



RE-308

***Evaluación Ex Post de
Proyectos
Informe Anual de 2004***

**Oficina de Evaluación y Supervisión, OVE
Banco Interamericano de Desarrollo
Washington, D.C.
10 de Agosto de 2005**

ÍNDICE

INTRODUCCIÓN.....	I
I. POLÍTICA DE EVALUACIÓN EX POST	1
II. MEDIDAS ADOPTADAS PARA EJECUTAR LA POLÍTICA SOBRE EVALUACIÓN EX POST	6
III. CONCLUSIONES DE LA EVALUACIÓN DE PROYECTOS.....	15
A. Revisión por temas: Proyectos sobre Regularización de la tenencia de la tierra ...	15
B. Revisión temática: Proyectos de mejoramiento de barrios.....	17
C. Evaluaciones aisladas.....	19
D. Lecciones aprendidas	21
E. Situación actual y costo.....	22
IV. LECCIONES APRENDIDAS EN 2004	25
V. CONCLUSIONES Y RECOMENDACIONES	32

ANEXOS DISPONIBLES EN: <http://idbdocs.iadb.org/Wsdocs/Getdocument.aspx?docnum=582891>

ANEXO 1: RESUMEN DE LAS CONCLUSIONES DE LA EVALUACIÓN DE PROYECTOS

ANEXO 2: NORMAS METODOLÓGICAS PARA LAS EVALUACIONES EX POST

ANEXO 3: EVALUACIONES EX POST PROPUESTAS PARA EL CICLO DE 2005

INTRODUCCIÓN

El Directorio Ejecutivo del Banco aprobó, en octubre de 2003, la nueva Política de evaluación ex post de Operaciones, en la cual se le encomendaron dos nuevas tareas a la Oficina de Evaluación y Supervisión (OVE): examinar y validar los informes de terminación del proyecto y realizar evaluaciones ex post de las operaciones¹. El presente informe se limita a describir la tarea de evaluación ex post². Su finalidad es presentar las conclusiones preliminares resultantes de la experiencia del primer año de aplicación de la Política aprobada. Estas conclusiones preliminares se concentran en los tres temas siguientes: i) proyectos individuales evaluados; ii) aspectos metodológicos del sistema de seguimiento y evaluación del Banco, y iii) estándares de los métodos de evaluación.

Cabe recordar que la Política de evaluación ex post surgió como respuesta a la percepción de que se había creado un vacío de evaluación en la efectividad en el desarrollo cuando el Banco abandonó la política de evaluaciones ex post obligatorias en 1994, y se cerró la Oficina de Evaluación de Operaciones³. El vacío percibido se ha vuelto mas relevante como resultado de numerosas iniciativas recientes del BID (véase el Recuadro 1) para mejorar la efectividad en el desarrollo de sus operaciones, que en conjunto evidencian su compromiso de aplicar un modelo gerencial de “gestión por resultados”. Esta reciente serie de acciones forma parte de la respuesta del BID a la solicitud formulada en 2002 por los Gobernadores del Banco en el sentido de que *todas las operaciones del Banco tengan objetivos claros que hagan factible definir parámetros de referencia para evaluar la efectividad de nuestra institución ... y que se aparte de una cultura basada en aprobaciones [para adoptar...] un enfoque basado en resultados*⁴.

¹ La nueva *Política sobre evaluación ex-post de las operaciones*, versión final, figura en el documento GN-2254-5 y en el *Documento de antecedentes sobre política de evaluación ex post de las operaciones*, GN-2254-6, de septiembre de 2003. *Informe del Presidente del Comité de Políticas y Evaluación del Directorio Ejecutivo*, documentos GN-2254-7, del 8 de octubre de 2003, y OP-305.

² Obsérvese que entre los proyectos evaluados se encuentran los “estudios piloto” propuestos en el documento RE-257, de septiembre de 2003.

³ Antes de 1994, todos los proyectos del BID debían ser evaluados por el prestatario (*Modificaciones al sistema de evaluación ex post por el prestatario*, marzo de 1993, CON/OEO-84/93). La reforma cambió el sistema obligatorio por uno voluntario. Simultáneamente, se cerró la Oficina de Evaluación de Operaciones, que tenía a su cargo las autoevaluaciones.

⁴ Palabras del señor Enrique V. Iglesias en la Sesión de Clausura, Reunión Anual de la Asamblea de Gobernadores, Fortaleza, AB-2281, 26 de marzo de 2003.

Recuadro 1: Iniciativas recientes para elevar la efectividad en el desarrollo

Las *iniciativas* incluyen las siguientes:

- Aprobación de una nueva Política de evaluación ex post.
- Creación de un sistema destinado a mejorar la calidad inicial de los proyectos, y definición de un itinerario para la agenda de la efectividad en el desarrollo, consagrada en un Plan de Acción a Mediano Plazo (MTAP).
- Reformulación de dos de los principales informes de seguimiento y evaluación del Banco¹ con miras a mejorar la información sobre los objetivos de desarrollo; readecuar la información sobre numerosos proyectos considerada insuficiente, y mejorar la información sobre resultados en el *Informe Anual sobre gestión, desempeño y resultados de la cartera* del Banco.
- Adopción de una nueva *Política sobre disponibilidad de información* que permite la divulgación pública de todos los documentos de evaluación presentados al Directorio Ejecutivo del Banco.
- Designación de un funcionario en la oficina central de cada Departamento Regional con la responsabilidad principal de verificar que las iniciativas se hagan realidad en el terreno, y creación de una Oficina de Efectividad en el Desarrollo, más tarde consolidada en un nuevo Departamento de Efectividad en el Desarrollo y Planificación Estratégica (DEV).

La descripción de estas iniciativas figura en los documentos *Plan de Acción a Mediano Plazo para la efectividad en el desarrollo en el BID*, GN-2324, del 2 de agosto de 2004; *Informe de avance sobre medidas adoptadas por la Administración en 2003 y futuras acciones destinadas a reforzar la efectividad del Banco en el desarrollo*, CA-456, del 28 de marzo de 2004; *Informe del Presidente del Comité de Políticas y Evaluación y Programa de Ejecución del Pilar Externo del Plan de Acción a Mediano Plazo para la Efectividad en el Desarrollo (PRODEV)*, GN-2346-2, de marzo de 2005. La ampliación del servicio crediticio PRODEF, de modo que incluya financiación para evaluaciones ex post figura en el documento GN-2351, de diciembre de 2004.

Estas iniciativas también reflejan que el Banco se encuentra comprometido a fomentar una mayor responsabilidad institucional, como surge de la declaración conjunta denominada *Medición, seguimiento y gestión por resultados en función del desarrollo* de 2002. Implica asimismo una mayor demanda de evaluación para hacer frente a los desafíos de un nuevo contexto externo (de mayor exigencia en la rendición de cuentas por parte de las partes interesadas) e interno (de reorientación hacia una gestión basada en los resultados). La gestión por resultados implica la acumulación de conocimientos en los que pueda apoyarse un proceso decisorio basado en pruebas para el diseño de políticas, estrategias y proyectos.

El presente informe está organizado de la siguiente manera. El capítulo I contiene un resumen de los elementos de diseño de la Política de evaluación ex post. El capítulo II presenta una relación de las medidas adoptadas por OVE para poner en práctica dicha política, incluida una descripción de las normas evaluativas. El capítulo III contiene un resumen de las principales conclusiones de una selección de proyectos evaluados durante el ciclo de evaluación ex post de 2004⁵. El capítulo IV utiliza los estudios de caso de las evaluaciones de los proyectos, con la adición de nuevos estudios sobre los aspectos metodológicos del sistema de seguimiento y evaluación del Banco y presenta las lecciones aprendidas durante el ciclo de 2004. El último capítulo presenta las conclusiones y recomendaciones.

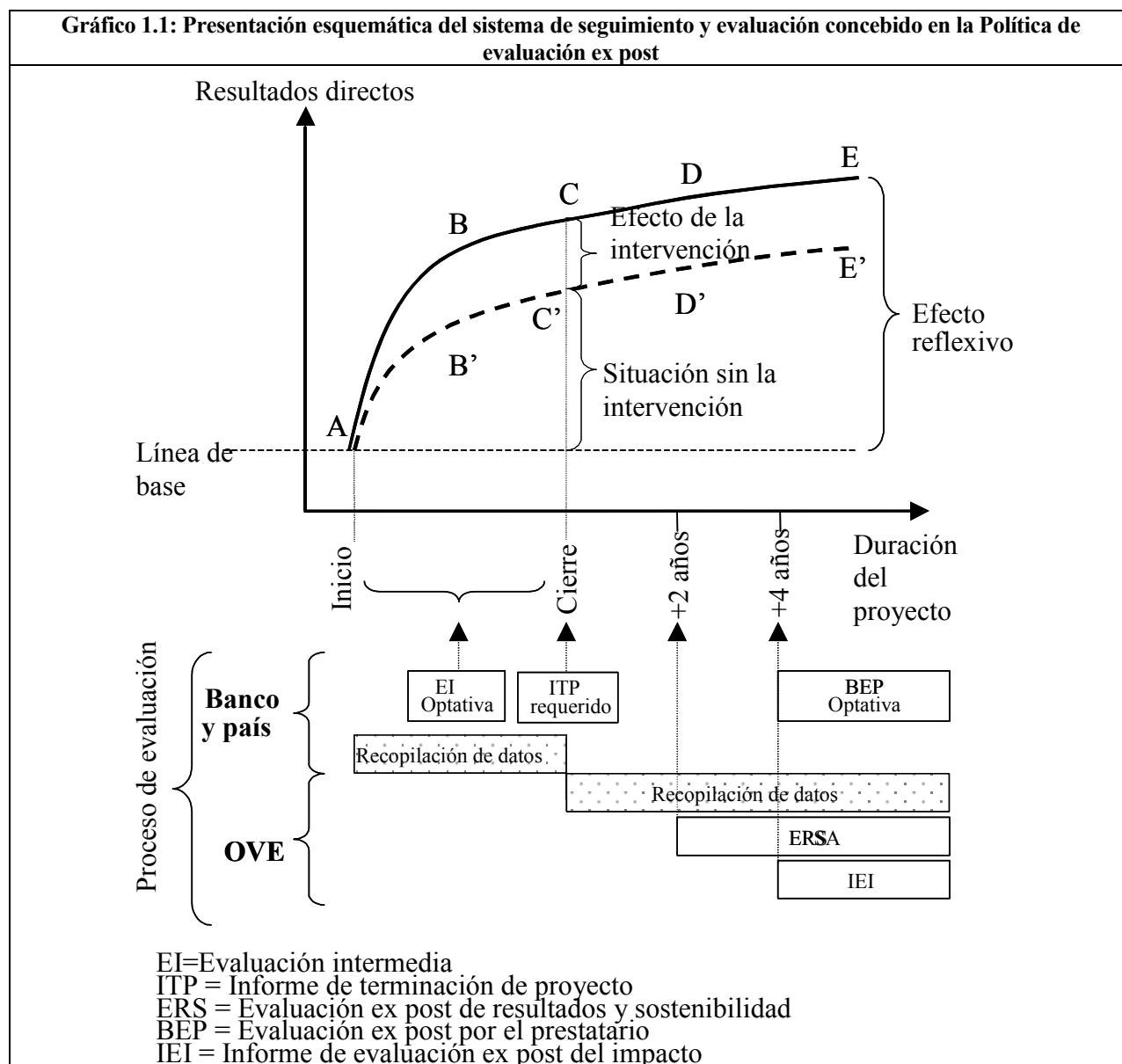
⁵ La Política sobre evaluación ex post también menciona los productos del sistema de Evaluación Ex Post por el Prestatario, aun cuando se espera recibir el informe de este ejercicio después de cerrado el proyecto; se supone que durante la ejecución del proyecto se estén recolectando datos para dicho informe.

