: 1137-1167. NBER Documento de Trabajo 6304, 1997. Reimpreso en Alexander Tabarrok, ed., *Entrepreneurial Economics: Bright Ideas from the Dismal Science*. Oxford University Press, 2002.
- «On the Effect of Changing Sexual Activity on HIV Prevalence». (Con Charles Morcom.) *Mathematical Biosciences* 151 (1), julio 1998: 99-122.
- «AIDS: The Economic Rationale for Public Intervention». En Martha Ainsworth, Lieve Franssen, y Mead Over, eds., *Confronting AIDS: Evidence from the Developing World*. The European Commission and the World Bank, 1998.
- «AIDS, Behavioral Choice, and the Composition of the Pool of Available Partners». *Advances in Mathematical Population Dynamics: Molecules, Cells and Man Part II, Chapter 8*. World Scientific Publishing Co., 1997: 487-510.

3. Selección de publicaciones de Esther Duflo

- «The economic life of the poor». (Con A. Banerjee.) *The Journal of Economic Perspectives* 21 (1), 2007: 141-167.
- «Using Randomization in Development Economics Research: A Toolkit». (Con A. Banerjee, Rachel Glennerster, y Michael Kremer.) *Handbook of Development Economics*, vol. 4., 2006.
- «Remedying Education: Evidence from Two Randomized Experiments in India». (Con A. Banerjee, S. Cole, y L. Linden.) *Quarterly Journal of Economics*, 122 (3), 2007: 1235-1264.
- «Monitoring Works: Getting Teachers to Come to School». (Con Rema Hanna.) Mayo 2006.

- «Saving Incentives for Low-and Middle-Income Families: Evidence from a Field Experiment with H&R Block». (Con William Gale, Jeffrey Liebman, Peter Orszag, y Emmanuel Saez.) *Quarterly Journal of Economics*, 121 (4), 2006: 1311-1346.
- Field Experiments in Development Economics*, enero de 2006.
- Gender Equality in Development*, diciembre de 2005.
- «Growth Theory Through the Lens of Development Economics». (Con Abhijit Banerjee.) *Handbook of Economic Growth*, vol. 1A, 2005: 473-52.
- «Dams». (Con Rohini Pande.) *Quarterly Journal of Economics*, 122 (2), 2007: 601-646.
- «Addressing Absence». (Con Abhijit Banerjee.) *Journal of Economic Perspectives*, 2005.
- «Why Political Reservations». *Journal of the European Economic Association* 3 (2-3), mayo 2005: 668-678.
- «Unappreciated Service: Performance, Perceptions, and Women Leaders in India». (Con Petia Topalova.) Octubre 2004.
- «Women as Policy Makers: Evidence from a Randomized Policy Experiment in India». (Con Raghavendra Chattopadhyay.) *Econometrica* 72 (5), septiembre 2004: 1409-1443.
- «Health Care Delivery in Rural Rajasthan». (Con Banerjee, y A. Deaton) Febrero 2004 (2-28-04). *Economic and Political Weekly*, 944-949.
- «The Medium Run Consequences of Educational Expansion Evidence from a Large Construction Program in Indonesia». *Journal of the Development Economics*, 74 (1), 2004: 163-197.
- «The Impact of Reservation in the Panchayati Raj: Evidence from a Nationwide Randomized Experiment». (Con R. Chattopadhyay.) *Economic and Political Weekly*, febrero 2004 (2-28-04): 979-986.
- «How Much Should We Trust Differences-in-Differences Estimates?». (Con Sendhil Mullainathan, y Marianne Bertrand.) *Quarterly Journal of Economics* 119 (1), 2004: 249-275.
- «Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment». *American Economic Review*, vol. 91 (4), 2001: 795-813.