I. POLÍTICA DE EVALUACIÓN EX POST

- 1.1 Para poner las principales conclusiones del informe en su debido contexto, conviene destacar los parámetros que fijaron la Política de evaluación ex post (documento GN-2254-5) y otras orientaciones aprobadas por el Directorio (documento RE-284) en 2004 para este tipo de evaluación. La información que había disponible en 2003 y que se utilizó en el diseño figura en el documento de antecedentes de la Política.
- 1.2 La Política de evaluación ex post procura asignar la responsabilidad de la evaluación a diferentes protagonistas. A la Administración le asigna la tarea de desarrollar y mantener los indicadores de productos y resultados directos durante la ejecución de los proyectos, mientras que a OVE le asigna la evaluación posterior a la ejecución (párrafo 1.2). Además, *“los prestatarios tienen la responsabilidad principal de recopilar la información básica necesaria para la preparación del ITP y las evaluaciones ex post. Tal información se reúne sistemáticamente para todos los proyectos hasta el nivel de resultados”* (párrafo 1.6). *“...Como parte integrante de la ejecución de todos los proyectos del Banco, el organismo ejecutor debe recopilar la información y los indicadores para las actividades de seguimiento y evaluación acordadas durante el diseño del proyecto* (párrafo 1.6). *OVE realizará, a modo de muestra, minuciosas evaluaciones ex post de los resultados (impactos y/o resultados) de las operaciones financiadas por el BID a dos años o más de su conclusión”* (párrafo 1.3). Además, la Política estipula que durante la ejecución de los proyectos *“Las actividades de seguimiento y evaluación en el BID son una responsabilidad compartida entre el Banco y el país prestatario”* (párrafo 1.3).
- 1.3 La política menciona cinco informes (véanse los párrafos 4.1 a 4.5) que han de emanar del sistema y que se describen en el Recuadro 1.1, a saber: i) Informe de terminación del proyecto (ITP); ii) Revisión del ITP; iii) Evaluación ex post de resultados y sostenibilidad (ERS); iv) Informes de evaluaciones de impacto (IEI); y v) la Evaluación Ex Post por parte del Prestatario.

Recuadro 1.1: Informes requeridos por la Política de evaluación ex post			
Tipo de informe	Número	Definición y cuestiones evaluativas	Responsabilidad
Informe de terminación del proyecto (ITP)	100% de operaciones terminadas	Resumen de la experiencia de la ejecución, información sobre productos y resultados directos, calificación del desempeño	Administración
Revisión de los ITP	100% de operaciones terminadas	Validación de los ITP	OVE
Evaluación ex post de resultados y sostenibilidad	20% de los proyectos concluidos por año, dos años o más después de concluido el proyecto	Evaluación de los resultados a corto y mediano plazos de un programa, incluidos sus productos y resultados directos, la sostenibilidad institucional y el flujo de productos.	OVE
Informes de evaluaciones de impacto	Al menos dos al año, cuatro años o más después de concluido el proyecto	Evaluación de los resultados a largo plazo, incluido su impacto y resultados directos, así como su efectividad, eficiencia y flujo de resultados beneficiosos.	OVE
Evaluación Ex Post por el Prestatario	Voluntaria	Evaluación de los resultados y sostenibilidad.	País prestatario

- 1.4 El sistema descrito en la Política de evaluación ex post puede representarse en forma esquemática como se hace en el Gráfico 1.1. Los proyectos que se encuentran en etapa de aprobación están diseñados con claros indicadores para sus objetivos de desarrollo y resultados directos (los “resultados directos” de la coordenada vertical). Cerca de la fecha de iniciación del proyecto, se recopila información de base sobre esos resultados directos, es decir, el estado de los objetivos de desarrollo antes del proyecto (valor “A”). Durante la ejecución del proyecto puede recopilarse más información sobre los indicadores de objetivos de desarrollo, valor “B”, para una evaluación intermedia. Hacia el final del proyecto, se recopilan datos sobre los objetivos de desarrollo y resultados directos (valor “C”). La información identificada como A, B y C constituye insumos para las afirmaciones que se hacen en los informes de terminación del proyecto sobre los objetivos de desarrollo, en los cuales se informa, y quizás se interpreta, que de todas las variaciones ocurridas en dichos indicadores son automáticamente atribuibles a la intervención.



- 1.5 La Política sobre evaluación ex post le encomienda a OVE la preparación de evaluaciones del “impacto” después de concluido el proyecto. Evaluar el impacto significa analizar minuciosamente los resultados logrados a fin de separar los efectos derivados del tratamiento (la operación) y los que se deban a otras fuerzas que influyan sobre las mismas situaciones durante el mismo periodo. Para evaluar estos efectos del tratamiento se requiere recolectar datos adicionales durante la ejecución del proyecto respecto a los no beneficiarios, a fin de poder estimar los valores de B’ y C’. Provista de esta información (A, B, B’, C y C’), OVE procederá a recopilar datos en el punto “D” (y estimar D’) para una *Evaluación ex post de resultados y sostenibilidad* (ERS), y datos en el punto E (y E’) para un *Informe de evaluación de impacto* (IEI).

- 1.6 Según esta interpretación, es oportuno establecer que la principal distinción entre la Evaluación ex post de resultados y sostenibilidad y el Informe de evaluación de impacto es el tiempo. La Política indica que la primera se realiza al menos dos años después de la conclusión del proyecto y, el segundo, al menos cuatro años después de dicha conclusión. Si bien esta distinción puede deberse a la intención de medir los efectos a corto y largo plazos, respectivamente, en consonancia con los principios evaluativos del Comité para el Desarrollo,¹ la heterogeneidad de los proyectos del Banco hace que sea improbable que todos los proyectos tengan efectos de corto y largo plazo en mismo periodo de tiempo. Además, es posible que la pertinencia de la cláusula sobre aprendizaje de lecciones y adaptación de políticas y programas con el fin de mejorar la efectividad en el desarrollo necesite una revisión adicional. Por ejemplo, la lógica del modelo indica que el efecto de un tratamiento o efecto de una intervención podría ser visible durante la ejecución del proyecto, particularmente para operaciones que apliquen tratamientos similares para diferentes grupos de población a medida que madura el proyecto. En tales casos, podría haber un valor de aprendizaje en la evaluación de los efectos del tratamiento durante la ejecución, así como también después del último desembolso.
- 1.7 El uso frecuente de la forma condicional en la redacción del documento de la Política indica que ésta también podría interpretarse como una afirmación prospectiva y normativa: no lo que es, sino lo que deba ser. Según esta interpretación, la Política de evaluación ex post reconoce implícitamente que si bien el sistema puede no ser válido para la actual serie de proyectos terminados, debería serlo para los futuros proyectos. Según esta óptica, tiene sentido la afirmación: “...OVE procurará identificar operaciones que estarán sujetas a este tipo de evaluación (del impacto) lo más pronto posible durante el ciclo de proyectos (v.g., durante la programación y el diseño de las operaciones) (párrafo 3.2). Esto implica la creación de un mecanismo de coordinación entre la Administración y OVE que permita comprobar que la información necesaria para la evaluación ex post sea recopilada en forma coincidente con la vigencia del proyecto. No obstante, la Política dispone que se realice la evaluación de todo el actual conjunto de proyectos financiados por el BID.
- 1.8 La Política define un sistema según el cual todos los protagonistas deben desempeñarse según lo previsto. Si alguno de ellos deja de cumplir la función asignada, las consecuencias posteriores serán negativas y afectarán al resto de actores. Si un proyecto los objetivos de desarrollo carecen de una métrica definida en la etapa de aprobación, o si el Banco no genera, recopila y mantiene efectivamente la información evaluativa prometida durante la ejecución del proyecto para el uso de terceros, resultará difícil o imposible realizar el ejercicio ex post tal como está concebido en la Política. Así sucederá para los proyectos aprobados y concluidos antes de la aprobación de la Política. En tales circunstancias, todo el ejercicio de evaluación se traslada a OVE, oficina que tiene que definir la métrica de los resultados directos y buscar información sobre ellos en forma retrospectiva: Desde C a B y hasta A, y estimar D y E. Teniendo en cuenta que el periodo promedio de ejecución de un proyecto es de aproximadamente 10 años, la tarea evaluativa tiende a parecerse más a una exploración arqueológica: la de buscar e indagar para obtener información.