4. Selección de publicaciones de Rohini Pande

- «Coordinating Development: Can Income-based Incentive Schemes Eliminate Pareto Inferior Equilibria?» (Con Philip Bond.) *Journal of Development Economics* 83 (2), julio 2007: 368-391.
- «Dams». (Con Esther Duflo.) *Quarterly Journal of Economics* 122 (2), mayo 2007: 601-646.
- «Profits and Politics: Coordinating Technology Adoption in Agriculture». *Journal of Development Economics*, vol. 81, 2, 2006.
- «Institutions and Development: A View from Below». (Con Christopher Udry.) De próxima aparición en *Proceedings of the 9th World Congress of the Econometric Society* (editado por R. Blundell, W. Newey, y T. Persson), Cambridge University Press, 2005.
- «Participatory Democracy in Action: Survey Evidence from India». (Con Tim Besley, y Vijayendra Rao.) *Journal of the European Economic Association Papers and Proceedings*, vol. 3 (2-3), abril-mayo 2005: 648-657.
- «Banking for the Poor: Evidence from India». (Con Robin Burgess, y Grace Wong.) *Journal of the European Economic Association Papers and Proceedings*, vol. 3 (2-3), abril-mayo 2005: 268-278.
- «Can Rural Banks Reduce Poverty? Evidence from the Indian Social Banking Experiment». (Con Robin Burgess.) *American Economic Review* 95 (3), junio 2005: 780-795.
- «Unmarried Parenthood and Redistributive Politics». (Con Lena Edlund, y Laila Haider.) *Journal of the European Economic Association*, vol. 3 (1), marzo 2005: 95-119.
- «The Politics of Public Good Provision: Evidence from Indian Local Governments». (Con Timothy Besley, Lupin Rahman, y Vijayendra Rao.) *Journal of the European Economic Association Papers and Proceedings*, vol. 2 (2-3), abril-mayo 2004: 416-426.
- «Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence? Theory and Evidence from India». *American Economic Review*, vol. 93 (4), septiembre 2003: 1132-1151.
- «Why Have Women Become Left-Wing? The Political Gender Gap and the Decline in Marriage». (Con Lena Edlund.) *Quarterly Journal of Economics*, vol. 117, agosto 2002: 917-961.

5. Selección de publicaciones de Vijayendra Rao

- «Governance, and the Karnataka Model of Development». (Con Ravi Kanbur, Gopal Kadekodi, y Vijayendra Rao.) *Economic and Political Weekly* 24, febrero 2007: 649-652.
- «The Political Economy of Panchayats in South India». (Con Tim Besley, Rohini Pande, y Vijayendra Rao.) *Economic and Political Weekly* 24, febrero 2007: 661-666.
- «The Economics of Dowries in India». (Con Vijayendra Rao.) En Kaushik Basu, ed., *Oxford Companion to Economics*. Oxford University Press, Delhi. De próxima aparición.
- «Symbolic Public Goods and the Coordination of Collective Action: A Comparison of Local Development in India and Indonesia». (Con Vijayendra Rao.) En Pranab Bardhan, e Isha Ray, eds., *The Contested Commons: Conversations Between Economists and Anthropologists*. Blackwell Publishers, 2006.
- «Governance in the Gullies: Democratic Responsiveness and Leadership in Delhi's Slums». (Con Saumitra Jha, Vijayendra Rao, y Michael Woolcock.) *World Development*. 2006.
- «The Social Impact of Social Funds In Jamaica: A "Participatory Econometric". (Con Vijayendra Rao, y Ana María Ibáñez.) «Analysis of Participation, Targeting, and Collective Action in Community-Driven Development». *Journal of Development Studies*, vol. 41, núm. 5, julio 2005: 788-838.
- «Participatory Democracy in Action: Survey Evidence from Rural India». (Con Tim Besley, Rohini Pande, y Vijayendra Rao.) *Journal of the European Economic Association*, 2005.
- «Community Based (and Driven) Development: A Critical Review». (Con Ghazala Mansuri, y Vijayendra Rao.) *The World Bank Research Observer*, vol. 19, núm. 1, 2004: 1-39.
- «Integrating Qualitative and Quantitative Approaches in Program Evaluation». (Con Vijayendra Rao, y Michael Woolcock.) Cap. 8 en Francois Bourguignon y Luiz A. Pereira da Silva, ed., *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*. World Bank/Oxford University Press, 2004.
- «Conclusion: Implications of Cultural Lens for Development Policy and Thought». (Con Vijayendra Rao, y Michael Walton.) En Vijayendra Rao, y Michael Walton, eds., *Culture and Public Action*. Stanford University Press, 2004.

- «Wedding Celebrations as Conspicuous Consumption: Signaling Social Status in Rural India». (Con Francis Bloch, Vijayendra Ra, y Sonal Desai.) *Journal of Human Resources*, vol. XXXIX, núm. 3, verano 2004.
- «The Politics of Public Good Provision: Evidence From Indian Local Governments». (Con Tim Besley, Rohini Pande, Lupin Rahman, y Vijayendra Rao.) *Journal of the European Economic Association*, abril-mayo 2004.
- «The Determinants of Gender Equity in India: Examining Dyson and Moore's Thesis with New Data». (Con Lupin Rahman, y Vijayendra Rao.) *Population and Development Review* 30 (2), junio 2004: 239-268.
- «Terror as Bargaining Instrument: A Case-Study of Dowry Violence in Rural India». (Con F. Bloch, y Vijayendra Rao.) *American Economic Review*, septiembre 2002.
- «Sex Workers and The Cost of Safe Sex: The Compensating Differential for Condom Use in Calcutta». (Con I. Gupta, M. Lokshin, y S. Jana.) *Journal of Development Economics*, vol. 71, Issue 2, agosto 2003: 585-603.
- «Celebrations as Social Investments: Festivals, Unit Price Variation and Social Status in Rural India». (Con Vijayendra Rao.) *Journal of Development Studies*, vol. 38, núm. 1, octubre 2001.
- «Price Heterogeneity and Real Inequality: A Case-Study of Prices and Poverty in Rural South India». (Con Vijayendra Rao.) *Review of Income and Wealth*, Series 46, #2, junio 2000.
- «The Marriage Squeeze Interpretation of Dowry Inflation: Response». (Con Vijayendra Rao.) *Journal of Political Economy*, vol. 108 (6), diciembre 2000.
- «The Rising Price of Husbands: a Hedonic Analysis of Dowry Inflation in Rural India». (Con Vijayendra Rao.) *Journal of Political Economy*, vol. 101 (4), 1993: 666-677.

Índice de cuadros

CUADRO 1.1:	Mecanismos de identificación e imputación de efectos	21
CUADRO 1.2:	Resultados básicos del proyecto STAR	27
CUADRO 1.3:	El procedimiento experimental frente a los resultados observacionales: la crítica de LaLonde	29
CUADRO 1.4:	Media del resultado de los exámenes	45
CUADRO 2.1:	Efecto de las campañas informativas sobre el comportamiento sexual de las niñas y los niños	63
CUADRO 2.2:	Efecto sobre los diferentes indicadores de servicios sanitarios	69
CUADRO 2.3:	Efecto medio por familia de indicadores	70
CUADRO 4.1:	Porcentaje de familias en cada país que usan las instituciones del sector financiero formal para ahorrar	107

Índice de gráficos

GRÁFICO 1.1:	Nivel de fraccionalización de los donantes	39
GRÁFICO 2.1:	Eficacia relativa de cada programa sobre los objetivos básicos	64
GRÁFICO 3.1:	Coste por año adicional de educación inducido	86
GRÁFICO 3.2:	Eficiencia de las actuaciones destinadas a mejorar los incentivos	95
GRÁFICO 3.3:	¿Qué saben los padres y los comités escolares sobre el nivel de aprendizaje de los alumnos?	99
GRÁFICO 4.1:	Penetración bancaria y PIB per cápita	105
GRÁFICO 4.2:	Cuentas de crédito y PIB per cápita	106
GRÁFICO 4.3:	Evolución temporal de la correlación entre desarrollo bancario inicial y nuevas oficinas en zonas rurales	118
GRÁFICO 4.4:	Evolución del número acumulado de oficinas en zonas que ya tenían oficinas inicialmente	120
GRÁFICO 4.5:	Tendencias en la relación entre desarrollo financiero inicial y apertura acumulada de oficinas	120
GRÁFICO 4.6:	Evolución del número de pobres en zonas rurales y urbanas	122