¹ Véase el documento: *Principles for Evaluation of Development Assistance*; “Review of DAC Principles For Evaluation of Development Assistance”, OCDE, 1998; y el Glosario de los principales términos sobre evaluación y gestión basada en resultados” en el sitio virtual <http://www.oecd.org/document/>

- 1.9 La Política de evaluación ex post dejó los detalles de cómo implementar la labor ex post para que se definan en la práctica. El *Documento de antecedentes sobre la Política* (GN-2254-6) propone que: "... el Banco (Administración y OVE) emitirá directrices operativas y metodológicas a fin de asegurar que en todo el proceso se apliquen normas homogéneas. Las directrices metodológicas incluirán una clara presentación de las distintas metodologías disponibles, sus requisitos de información, la capacidad institucional necesaria para su exitosa aplicación, y una estimación del rango de costos. Una definición importante será definir el estándar de evaluación mínimo ha de considerarse aceptable como evaluación ex post para el Banco ..." (Párrafo 4.8). El presente informe presenta los insumos preliminares para dichas directrices². En el capítulo II se describen los métodos de evaluación en general y la estándar metodológico mínimo en particular (cuyos requisitos de datos se detallan en el Anexo 2). La información sobre costos figura en el capítulo III.
- 1.10 La Política de evaluación ex post y su documento de antecedentes fueron aprobados junto con un plan de puesta en marcha de seis evaluaciones ex post con carácter experimental (documentos RE-284, RE-284-1, RE-284-2), con el propósito de "ofrecer orientación en cuestiones operativas clave que afectarán a la exitosa implantación de la nueva Política de evaluación ex post en el Banco."³ Las preguntas cruciales propuestas en el documento RE-284 incluyeron las siguientes: i) ¿Cómo se puede coordinar de manera óptima la ejecución de las evaluaciones ex post a fin de maximizar su valor agregado?; ii) ¿Cuáles son las implicaciones presupuestarias y de datos de los distintos criterios metodológicos para la evaluación ex post?; iii) ¿Cuáles son los aportes relativos de las evaluaciones reflexivas y de impacto para determinar la eficacia de los proyectos?; y iv) ¿Cómo deberían difundirse las lecciones aprendidas al interior del Banco y en los países prestatarios? Entre sus comentarios a la propuesta aprobada (documento RE-284-1), la Administración observó lo difícil que resulta separar analíticamente la *Evaluación ex post de resultados y sostenibilidad* y los *Informes de evaluación ex post del impacto*, y añadió que debería darse atención a las preguntas evaluativas planteadas en la Política durante el ejercicio piloto. Las seis evaluaciones ex post experimentales fueron financiadas por OVE, RE1, RE3 y DPP.
- 1.11 Los seis proyectos piloto aprobados en octubre 2003 fueron incorporados al programa de trabajo de OVE para 2004, junto con un plan de trabajo para la conducción de 10 evaluaciones ex post adicionales (definidas como IEI o ERS), según lo estipulado en la Política. Dado que las preguntas evaluativas, operativas y metodológicas que deben abordarse en ambas series de evaluaciones ex post son similares y que 2004 fue el primer año de ejecución de la Política, puede considerarse que las 16 evaluaciones forman parte

² "Preliminares" porque OVE considera que el ejercicio de un año más generará una base más sólida de información para esas directrices.

³ En sus comentarios, que figuran en el documento RE-284-2, el Comité de Políticas y Evaluación del Directorio Ejecutivo coincidió con que el programa piloto "se usaría para refinar la metodología usada por OVE para llevar a cabo futuras evaluaciones del impacto." Además, "los Directores se pronunciaron interesados en conocer las recomendaciones de OVE sobre "cómo asegurar que el Banco obtenga un valor óptimo con los recursos usados en tales evaluaciones, cómo asegurar que se difundan eficazmente las lecciones aprendidas a nivel de proyecto, y cómo hacer el mejor uso posible de estos ejercicios para coordinarlos con los propios procedimientos de evaluación de los prestatarios y cómo reforzarlos."

del ejercicio experimental y que las conclusiones presentadas en este informe se refieren a todo el conjunto de evaluaciones ex post realizadas en 2004, lo que permite sentar una plataforma más sólida para el análisis y las recomendaciones que se presenten. En el capítulo II se detallan las medidas adoptadas para ejecutar las evaluaciones ex post en 2004.

II. MEDIDAS ADOPTADAS PARA EJECUTAR LA POLÍTICA SOBRE EVALUACIÓN EX POST

- 2.1 Para cumplir con el nuevo mandato, OVE realizó numerosas acciones para implantar la Política para el ciclo de 2004, actividades se documentan en este capítulo. Estas actividades deben considerarse dentro de la tarea global de establecer un sistema que permita preparar un gran número de evaluaciones de proyectos cada año. El trabajo de construir dicho sistema se encuentra en marcha.
- 2.2 Las acciones realizadas en 2004 fueron, entre otras, las siguientes: i) determinar el alcance de las preguntas evaluativas; ii) jerarquizar los métodos evaluativos; iii) definir los criterios de selección de muestras de los proyectos a evaluarse; y iv) realizar las evaluaciones ex post de 16 proyectos. En este capítulo se señalan los proyectos seleccionados para el ciclo de 2004, y se ofrece una descripción de los mecanismos de implementación adoptados en dicho ciclo, así como algunas medidas adoptadas en preparación para el ejercicio de 2005.
- 2.3 **Preguntas evaluativas.** Las preguntas evaluativas generales que propone la Política de evaluación ex post son, primero, *“la medida en la que se han alcanzado los objetivos de desarrollo de las operaciones financiadas por el BID”* y segundo, *“...la eficiencia con la cual se han conseguido tales objetivos”* (párrafo 1.1)⁴. El alcance de las preguntas evaluativas debe definirse de una manera que aporte valor agregado a la información ya generada por el Banco. En la forma originalmente diseñada, este sistema continúa siendo uno que genera información sobre ejecución y sobre los aspectos operativos y de procesamiento de sus operaciones; es decir, la información necesaria para elevar la eficiencia de la ejecución de un proyecto y mejorar los datos para la rendición de cuentas a las partes interesadas. Recientemente, se ha establecido el requisito de dar seguimiento a los datos sobre resultados, con el supuesto de que ello basta para afirmar algo respecto de la efectividad de la intervención en función del desarrollo. Pero esto puede no ser suficiente para cumplir la tarea. Aun cuando el seguimiento de los datos de resultados es importante para los gerentes de los programas y da lugar a información sobre la evolución de un indicador dado a través del tiempo, carece de la información necesaria para determinar el impacto del proyecto en ese resultado así como tampoco ofrece explicaciones en cuanto al por qué de los cambios en el indicador de resultados. Este es el vacío que llena una

⁴ Las evaluaciones del ciclo de 2004 no incluyeron análisis de eficiencia de costos, pero el de 2005 lo hará. Se ha procurado efectuar un análisis de costo y beneficios de dos proyectos en preparación para el ciclo de 2005. En ambos casos, hubo un ejercicio ex ante; en un caso hubo disponible tanto un ejercicio ex ante como una evaluación ex post. Estos ejercicios revelan que si no se dispusiera de una referencia metodológica ni de recolección de datos de seguimiento, habría dificultades para realizar una evaluación ex post de características similares.

evaluación a profundidad, de tratamiento o de impacto y es el enfoque de las evaluaciones ex post emprendidas por OVE en 2004.

- 2.4 **Método.** Se escogió la metodología con el fin de maximizar la información adicional que OVE podría proporcionar. A fin de dar respuesta a las preguntas evaluativas planteadas, debe seleccionarse un método que admita afirmaciones cuantitativas, con cierto grado de confianza, sobre la evolución de los resultados u objetivos de desarrollo (tanto los declarados explícitamente como previstos, y también los no anticipados) que sean atribuibles al proyecto. Teniendo en cuenta las principales preguntas operativas planteadas en el documento RE-284, OVE realizó sus evaluaciones ex post conforme a dos métodos: i) el método simple o reflexivo, que supone que todas las variaciones registradas en los objetivos de desarrollo, son atribuibles a la intervención; y ii) el método de efecto tratamiento, que intenta diferenciar los efectos que se deben a la intervención, de los efectos brutos o totales, estableciendo una causalidad entre la intervención y los resultados en términos cuantitativos⁵. En este punto es útil revisar nuevamente la Figura 1.1. La magnitud del efecto simple o reflexivo está representado como la distancia entre la línea de base y la línea continua. La magnitud del efecto tratamiento está representada como la distancia entre la línea punteada y la línea continua. La diferencia entre estas dos magnitudes – la distancia entre la línea de base y la línea punteada – mide el efecto de los factores externos al programa que también tuvieron incidencia en los resultados. Esto se conoce como “contra-factual” y representa lo que hubiera sucedido en ausencia del programa. Así mismo es útil, para clarificar, mencionar en la literatura de evaluación se usan diversos términos para denominar los efectos reflexivos y de tratamiento, y sus evaluaciones correspondientes. En particular, las evaluaciones que no pretenden desagregar los efectos de tratamiento y los contra-factuales son llamadas en la literatura como evaluaciones reflexivas o simples y los efectos que éstas miden se denominan simples, reflexivos o efectos brutos. Las evaluaciones que miden efectos del tratamiento son conocidas como evaluaciones de impacto; los efectos de tratamiento también son conocidos como efectos netos. En cualquier caso, en la medida en que la estimación de los efectos reflexivo y de tratamiento de un proyecto difieran, solamente debe usarse la última para calcular la eficiencia de la intervención en función del costo. En el ciclo de 2004, se adoptó como estándar mínimo el método de los efectos del tratamiento, pero con la tarea adicional de comparar y contrastar las estimaciones de los efectos brutos (simples, reflexivos) con las de los efectos netos (del tratamiento o de impacto).

5 La norma metodológica de los efectos del tratamiento depende críticamente de la posibilidad de establecer situaciones hipotéticas sin la intervención – “contra-factual”. El estándar más elevado para obtener los efectos netos proviene de una estrategia de muestreo que utiliza una selección aleatoria. Realizando una asignación aleatoria de beneficiarios elegibles en grupos de tratamiento y de control y captándolos de la misma distribución poblacional, se logra que ambos grupos sean perfectamente comparables. Esto no ha de implicar necesariamente la denegación de beneficios a los participantes previstos. Basta con asignar los beneficiarios en forma aleatoria en una primera y segunda rondas de aplicación. De no realizarse la asignación aleatoria, hay dos consideraciones que surgen con respecto a las estrategias de muestreo. Una es cómo incluir a sujetos no participantes comparables en un conjunto de datos cuando el evaluador no tiene forma de saber a priori exactamente quiénes son comparables con los beneficiarios previstos. La otra es que deben tomarse en cuenta otras variables, tanto para explicar la participación en el programa como para considerar otras influencias o resultados directos anticipados o fortuitos en función del desarrollo.