Índice alfabético

- absentismo, 16, 50, 51, 54, 55, 67, 76, 84, 85, 87, 92, 93, 94
- África subsahariana, 11, 50
- Agencia Europea de Evaluación de Medicamentos, 23
- análisis de textos, 138
- ANDERSON, M., 41
- ANGRIST, J., 40, 96
- balsakhi* (profesora de apoyo), 89, 90, 94
- Banco Africano de Desarrollo, 68
- Banco Mundial, 16, 37, 38, 68, 97, 140
- BANERJEE, A., 33, 34, 88, 98, 110, 112, 125
- Bangladesh, 50, 76, 124
- becas, 16, 43, 84, 86g, 87, 94, 95g, 95
- BECK, T., 105g, 106g
- BLEAKLEY, H., 55
- BOONE, P., 32
- Brasil, 59, 107c
- buenas prácticas (*best practices*), 80, 81
- BURGESS, R., 105, 117, 123
- BURNSIDE, C., 32, 33
- Busia, 44, 86g, 87, 94, 95
- cámaras (fotográficas), 16, 92, 93, 94, 95, 95g
- Camboya, 15, 66, 67, 71, 72, 82
- campañás
- educativas, 15, 60
 - de vacunación, 50, 59
 - informativas, 35, 63c
- CHAUDHURY, N., 76
- CHEN, S., 42
- cheques escolares, 15, 25, 77, 78, 96, 97
- Chile, 77
- China, 38, 42, 91, 92
- circuncisión, 65
- colateral, 112, 113, 124, 125, 126
- Colombia, 25, 96
- comités
- educativos, 97, 98
 - escolares, 78, 79, 97, 99g, 99
- Community Driven Development* (CDD o desarrollo basado en la comunidad), 16, 129, 130, 139, 140, 142
- Computer Assisted Learning* (CAL), 89
- contrafactual, 18, 19, 20, 28
- corrupción, 14, 34, 35, 71, 97, 98, 138, 143, 144
- coste-beneficio (o eficiencia o coste efectivo), 45, 72, 97, 104
- Current Population Survey (CPS), 29
- datos observacionales, 27, 28, 29, 30, 42, 43, 44, 45
- DEMIRGUC-KUNT, A., 105g, 106g
- descentralización, 15, 96, 97
- distribución de:
- comida, 16, 18, 54, 84, 84n, 86g
 - gafas, 91, 92
 - libros, 42, 43, 44, 45, 88, 95
 - uniformes, 42, 54, 63c, 83, 83n