- 2.5 El objetivo de este criterio fue ilustrar las diferencias que surgen entre los cálculos de los efectos reflexivo y de tratamiento, en respuesta a la pregunta de la Administración sobre este tema planteada en el documento RE-284-1⁶. Una evaluación del efecto de tratamiento (ya sea que coincida con la duración del proyecto o que se realice con posterioridad) difiere con respecto a la evaluación reflexiva en tres aspectos cruciales: su objetivo primordial (preguntas evaluativas que puedan responderse usando el método), qué y cómo se analiza, y la intensidad y el alcance de la recopilación de datos (véase el Recuadro 2.1). No son mutuamente excluyentes ni sustituibles, pero se necesitan las dos para tener una idea completa de los resultados de un proyecto.

Recuadro 2.1: Seguimiento del desempeño frente a la evaluación de los efectos del tratamiento	
Evaluación reflexiva	Evaluación de efecto tratamiento
Objetivo primordial: rendir cuentas a las partes interesadas y resolver problemas de ejecución; costo-eficiencia	Objetivo primordial: Crear conocimiento (comprender y perfeccionar los efectos tratamiento de un programa), costo-eficiencia y costo-beneficio.
Análisis de productos y resultados directos brutos para mejorar la implementación	Análisis de los efectos netos (efectos tratamiento) de los objetivos de desarrollo para perfeccionar el diseño de proyectos (coincidentes o para futuros proyectos similares)
La recopilación de datos es constante, se cuenta con datos recolectados regularmente y de fácil acceso.	La recopilación de datos es periódica, más intensa y requiere información tanto sobre beneficiarios como no beneficiarios a través del tiempo.

- 2.6 Cabe observar que el método efecto tratamiento no es nuevo para el Banco. La norma adoptada por OVE está basada en evaluaciones de rigurosa calidad de los proyectos del BID, realizadas en forma conjunta por profesionales del Banco y autoridades del país (véase el Recuadro 2.2). El problema es no solamente que su precariedad de cifras indica que estas evaluaciones constituyen una excepción a la práctica normal, sino también que se trata exclusivamente de evaluaciones de proyectos sociales. Debido a la falta de enlace entre estas evaluaciones y los sistemas regulares de información del Banco, también parece haber una insuficiente difusión de sus conclusiones, las cuales no figuran en los sitios virtuales del Banco ni entre los documentos accesibles al público.

Recuadro 2.2: Ejemplos de autoevaluaciones de calidad de los proyectos del BID
<ul style="list-style-type: none"> • Oportunidades/ programa de transferencia condicional de efectivo / México/ evaluación de diseño cuasiexperimental. (Véase http://www.iadb.org/ove/DefaultNoCache.aspx?Action=WUCPublications@ImpactEvaluations) • PRAF/ Programa de transferencia condicional de efectivo / Honduras/ diseño experimental. (Véase http://www.iadb.org/ove/DefaultNoCache.aspx?Action=WUCPublications@ImpactEvaluations) • Red social/ Programa de transferencia condicional de efectivo / Nicaragua/ diseño experimental. (Véase http://www.iadb.org/ove/DefaultNoCache.aspx?Action=WUCPublications@ImpactEvaluations) • Red social / Colombia/ evaluación cuasiexperimental. Las evaluaciones incluyen los siguientes subprogramas: i) Familias en acción (transferencia condicional de efectivo); ii) / Empleo en servicios públicos para beneficiarios de asistencia social; y iii) Jóvenes en acción (capacitación para el trabajo)

⁶ “Es menester aportar una definición más precisa de lo que constituye una evaluación ‘reflexiva’ ... No se desprende claramente de la propuesta por qué las evaluaciones ‘reflexivas’ y de impacto efectuadas una vez ‘concluido el hecho’ contribuirían de modo diferente a la efectividad de una operación.”

- 2.7 Considerando las preguntas evaluativas planteadas, el método del efecto tratamiento es capaz de estimar el impacto del tratamiento de un programa existente y es el más apropiado para la aplicación de la Política sobre evaluación ex post. Tiene, sin embargo, ciertas desventajas (véase el Recuadro 2.3). Primero, la serie de preguntas que pueden hacerse está limitada a un programa existente, y es incapaz de prever los efectos de un programa nuevo o del mismo programa en otro contexto. Segundo, las series de programas que pueden evaluarse en general se limitan a los de cobertura parcial, de manera que puedan construirse grupos de tratamiento y de control, esto usualmente excluye a programas de cobertura nacional. Finalmente, las conclusiones de una evaluación de los efectos de un solo tratamiento puede ser de comparabilidad limitada entre varios estudios. Para responder a estas diversas preguntas evaluativas, sería más adecuado aplicar un método econométrico estructural.

Recuadro 2.3: Comparación de métodos alternativos para la evaluación de programas		
	Método econométrico estructural	Método de efecto del tratamiento
Serie de preguntas atendidas	Evalúa el efecto tratamiento de programas existentes. Pronostica el efecto del programa en un nuevo entorno. Predice los efectos de un programa jamás intentado antes.	Evalúa el efecto del tratamiento de un programa existente. Evalúa un programa en un solo entorno. Incapaz de predecir los efectos de un programa nuevo.
Variedad de programas que pueden evaluarse	Programas de cobertura ya sea parcial o universal dependiendo de la variación de los datos (precios y dotaciones)	Programas de cobertura parcial (grupos de tratamiento y control)
Comparabilidad entre varios estudios	Elevada comparabilidad entre varias evaluaciones (parámetros invariables de programas)	No comparables en general, a menos que las evaluaciones estén diseñadas para una metaevaluación de programas similares
Fuente: Adaptado del Cuadro V en <i>Structural Equations, Treatment Effects And Econometric Policy Evaluations</i> por James J. Heckman y Edward Vytlacil, documento de trabajo NBER No. 11259, marzo de 2005. Obsérvese que en este artículo se propone una síntesis de ambos métodos, que no se toma en cuenta en este cuadro modificado.		

- 2.8 Hay además otras preguntas evaluativas, como las que se refieren a la sostenibilidad institucional o a los resultados cualitativos, que requieren otros métodos de análisis. Estas preguntas no se prestan para evaluaciones ya sean reflexivas o de efecto tratamiento, dado que: (i) las dimensiones que uno quisiera evaluar no son fácilmente medibles, y (ii) la unidad de análisis (instituciones, etc.) son por lo general no lo suficientemente numerosas como para aplicar herramientas estadísticas y econométricas estándar.
- 2.9 Un tema que ha surgido en discusiones con la Administración es el uso del análisis de costo-beneficio en lugar de evaluaciones de impacto. Hay tres puntos que vale la pena mencionar. El primero se refiere a los análisis costo-beneficio ex-ante o análisis de costo-beneficio que se realizan antes de la implementación del programa. El análisis costo-beneficio ex-ante es uno de una serie de simulaciones que se pueden realizar antes de una intervención para obtener una idea de la magnitud de los beneficios y costos⁷. Este análisis

⁷ Recientemente ha habido una proliferación de otras técnicas de simulación que buscan obtener proyecciones de los beneficios de un programa y la distribución de estos beneficios. Referirse por ejemplo a Bourguignon, Ferreira y Leite (2002)

ha sido la principal técnica de simulación ex ante durante los últimos 30 años aproximadamente. Para crédito del Banco, durante la década de los 70 y los 80, éste fue pionero en aplicar ésta metodología como un criterio ex-ante para la elegibilidad de inversiones⁸. Sin embargo, a pesar de que estas simulaciones pueden ayudar a guiar buenas políticas, no pueden responder preguntas referentes a la efectividad de las intervenciones; únicamente proveen ex-ante “las mejores conjeturas” acerca de su efectividad. Esto lleva al segundo punto: el análisis costo-beneficio también se puede realizar ex-post, como una herramienta de evaluación *sui generis*. Sin embargo, en estos casos no son independientes de evaluaciones de impacto; por el contrario, cuando se realizan correctamente éstas dependerán de manera crítica en los resultados de las evaluaciones de impacto para poder asignar parámetros a los beneficios del análisis. En otras palabras, los parámetros obtenidos de una evaluación de impacto debieran proveer información acerca de los beneficios para la realización del análisis costo-beneficio. De hecho, esto se aplica no solo al análisis costo-beneficio sino también a cualquier otro tipo de análisis evaluativo que pretenda comparar los beneficios de diferentes intervenciones y compararlos bajo un criterio similar. El tercer y último punto se refieren a los supuestos que están implícitos en el análisis costo-beneficio. A pesar de que esta técnica posee la ventaja de que identifica todos los costos y beneficios de una política o programa y los resume en una medida monetaria del cambio agregado en términos de bienestar (y por lo tanto ofrece una regla sencilla para los responsables de tomar decisiones sobre la conveniencia de adoptar una política o programa determinado) es neutral con respecto a la distribución de los costos y los beneficios. Para el BID esto constituye un problema dada la preocupación de la institución con respecto a la pobreza: un dólar para los pobres no es lo mismo que un dólar para un millonario.

- 2.10 **Estrategia de muestreo.** La selección de proyectos que realizó OVE fue más planificada que aleatoria⁹, lo que se refleja en la opción de hacer ya sea evaluaciones de proyectos aislados o evaluaciones simultáneas de grupos de proyectos similares (o subcomponentes de proyectos). Una evaluación en grupo permite cuantificar la importancia del contexto del proyecto, a diferencia del modelo de proyecto, en la consecución de los objetivos de desarrollo, incrementando de esa manera la comparabilidad entre diferentes evaluaciones. En la medida en que el contexto del proyecto varíe con el tiempo y en diferentes países, el hecho de considerar proyectos similares en diferentes países y en distintos años permitirá mantener constante el “modelo” y variar el “contexto”. Esto implica evaluar proyectos similares en forma simultánea como “grupos” de proyectos. El término “proyectos similares” se refiere a las operaciones que siguen un modelo común definido por la

⁸ Cuando se comparan las evaluaciones de proyectos realizadas en el contexto de proyectos antiguos del Banco con las prácticas de evaluación actuales, se torna evidente que a pesar de que el Banco estaba a la vanguardia en cuanto a evaluaciones hace dos décadas, no ha logrado mantenerse al corriente de la evolución tecnológica y desde entonces ha perdido sus ventajas comparativas.

⁹ Aparte de la idea de los grupos de proyectos, los criterios de selección concuerdan con los sugeridos en la Política de evaluación ex post: “Los criterios para decidir qué proyectos estarán sujetos a evaluación ex post y la metodología específica que se aplicará en cada caso son, entre otros, los siguientes: i) Las prioridades de OVE según las áreas seleccionadas para evaluación. ii) Los aspectos innovadores de la operación y su potencial para generar lecciones. iii) La pertinencia estratégica en el ámbito del Banco, del sector o del país. iv) El interés particular y la capacidad del prestatario, especialmente considerando que el ejercicio de evaluación puede contribuir a fortalecer la capacidad evaluativa del país.” (página 4 del documento GN-2254-5).

comunalidad del conjunto de insumos, productos y resultados previstos. La comunalidad permite aglutinar esos factores en un solo tema. A su vez, este método facilita la elaboración de una metaevaluación, vale decir, la evaluación de evaluaciones individuales con el fin de extraer lecciones generales. Una metaevaluación tendría mayor valor agregado en términos de generación de conocimiento que la evaluación de un solo proyecto, ya que las conclusiones son comunes a diferentes contextos de ejecución.

- 2.11 En la estrategia de muestreo se adoptaron los siguientes criterios para determinar cuáles grupos de proyectos se seleccionarían. Primero, la pertinencia, que en general se define como la selección de los grupos de proyectos en los que haya falta de conocimiento evaluativo y en los que haya varios proyectos similares en marcha y otros en la programación del Banco. Sin embargo, en casos excepcionales podrían seleccionarse evaluaciones aisladas de proyectos innovadores, si éstos se consideran como posibles candidatos para un futuro “grupo”. Segundo, la viabilidad, que se define como un grupo que contenga muchos proyectos realizados o en ejecución y también en el cual exista una elevada probabilidad de disponer de la información requerida. Este último criterio implicó: i) seleccionar proyectos cuyos resultados directos previstos en función del desarrollo pudieran tal vez obtenerse de otras fuentes (por ejemplo, censos de población o agrarios o encuestas domiciliarias, etc.) en caso de no existir información disponible en el sistema de seguimiento y evaluación del Banco; y ii) seleccionar los proyectos más recientes aprobados o concluidos; vale decir, limitar los proyectos a los aprobados desde 1990. De esta manera se espera poder incrementar la probabilidad de tener acceso tanto a las personas que participaron en los proyectos como la documentación y datos pertinentes.
- 2.12 **Temas y proyectos seleccionados.** Cabe observar que el ciclo de 2004 fue *sui generis*. En primer lugar, para este ciclo, OVE no evaluó “programas” (es decir, operaciones con efectos en toda la economía) a fin de asegurar que los proyectos evaluados tuvieran un conjunto definible de sujetos no beneficiarios, y facilitar así el uso de técnicas convencionales de evaluación centrada en los efectos del tratamiento¹⁰. En segundo lugar, OVE y la Administración se reunieron reiteradamente en 2003 para seleccionar los seis proyectos piloto incluidos en el documento RE-284. Estos proyectos se incorporaron al programa de trabajo de OVE para 2004¹¹, que también incluyó la propuesta de 10 evaluaciones adicionales, escogidas según dos grupos temáticos de proyectos, a saber: *Regularización de la tenencia de la tierra* y *Mejoramiento de barrios*. El diálogo con la Administración (representada por los coordinadores de efectividad en el desarrollo de los tres departamentos operativos y un funcionario de DPP) generó no solamente una lista consensuada de proyectos a ser evaluados, sino que dentro del contexto de un diálogo constructivo, la Administración trajo a la mesa una experiencia colectiva que facilitó la definición más clara de un itinerario de evaluación ex post, en el cual se identificaron a priori muchos posibles obstáculos y se definieron rutas alternativas, en caso de tropezarse con alguno de ellos. De esta manera, la colaboración redujo las asimetrías de información y aceleró el avance a lo largo de la curva de aprendizaje. Finalmente, como se había definido

¹⁰ Este criterio implicó evitar los préstamos en apoyo de reformas de política (tradicionales y en condiciones de emergencia). También se excluyeron del ciclo de 2004 las operaciones de asistencia técnica y las operaciones para el sector privado.

¹¹ Los seis proyectos tuvieron el apoyo financiero de la Administración (Regiones I y II y DPP).

en las conversaciones con la Administración, la conjugación de las metas de numerosos proyectos simultáneamente con una meta metodológica implicó que algunos proyectos originalmente seleccionados tuvieran que desestimarse y adoptarse otros nuevos durante el ejercicio. Por ejemplo, entre los seis estudios piloto originales, el proceso de retroalimentar la información tuvo un éxito relativo y resultó en evaluaciones de efecto tratamiento solamente para tres proyectos (BR-0182, EC-0025, y AR-0169).

- 2.13 Por consiguiente, a principios de 2004 se eliminaron tres de las evaluaciones originalmente propuestas, a saber: i) *Wawa Wasi*, PE-0167, que fue reemplazada por el *Programa de titulación y registro de tierras* (PE-0037). Se escogió este programa por dos razones: primero porque podría incorporarse al grupo de proyectos sobre regularización de la tenencia de la tierra; segundo, porque con el programa peruano se distribuyeron títulos de propiedad a más de 1,1 millón de familias rurales, hecho que lo convirtió en uno de los programas más grandes de reconocimiento oficial de títulos enfocado en zonas rurales en el mundo en desarrollo. Por ello, el programa ofrece la oportunidad de poner a prueba los medios a través de los cuales los títulos de propiedad pueden influir en el bienestar de la familia; y ii) el proyecto ME-0186 que fue sustituido, a mitad del año, por el denominado PROCAMPO (ME-0185). La coexistencia de PROCAMPO con el programa PROGRESA, en México, representaba una singular oportunidad para comprobar la necesidad y la función de condicionalidades en los programas de transferencia de dinero para alivio de la pobreza, ya que había familias que estaban recibiendo transferencias a través de PROCAMPO, y también transferencias condicionadas de dinero a través de PROGRESA. Igualmente se pudo poner a prueba el supuesto frecuentemente admitido de que los beneficios de los resultados directos de transferencias de dinero se maximizan si la destinataria de la transferencia es la mujer de la casa (PROGRESA) y no el jefe de familia (PROCAMPO), sea este último hombre o mujer. El último de los seis estudios experimentales, el *Proyecto nacional del medio ambiente* (ES-0024), a pesar de haberse iniciado en 2004, tuvo que aplazar su fecha de conclusión a 2005 (ya que se necesitaba recolectar datos primarios para proseguir con las actividades, el costo de lo cual quedaba fuera de las posibilidades presupuestarias). Este proyecto fue reemplazado por PROAPS/REMEDiar (AR-0120). REMEDIAR fue seleccionado como evaluación aislada porque constituye un ejemplo de proyecto innovador y una muestra de cómo el BID puede reaccionar rápidamente ante circunstancias imprevistas de crisis económica extrema y además determinar si el diseño del proyecto fue eficaz para atenuar las consecuencias negativas de la crisis sobre las condiciones de salud de los pobres. Si bien el criterio para abandonar proyectos piloto originalmente seleccionados se refirió principalmente a la falta de datos, la mayor relevancia y disponibilidad de datos de los proyectos alternativos, ello requirió rectificar constantemente los acuerdos alcanzados anteriormente con la Administración, factor que debe tomarse en consideración en los futuros ciclos de evaluación ex post.
- 2.14 El Cuadro 1.1 presenta una lista resumida de los proyectos seleccionados en forma original y definitiva. En términos de criterios de relevancia, al momento de seleccionar los proyectos sobre *Regularización de la tenencia de la tierra*, había seis que estaban concluidos, catorce en marcha y cinco que figuraban en la programación del Banco, mientras que para los proyectos sobre *Mejoramiento de barrios*, había nueve proyectos concluidos (de los cuales se seleccionaron siete), ocho en marcha y 10 programados. Con

respecto a los proyectos aislados el Banco aprobó durante los años noventa, los siguientes: 67 proyectos sobre agua y alcantarillado; 18 programas sobre capacitación laboral; 78 fondos de inversión social, y seis programas de transferencias condicionales de dinero. En el cuadro también se aprecia que la mayoría de los proyectos eran operaciones de seguimiento o tuvieron una operación de seguimiento: de los 16 proyectos seleccionados, 13 eran operaciones de seguimiento o tuvieron operaciones de seguimiento, y muchas de estas operaciones de seguimiento fueron aprobadas mientras la anterior se encontraba aún vigente. Finalmente, según la interpretación descrita en el capítulo II, los proyectos seleccionados implicaron un intento de llevar a cabo: i) ocho evaluaciones ex post de resultados y sostenibilidad, es decir, proyectos concluidos dos años o más, pero menos de cuatro, antes de 2004; ii) cinco informes de evaluación ex post del impacto, es decir, proyectos terminados antes de 2000; y iii) cinco de la categoría “otros” (incluidos dos proyectos concluidos después de 2000) y tres que permanecían en marcha. La razón de esta última categoría (que queda fuera de la regla de los dos o cuatro años de la Política de evaluación ex post) fue poner a prueba la hipótesis de que las metodologías de efecto de tratamiento y reflexiva podrían aplicarse a proyectos terminados más recientemente o incluso vigentes.