- DJANKOV, S., 37, 38, 39g, 41
 DOLLAR, D., 32, 33
 DUFOLO, E., 15, 16, 42, 60, 110, 112, 125, 133n, 147
 DUPAS, P., 62
- EASTERLY, W., 32, 41
 efectos secundarios de la ayuda, 41, 42
 embarazo, 58, 61, 62, 63, 64, 65, 66, 63c
 emparejamientos (o *matching*), 21c, 30, 135
 entrevista en profundidad, 138, 144
 escuela,
 primaria, 76, 83, 96
 secundaria, 96
 escuelas,
 privadas, 77
 públicas, 77
 Estados Unidos, 14, 23, 23n, 25, 29, 39, 50,
 55, 60, 107c, 111, 113
 evaluación experimental (o método experi-
 mental), 14, 15, 17, 18, 23, 24, 31, 40, 46,
 52, 55, 129
 experimento
 natural, 109, 110, 117, 121, 127, 138
 aleatorizado, 23, 23n, 24, 27, 28, 30, 45,
 58, 92, 94, 97, 109, 111, 135
- Filipinas, 126
flip charts, 18, 18n, 42, 43, 44, 44n, 45, 82, 88
focus group (o entrevista en grupo), 137
 Food and Drug Administration (FDA), 23
 fundación Rockefeller, 55
- GERTLER, P., 58
 GIBSON, C., 36n
 GINE, X., 125
 GLENNERSTER, R., 42, 133n
 GLEWWE, P., 42, 44, 45c, 88, 91, 92
- HONOHAN, P., 108
- incentivos, 13, 15, 16, 22, 35, 36, 37, 46, 50, 51,
 52, 53, 57, 66, 70, 88, 89, 92, 93, 94, 95g, 130
 India, la, 16, 50, 51, 55, 59, 76, 77, 79, 84, 86,
 87, 88, 89, 91, 93, 97, 98, 100, 104, 107c,
 110, 112, 115, 116, 117, 121, 123, 124,
 125, 127, 132, 133, 138, 148
 Indonesia, 35, 76, 97, 130, 136, 143, 144, 145
- JENSEN, R., 94
- KARLAN, D., 110, 111, 125
 Kenia, 15, 42, 43, 44, 54, 55, 57, 60, 61, 83, 84,
 86g, 87, 88, 93, 94, 95, 97, 100, 107c
 KHANDKER, S., 124, 125
 KING, R. G., 108
 KREMER, M., 15, 54, 56, 57, 76, 77, 78, 94,
 133n, 147
- LaLONDE, R., 28, 29, 30, 29c, 45
 LEVINE, R., 32
- Madagascar, 100
 MANSURI, G., 139n
 MAREN, M., 41
 MARTÍNEZ PERIA, M., 105g, 106g
 México, 25, 53, 59, 86g
 microcréditos, 16, 103, 104, 109, 111, 113,
 114, 115, 124, 125, 126, 143
 MIGUEL, E., 54, 56, 57
 MONTALVO, J. G., 30, 35n, 37, 38, 39, 41
- Naciones Unidas, 76
 National Supported Work Demonstration
 (NSWD), 29
- OLKEN, B., 35, 97, 98
 ONG, 37, 38, 44, 46, 47, 54, 59, 66, 71, 77, 80,
 87, 95, 97, 131
 Oportunidades, 57n, 58, 59
- Pakistán, 71, 77, 107c
 PANDE, R., 16, 115, 117, 123, 126, 147
 Panel Study of Income Dynamics (PSID), 29
 parásitos/desparasitación, 13, 15, 42, 45, 53,
 54, 55, 56, 57, 60, 72, 84, 86g, 87, 100
 PARK, A., 88, 91, 92
 PAULSON, A., 108
 PITT, M., 124, 125

- Plan Director de Cooperación Española, 37
 Poverty Action Lab (Laboratorio para Acción
 contra la Pobreza), 75
 PRITCHETT, L., 40
 Programa de Ampliación de la Cobertura de
 la Educación Secundaria (PACES), 25, 96
Programa de Banca Social (o licencias «1:4»),
 116, 117, 119, 122, 123
 programa Moving to Opportunity, 25
 programa Read India, 91
 programa Sarva Shiksha Abhiyan (SSA), 97
 programas de créditos, 103, 104, 113, 115
 Progresía, 25, 53, 57, 58, 59, 83, 85, 86, 100
 proyecto STAR, 25, 27*c*, 27, 90*n*
 Public Expenditure Tracing System (PETS),
 34, 35
- RAJAN, R., 33, 34, 41
 Rajastán, 59
 RAO, V., 16, 137*n*, 139*n*, 147
 RAVALLION, M., 42, 134*n*
Regression Discontinuity Design (RDD o diseño
 de discontinuidad), 21*c*, 30
 REINIKKA, R., 34, 35, 98
 República Dominicana, 94
 restricciones crediticias, 103, 108, 109, 110
 REYNAL-QUEROL, M., 30, 37, 38, 39, 41
 ROODMAN, D., 32
 RUBIN, D., 30
 Rusia, 142
- scaling up*, 55
 Schultz, P., 58
 servicios públicos, 79, 96, 140
- sida, 15, 50, 53, 60, 61, 62, 63, 65, 83, 132
 Sierra Leona, 71, 142
 Somalia, 12, 41
 Subramanian, A., 33, 34
 subsidios (o transferencias monetarias, con-
 dicionadas o subvenciones), 15, 16, 42,
 45, 57, 59, 72, 94, 114, 123, 124, 143
 Sudáfrica, 65, 107*c*, 110, 111, 113
sugar daddy, 64*g*, 64
 suplementos nutricionales, 58
 SVENSSON, J., 35
- Tailandia, 108
 Tanzania, 35, 107*c*
 teoría del paracaídas, 40, 41, 42
 Teso, 44, 86*g*, 87, 94, 95
 THORNOTN, R., 94
 TOWNSEND, R. M., 108
- Uganda, 35, 50, 76, 98, 100
- vacunación, 15, 45, 50, 53, 59, 60, 66, 69*c*, 71, 72
 validez,
 externa, 31, 130, 131
 interna, 31, 130, 131
 variables instrumentales, 21*c*, 23, 69*c*, 135,
 136
 Vietnam, 50, 67, 107*c*
 votación con los pies, 77
- WOOLCOCK, M., 137*n*
- ZHAO, M., 88, 91, 92
 ZINMAN, J., 110, 111