Cuadro 1.1: Proyectos seleccionados para evaluación ex post						
Cuadro 1.1: Proyectos seleccionados para evaluación ex post						
Proyecto	Título	Fecha aprob./ conclusión	Programa anterior	Detalles	Programa posterior	Detalles
Proyectos de regularización de la tenencia de la tierra						
PE-0037	Programa de Titulación de Tierras	1996/2001	no	-	sí	PE-0107 Progr. Titulación y Registro de Tierras II (2001/2005)
PR-0083	Consolidación de Colonias Rurales	1992/2000	no	-	no	-
EC-0048	Préstamo CT y Sectorial Agrícola	1994/2002	no	-	sí	EC-0191 Regular. y Administr. Tierras Rurales (2001/2007)
JA-0030	Proyecto de Titulación de Tierras	1987/1997	no	-	sí	JA-0050 Prog. Admin. y Ordenación de Tierras (1999/2005)
BL-0007	Proyecto de Administración de Tierras	1997/2001	no	-	sí	BL-0017 Progr. Ordenación Tierras II (2001/2006)
NI-0020	Programa Sectorial Agrícola y CTR	1992/1997	no	-	no	-
Proyectos de Mejoramiento de Barrios						
BR-0182*	Progr. Mejoramiento Urbano Río Janeiro	1996/2000	no	-	sí	BR-0250 Mejoramiento Urbano Río Janeiro II (2000/2006)
EC-0138	Apoyo al Sector Habitacional	1997/2000	no	-	sí	EC-0207 Apoyo al Sector Habitacional II (2002/2006)
CH-0118	Mejoramiento de Lotes con Servicios III	1989/1994	sí	CH0134 Programa Infraestructura Urbana (1982/1988)	sí	CH-0032 Programa Vivienda Progresiva y Mejoramiento de Barrios (1994/1999)
UR-0070	Programa Desarrollo Municipal II Etapa	1990/1998	sí	UR-0028 Programa Desarrollo Municipal (1984/1992)	sí	UR-0111 Programa Desarrollo Municipal III (1997/2004)
TT-0016	Programa Nacional de Asentamientos	1991/2001	no	-	no	-
Proyectos aislados						
EC-0025*	Proyecto de Agua y Alcantarillado Quito	1994/2002	no	-	sí	EC-0200 Programa Saneamiento Ambiental del Distrito Metropolitano de Quito (2002/2006)
PN-0054	Programa de Inversión Social	1994/2001	no	-	sí	PN-0111 Alivio a la Pobreza y Desarrollo Comunitario (1999/2005)
AR-0120*	Reforma de Atención Primaria de Salud	1999/2006	no	-	no	-
AR-0169*	Apoyo Prod. Empleabilidad de Jóvenes	1997/2005	sí	AR-0062 Programa Apoyo Reconversión Productiva (1994-2000)	no	-
ME-0186*	Modernización Mercados de Trabajo II	1996/2000	no	-	sí	ME-0118 Modernización Mercados de Trabajo II (2000/2004)
ES-0024*	Programa Nacional de Medio Ambiente	1995/2006	no	-	no	-
PE-0167*	Programa Conversión Productiva Rural	1998/2005	no	-	no	-
ME-0185	Programa Conversión Productiva Rural	1996/1998	no	-	sí	ME-0213 Apoyo a Pequeños Agricultores a través de PROCAMPO (2001/2005)

Se identifica con letra cursiva los seis estudios piloto propuestos en el documento RE-284. Los proyectos ME-0186, ES0024 y PE0167 fueron eliminados en 2004 por falta de datos, con lo que quedó el grupo de 16 evaluaciones ex post realizadas.

2.15 **Ejecución.** Durante 2004 y principios de 2005, una vez seleccionados los proyectos, se decidió la mecánica de implementación efectivamente aplicada, con frecuentes rectificaciones de enfoque según la viabilidad y otros factores. La implementación consistió en el siguiente conjunto de actividades: Primero, se realizó un estudio documental para determinar lo que se sabía acerca de las intervenciones, para lo cual se recurrió a varias

fuentes del Banco: i) documentos de préstamo y antecedentes complementarios; ii) evaluaciones ex post generadas en el sistema de Evaluación ex post por el prestatario; iii) Informes de terminación del proyecto; iv) Informes de seguimiento del desempeño de los proyectos, y v) metaevaluaciones elaboradas por diversas entidades del Banco. Segundo, OVE entrevistó a los integrantes de los equipos de proyecto y a los especialistas de las Representaciones y, cuando fue posible, a funcionarios del país. Estas medidas tuvieron por finalidad identificar todas las preguntas evaluativas pertinentes, fuentes de información sobre resultados directos y objetivos de desarrollo, y confeccionar listas de contactos en el país, etc., para añadir a la información recopilada en un estudio documental. Esta tarea fue decisiva, ya que los equipos de proyecto y participantes locales cuentan con información que no llega al sistema oficial de información del Banco. Tercero, se decidió sobre la conveniencia de llevar a cabo la evaluación en forma interna o mediante contratación externa (parcial o total), así como sobre los mecanismos de coordinación intrainstitucional que habían de usarse. Se adoptaron diversos mecanismos de coordinación institucional en diferentes contextos, variación que permitió llegar a una primera aproximación para descubrir lo que da “mejores resultados”. Finalmente, se procuró controlar la calidad de las evaluaciones individuales mediante seminarios técnicos ofrecidos al personal del BID y sometiendo los informes individuales a la revisión de técnicos externos.

2.16 Por otra parte, OVE ha comenzado a establecer numerosos mecanismos para edificar un sistema capaz de asistir en la tarea de realizar un gran número de evaluaciones de proyectos en el futuro. Actualmente hay tres redes que se encuentran en proceso de consolidación y que se utilizarán en el ciclo de 2005, a saber:

- (i) IntraEvalNet, integrada por funcionarios del BID. Sus objetivos son agrupar a los funcionarios del Banco interesados en la agenda evaluativa e intercambiar experiencias, lecciones aprendidas, términos de referencia, técnicas, etc., y de esa manera ampliar la capacidad de evaluación.
- (ii) EvalNet, formada por profesionales de la evaluación de programas en América Latina y el Caribe. Mediante la formación de una base de evaluadores cada vez más experimentados que trabajen en América Latina y el Caribe, esta red se propone construir un registro de consultores para facilitar la contratación externa de evaluadores por parte de OVE y el Banco. También aumentará la capacidad evaluativa no gubernamental del país trabajando con técnicos locales especializados en evaluación. Teniendo en cuenta que la contratación externa del trabajo de evaluación es la norma tanto para el BID como para los gobiernos, existe un elevado potencial de rentabilidad de esta red. Hasta abril de 2005 se habían registrado 267 evaluadores en representación de 26 países, de los cuales 230 provienen de países miembros prestatarios.
- (iii) RedEval, formada por entidades públicas involucradas en labores de evaluación. Los objetivos de RedEval consisten en establecer relaciones con las contrapartes oficiales de los países para las evaluaciones de OVE, e intensificar la capacidad evaluativa institucional interna del país. La “red” se encuentra aún en estado de

gestación y requiere bastante esfuerzo, tiempo y presupuesto para empezar a funcionar¹².

III. CONCLUSIONES DE LA EVALUACIÓN DE PROYECTOS

- 3.1 En este capítulo se presentan las conclusiones a las que llegó la oficina a partir de la muestra de proyectos seleccionados para el ciclo de 2004. Las conclusiones se ordenan brevemente por tema y también resaltan algunas evaluaciones de proyectos aislados¹³, para terminar el capítulo con el costo del ejercicio de 2004. Los informes completos pueden consultarse en la red IntraEvalNet:

(<http://ove/oveIntranet/Default.aspx?Action=WUCHtmlContent@EXE>).

A. Revisión por temas: Proyectos sobre Regularización de la tenencia de la tierra

- 3.2 Los proyectos sobre regularización de la tenencia de la tierra pueden comprenderse mejor cuando se consideran como parte de una modificación de la política general de desarrollo. En los años noventa, la política de América Latina evolucionó hacia un régimen pro mercado, en el cual los derechos de propiedad fueron considerados como un requisito indispensable para un desarrollo económico exitoso y para la reducción de la pobreza. En particular, se consideró que los títulos legales de propiedad eran decisivos para que los pequeños agricultores pobres se beneficiaran de los mercados. Se supuso que los derechos de propiedad legalizados reducirían el riesgo de expropiación, que se consideraba que actuaba como un impuesto sobre la rentabilidad de las inversiones y por ende daba lugar a una subinversión. Además, se esperaba que las mejoras realizadas en los derechos de propiedad redujeran los costos de transacción, y por ende incrementaran igualmente el mercado de tierras. De modo similar, bajo el supuesto de que los mercados de crédito eran competitivos, los títulos de propiedad podían utilizarse como garantía prendaria para reducir la prima por riesgo en la contratación de préstamos, y disminuir así la restricción al acceso y los intereses para los prestatarios rurales pobres. Los proyectos de regularización de títulos de propiedad tenían un modelo subyacente común. En términos de productos de los siete proyectos estudiados, cinco contaban con fortalecimiento institucional de la administración catastral y en cinco se identificaron los títulos de propiedad emitidos o registrados. En términos de resultados esperados, siete de ellos pronosticaban mejor seguridad en cuanto a propiedad de la tierra, aumento de la inversión productiva en la tierra; y cinco, mercados de tierras más eficientes y transparentes. Por consiguiente, las evaluaciones de este tema intentaron determinar los efectos tratamiento de los títulos de propiedad sobre las inversiones en el hogar o predio, así como en la compraventa de tierras y el acceso al crédito.

¹² Los beneficios potenciales quedaron en evidencia en la primera conferencia de organismos evaluadores públicos realizada en Brasilia, en enero de 2004, organizada por IPEA, de Brasil, en la cual participaron representantes de más de ocho países, que intercambiaron experiencias sobre los diversos sistemas de evaluación. Una de las principales conclusiones de esa reunión fue la necesidad de establecer una red de entidades públicas involucradas en funciones de evaluación.

¹³ El Anexo 1 contiene una descripción más detallada.