Nota sobre los autores

ESTHER DUFLO es doctora en Economía por el Massachusetts Institute of Technology (MIT), catedrática en Economía (Professor of Poverty Alleviation and Development Economics) en dicha universidad y directora del Poverty Lab in Action. Su investigación se centra en la aplicación de procedimientos experimentales al análisis de programas contra la pobreza en países en vías de desarrollo. Es una gran experta en temas relacionados con el desarrollo en la India. Ha escrito multitud de artículos sobre desigualdad, el papel de la mujer en los países en vías de desarrollo y el efecto de la construcción de escuelas en los países más pobres.

MICHAEL KREMER es doctor en Economía por la Universidad de Harvard, donde es catedrático de Economía (Gates Professor of Developing Countries), y fue fundador de la ONG World Teach. Ha trabajado extensamente en la financiación de las vacunas para el tercer mundo, la evaluación científica de programas de desarrollo económico, las externalidades de los programas de eliminación de parásitos de los niños y la ausencia de personal sanitario y de profesores en países subdesarrollados, la evaluación de las políticas educativas en países en vías de desarrollo y la influencia del comportamiento humano en la expansión del sida. Sus trabajos recientes también inclu-

yen el análisis de la deuda odiosa, el escaso uso de fertilizantes en muchas zonas subdesarrolladas o el impacto de la construcción de fuentes en salidas de aguas subterráneas.

JOSÉ GARCÍA MONTALVO, doctor en Economía por la Universidad de Harvard, es catedrático de Economía de la Universidad Pompeu Fabra (UPF), donde actualmente tiene el cargo de director del Departamento de Economía y Empresa. Ha trabajado como consultor para el Banco Mundial, la Organización para la Cooperación y el Desarrollo Económico (OCDE) y el Banco Interamericano de Desarrollo. Sus últimos trabajos en el campo del desarrollo económico se centran en el efecto sobre la malaria de los desplazamientos masivos de población causados por las guerras civiles en África, las causas de los conflictos armados y sus consecuencias sobre el desarrollo, la provisión de servicios sanitarios en Mozambique así como la pobreza en China.

ROHINI PANDE, doctora por la London School of Economics, es profesora de Política Pública (Mohammed Kamal Professor of Public Policy) en la Kennedy School of Government de la Universidad de Harvard. Su investigación se centra en el desarrollo económico, la economía política y las diferencias de género. Sus trabajos más influyentes tienen relación con los efectos de las intervenciones financieras, en particular los bancos rurales, sobre el desarrollo económico.

VIJAYENDRA RAO, doctor en Economía por la Universidad de Pensilvania, es Lead Economist del Departamento de Investigación del Banco Mundial en la división de pobreza y desigualdad. Ha escrito multitud de trabajos sobre la pobreza en la India y es el investigador de referencia en los temas de desarrollo basado en la comunidad y métodos mixtos de evaluación, así como una autoridad en los fundamentos económicos de la institución de la dote en la India.

Fundación **BBVA**

Gran Vía, 12
48001 Bilbao
España
Tel.: +34 94 487 52 52
Fax: +34 94 424 46 21

Paseo de Recoletos, 10
28001 Madrid
España
Tel.: +34 91 374 54 00
Fax: +34 91 374 85 22
publicaciones@fbbva.es
www.fbbva.es

ISBN 978-84-96515-79-6



9 788496 515796