- 3.3 Para ninguno de los proyectos estudiados en esta categoría se definió un conjunto completo de indicadores de objetivos de desarrollo, ni se recopiló ni se mantuvo información adecuada en el sistema de seguimiento y evaluación del Banco. Tampoco se pudieron obtener exitosamente datos de los resultados directos recurriendo al organismo ejecutor o a la unidad ejecutora del proyecto. Seis de los proyectos mencionan evaluaciones ex post en las propuestas de préstamo, pero hasta la fecha no se ha completado ninguna. OVE logró readecuar con éxito un subconjunto de resultados directos previstos para tres proyectos: vale decir, una tasa de atrición de 50%. Esta realidad limita la información disponible en materia de efectividad en el desarrollo, razón por la cual los datos que ofrece para una metaevaluación de calidad son insuficientes.
- 3.4 Las evaluaciones sobre regularización de títulos de propiedad demuestran un efecto tratamiento ambiguo sobre las inversiones productivas. Estos efectos fluctuaron entre no existente en Paraguay hasta un efecto sustancial y positivo en el caso de Nicaragua. El efecto tratamiento sobre el acceso al crédito fue igualmente ambiguo, sin un incremento demostrable de acceso ni en Nicaragua ni en Perú. Se concluyó que los efectos tratamiento sobre los mercados de alquiler fueron positivos solamente en el caso de Nicaragua, como también lo fue el efecto tratamiento sobre productividad agrícola. El único efecto tratamiento que fue grande e inequívoco fue el del valor de las propiedades. Los beneficiarios de los proyectos de regularización de títulos de propiedad comprobaron que los valores de sus tierras aumentaron en los tres proyectos. Sin embargo, para los demás efectos de desarrollo que supuestamente se generarían (mayor productividad, mayor inversión y mejor acceso al crédito), no se detectaron efectos tratamiento claros e inequívocos.
- 3.5 Las conclusiones indican que con estos programas no es dable esperar que solamente por el hecho de poseer un título de propiedad las instituciones crediticias oficiales aceptarán automáticamente el bien raíz como garantía prendaria, especialmente en contextos de racionamiento del crédito y cuando las asimetrías de información predominan en el entorno de las microfinanzas. Específicamente, para granjas más pequeñas (de menos de 20 hectáreas) el mercado crediticio sigue estando racionado y los propietarios no se benefician de la mayor oferta crediticia que supone un título. Este efecto diferencial entre propiedades pequeñas y grandes, indica que en entornos en que los mercados (como los de oferta de crédito formal) implican distorsiones, los agricultores más pequeños y pobres se encuentran en desventaja y puede implicar, debido a que las ventas de terrenos tienden hacia los productores más grandes, una mayor desigualdad en términos de propiedad. Para que los productores pequeños y pobres se beneficien de regímenes pro mercado se requiere que, junto con la emisión de títulos, se reduzcan al mismo tiempo los costos de transacción y las distorsiones del mercado que limitan el acceso al crédito.

B. Revisión temática: Proyectos de mejoramiento de barrios

3.6 Los proyectos de mejoramiento de barrios basados en las primeras experiencias realizadas en Chile y, más tarde, en Brasil se han repetido (o se encuentran programados) en la mayoría de los países prestatarios. Sin embargo, una evaluación de los componentes y objetivos de estos proyectos deja entrever, sorprendentemente, que tienen pocos rasgos comunes. La lista de resultados directos previstos varía considerablemente de un proyecto a otro. El objetivo más frecuentemente citado es el del suministro de servicios públicos (agua, alcantarillado,

Recuadro 3.1: Resultados de las entrevistas con los miembros del equipo del Programa de mejoramiento vecinal

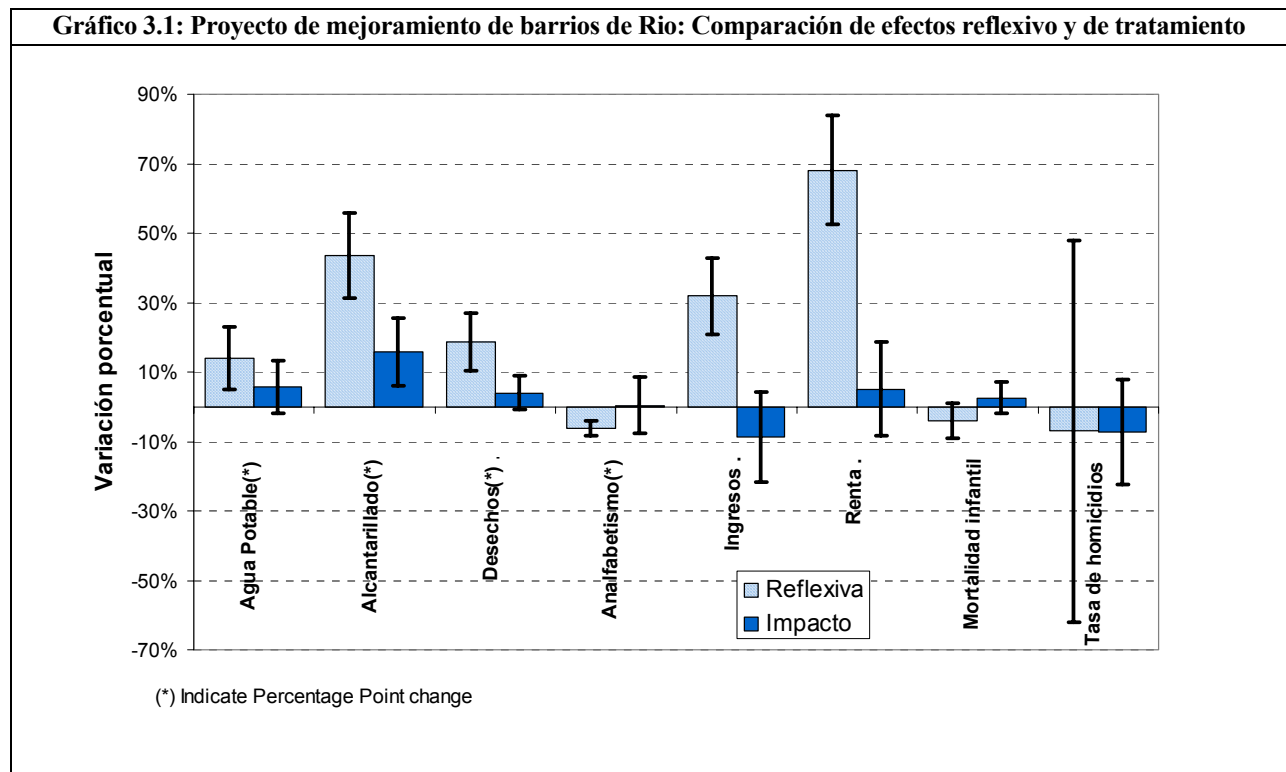
La entrevista realizada por OVE con los integrantes del equipo reveló disparidad de opiniones en cuanto a precisamente qué objetivos se planea lograr al diseñar los proyectos. En algunos casos, los entrevistados mencionaron que el objetivo principal era integrar los barrios en la ciudad, y que esto se lograría reduciendo las diferencias de acceso y en los derechos legales de los habitantes formales e informales. En otros casos, los encuestados pensaban que el aspecto más importante era la reducción de los costos de transacción que genera el mejoramiento vial. En otros casos más, se mencionó el potencial de estas intervenciones para influir en los indicadores sociales, por ejemplo, los de morbilidad y desarrollo infantil. En otros casos se intentó vincular las metas de mejoramiento de barrios con el concepto más amplio de capacidades desarrollado por Sen. En este caso, la opinión dada fue que hay que tratar de ayudar a los barrios informales a alcanzar las metas que se fijan internamente; y por tal razón los distintos proyectos tendrían diferentes metas. Por otra parte, otros entrevistados expresaron una opinión opuesta. En un caso, la respuesta puso énfasis en la dispersión de los objetivos a través de los proyectos como un buen ejemplo de “deslizamiento de metas” dentro de la institución (BID).

etc.), que es en realidad un producto de la mayoría de los proyectos y no tanto un efecto¹⁴. Otros objetivos comunes fueron los siguientes: i) desarrollar los mercados de vivienda, ii) integrar los barrios a la ciudad, y iii) mejorar los indicadores sociales. En términos de productos, se encontraron dos componentes en todos los proyectos: i) obras públicas y ii) fortalecimiento institucional, aun cuando hay bastante variación en la naturaleza de los componentes institucionales. El resto de componentes son muy variados y han llegado a incluir títulos de propiedad, guarderías infantiles, consejería en salud, inversiones orientadas a la comunidad y capacitación de jóvenes, entre otros. La falta de un marco común de Producto→Efecto→Meta parece indicar que, a diferencia de los proyectos de titulación de tierras, no hubo modelos subyacentes comunes en los proyectos de mejoramiento de barrios. Las entrevistas sostenidas con miembros del equipo del programa de mejoramiento de barrios (véase el Recuadro 6, donde se parafrasean las respuestas de los entrevistados) corroboraron esta apreciación. Sin embargo, la característica común de diseño son las intervenciones multidimensionales integradas (multitratamientos), bajo el supuesto de que se produce un efecto de sinergia. Este punto de vista quedó incorporado en la estrategia de desarrollo social del Banco aprobada en 2004.

3.7 OVE tuvo menos éxito en retroalimentar la información de proyectos de mejoramiento de barrios que en los de titulación de tierras. La oficina pudo obtener información previa y

¹⁴ En muchos proyectos se usan las necesidades básicas insatisfechas (NBI) como medida de resultados directos, que simplemente miden el grado de acceso a diferentes servicios (agua, saneamiento, recolección de basura, electricidad, etc.). No obstante, existen estudios publicados que sustentan el punto de vista de que dichas necesidades básicas son una aproximación razonable al bienestar y, por lo tanto, un efecto.

posterior al programa en el caso de un proyecto (Brasil) recurriendo a terceras fuentes (censos, encuestas domiciliarias y datos de la secretaría de salud): una tasa de atrición de 80%. Al menos en un caso (Chile), este enfoque pudo haberse repetido, pero las restricciones presupuestarias no lo permitieron. En dos casos (Ecuador y Uruguay), se intentó identificar el efecto tratamiento del programa observando los gastos por municipalidades en barrios y saneamiento (Uruguay) y observando la amplitud de la cobertura del programa (Ecuador). En el caso de Uruguay, los resultados indican que el gasto subnacional adicional en saneamiento y barrios conduce a una reducción en las necesidades básicas insatisfechas, especialmente las relacionadas con saneamiento. Este efecto tratamiento también se detectó en el caso de Brasil. Por otra parte, este efecto está bien dirigido hacia los pobres, tanto en el caso de Brasil como en el de Uruguay. Sin embargo, en el caso del Ecuador, no se percibió tal efecto. Además, tanto en Brasil como en Uruguay, no hubo efectos tratamiento demostrables del proyecto sobre un índice de calidad habitacional. Lo mismo sucedió en Ecuador. El contraste entre las estimaciones reflexivas y de efectos tratamiento en Brasil, como se aprecia en el Gráfico 3.1, por ejemplo, demuestra la medida en que la efectividad en el desarrollo puede subestimarse o exagerarse en una evaluación reflexiva.



3.8 La conclusión general sobre la pequeña muestra de proyectos evaluados es que lograron proporcionar exitosamente a las poblaciones previstas una mayor cobertura de ciertos servicios públicos. En la medida en que estos servicios sean considerados como bienes meritorios, su suministro constituye un resultado positivo. Además, en dos de los casos, este impacto fue más pronunciado para los segmentos más pobres incluidos en la población

tratada. Más allá de esto, muy poco es lo que puede decirse. El impacto sobre los objetivos relacionados con la formación de capital humano y los ingresos no logró demostrarse. En el caso de las intervenciones en salud —tal vez por el tipo de intervención más directamente vinculada a servicios de saneamiento— no se ha demostrado vínculo alguno entre las intervenciones y los resultados directos, incluso para los segmentos más necesitados de la población beneficiaria. Tampoco hubo evidencia reiterada que demostrara un aumento en las variables relacionadas con los valores habitacionales en los tres proyectos estudiados¹⁵. Con respecto al impacto del componente de titulación de los proyectos, tendrá que realizarse una evaluación adicional debido a que estos componentes no se han desembolsado. La precariedad de las evaluaciones de los efectos tratamiento realizadas en esta área temática deja sin respuesta la interrogante de si un “tratamiento múltiple” conlleva efectos de sinergia o no.

C. Evaluaciones aisladas

- 3.9 Aparte de los resultados ya resumidos de las evaluaciones temáticas, los resultados de los proyectos individuales fueron igualmente útiles para destacar la importancia crucial de la evaluación. El primero es el nexo entre los estudios publicados sobre un modelo y la necesidad de realizar más evaluaciones. En el caso de la evaluación del proyecto de agua, en Quito, la evaluación indica una evidencia preponderante de que el hecho de aumentar la cantidad y mejorar la calidad del agua tiene un importante efecto de tratamiento sobre las condiciones de la salud. Esta conclusión obtenida de estudios publicados indicaría que muy poco es lo que podría aprenderse con evaluaciones adicionales que estudian el agua y la condición de la salud. Sin embargo, incluso siendo así, los resultados de la evaluación revelaron conclusiones sorprendentes y potencialmente cruciales en términos normativos. Los resultados de la evaluación de los efectos tratamiento señalan, como se suponía, un efecto positivo de tratamiento del abastecimiento de agua y servicios de alcantarillado sobre los resultados de salud sustituidos por la mortalidad infantil (en promedio, 8,0% más alta sin él, cifra que es la mitad del 16,5% obtenido con la evaluación reflexiva). Sin embargo, la heterogeneidad de los resultados demuestra que las madres más educadas fueron más capaces de aprovechar los beneficios de la mayor cobertura del recurso hídrico. Esto llevó a una relación regresiva entre el efecto del tratamiento y el ingreso, en cuyo caso las familias más educadas (y más con más recursos) tenían mejor situación que las menos educadas (y más pobres). Estos resultados indican que entre las familias más pobres, la educación de la mujer es un factor decisivo para obtener los beneficios previstos en materia de salud infantil de las intervenciones relacionadas con el abastecimiento de agua y servicios de alcantarillado. La implicación potencial es que los proyectos futuros deberían incluir o ser coordinados (como una hipótesis que cabe intentar) con un componente de educación sanitaria junto con el aumento del suministro de agua.

¹⁵ En el caso del proyecto de Río, hubo dos estudios sobre los valores habitacionales. En un caso, los resultados demostraron una apreciación sustancial en dichos valores; en el otro, no hubo tal apreciación. El análisis realizado por OVE no halló apreciación en los valores habitacionales. En el caso del Ecuador, se seleccionaron las características de las viviendas como sustituto del valor habitacional; no hubo mejora en el tipo de murallas o pisos, el tipo de cielo raso ni el número de habitaciones. En Uruguay, la medida observada fue el número de personas por habitación, y tampoco hubo indicios de un aumento en la construcción. Cabe señalar aquí que el parámetro del impacto se basa en el gasto municipal, que incluye fondos del BID y de otras fuentes.

- 3.10 El Fondo de Inversión Social de Panamá-FIS (PN-0054) destaca una segunda conclusión evaluativa, que demuestra que en algunos casos la evaluación reflexiva de hecho subestimó el verdadero efecto tratamiento del programa. Los resultados directos de este proyecto en términos de desarrollo fueron la reducción de la pobreza. Una evaluación reflexiva (efectos brutos) de la incidencia de pobreza sugirió que no solamente el proyecto no había sido exitoso, sino que de hecho había contribuido a empeorar la pobreza, vale decir, lo contrario de su propósito. Sin embargo, una evaluación de los efectos tratamiento (efecto neto) que comparó “municipalidades similares” demuestra que las municipalidades que se benefician de los recursos del Fondo de Inversión Social tuvieron una reducción significativa de la pobreza en contraste con las municipalidades comparables que no recibieron financiamiento de dicho Fondo; de manera que el proyecto tuvo claros resultados directos positivos en sus objetivos de desarrollo.
- 3.11 En algunos casos, los resultados de la evaluación de los efectos tratamiento fueron críticos para identificar las características de un rediseño óptimo de un programa durante su ejecución. Así sucedió con el *Programa de becas escolares*, en Argentina, un programa típico de transferencia condicional de dinero. La oportuna evaluación logró proporcionar insumos respecto de los plazos y el monto de los subsidios. La pregunta normativa fue ¿cuántas intervenciones (subvenciones) repetidas por año generan el resultado más positivo o es este resultado igual para el mismo número de beneficiarios anuales sin la repetición? Esta pregunta es decisiva para la definición de beneficiarios en un programa que cuenta con un presupuesto limitado. Los efectos estimados del tratamiento demostraron que el resultado más alto en función del costo fue la repetición de las becas durante tres años. Otro punto crucial en el programa de alivio de la pobreza mediante transferencias condicionales de dinero es si el éxito de tales intervenciones orientadas por la demanda depende de las inversiones orientadas por la oferta en las escuelas y en la implementación de programas. La evaluación también abordó esta pregunta. Se intentó estimar las contribuciones de la administración del programa y las características escolares para explicar la variación detectada entre las escuelas en los resultados de los estudiantes. Las conclusiones fueron que la capacidad y las condiciones institucionales para el aprendizaje en las escuelas y la mejor comunicación y ejecución de los procedimientos del programa explican una significativa variación en los resultados directos del programa. Por ende, la oferta importa.
- 3.12 Otro punto de controversia en los programas condicionales de dinero¹⁶ similares al descrito en el párrafo anterior es la condicionalidad misma. En los países pobres, cuya infraestructura e instituciones son menos desarrolladas, el costo administrativo del seguimiento de las condiciones puede ser elevado. Si una transferencia de efectivo por sí sola es suficiente para inducir un cambio de conducta mediante los efectos del ingreso, el ahorro de costo que significa el monitoreo del cumplimiento de las condiciones podría traducirse en una mayor cobertura o en niveles más altos de beneficios para los pobres. Sin embargo, la hipótesis de que la condicionalidad agrega valor y es eficaz en función del costo no se ha evaluado aún. Hay dos programas en México que ofrecen la oportunidad de hacerlo. El primero es PROGRESA, un programa de transferencias condicionales de

¹⁶ Basados en el éxito del programa PROGRESA, de México, iniciado en 1997 y *Bolsa Escola*, de Brasil, iniciado en 1994. Estos programas existen ahora en Argentina, Brasil, Colombia, Ecuador, Honduras, Jamaica, Nicaragua y Uruguay. Esta iniciativa ha pasado a ser la más influyente en materia de alivio de la pobreza en la región.

efectivo que sirve de modelo para otros programas similares en el resto de la región. Las condiciones para que las familias participantes reciban las transferencias incluyen un número mínimo de visitas de atención preventiva a clínicas de salud, un nivel mínimo de asistencia a la escuela (~80%), y participación en un seminario mensual sobre educación sanitaria. El segundo programa, PROCAMPO, iniciado en 1994, es el primero y único programa de apoyo agrícola en América Latina y el Caribe que ofrece beneficios directos en efectivo a los agricultores a fin de compensar las pérdidas que éstos prevén a raíz del Tratado de Libre Comercio de América del Norte. Este programa no incluye ninguna de las condiciones que se imponen en PROGRESA, por lo cual la existencia en el mismo grupo de familias pobres rurales de PROCAMPO y familias de PROGRESA ofrece la singular oportunidad de observar el efecto tratamiento de estos dos programas y tratar de determinar la necesidad que hay de la condicionalidad y su función. Las conclusiones son que el efecto general del tratamiento de ambos programas sobre el gasto total y el gasto en alimentación es el mismo. Sin embargo, el gasto en inversiones en salud y educación es más alto en el caso de PROGRESA. La condicionalidad genera efectivamente un efecto más allá del efecto en el ingreso de la transferencia con respecto a la educación y la salud. PROCAMPO, por otra parte, da como resultado una mayor inversión en actividades productivas.

D. Lecciones aprendidas

- 3.13 Una conclusión crítica a la que se llega en todos los proyectos es la falta de correspondencia entre las estimaciones reflexivas y las estimaciones de los efectos tratamiento. Prácticamente en todos los casos, las estimaciones fueron diferentes. La evaluación del Fondo de Inversión Social de Panamá ilustró el caso en que la evaluación reflexiva condujo a una conclusión de fracaso, mientras que el método de los efectos del tratamiento demostró éxito. El estudio sobre los barrios de la ciudad de Rio de Janeiro demuestra que en un solo proyecto pueden encontrarse prácticamente todas las diferencias posibles entre el criterio reflexivo y el efecto tratamiento (véase el Gráfico 3.1). Comparando una conclusión reflexiva con una del efecto tratamiento, puede afirmarse lo siguiente: i) se sobrestimó un efecto positivo en el caso del abastecimiento de agua potable, alcantarillado y desechos; ii) se subestimó un efecto positivo en el caso de la mortalidad infantil; iii) se detectó un efecto positivo cuando el efecto tratamiento fue negativo (para ingresos y renta); iv) se detectó un efecto negativo cuando el efecto tratamiento fue positivo (para reducción del analfabetismo); v) se detectó que son iguales (para la tasa de homicidios). Las ramificaciones son obvias: los datos de resultados directos únicamente de beneficiarios darán probablemente información engañosa sobre la efectividad en el desarrollo. Además, se demuestra la falsedad del supuesto de que las evaluaciones de los efectos tratamiento siempre arrojan resultados negativos.
- 3.14 Se realizaron seminarios técnicos con funcionarios del Banco para tratar los resultados preliminares de cada una de las evaluaciones efectuadas. Estos seminarios se divulgaron en el Banco en general y no siempre se logró captar el interés de personal clave en cada área temática. Además, los documentos de trabajo se distribuyeron a funcionarios clave para comentarios y crítica. Los puntos planteados en estos ejercicios ofrecen importantes conocimientos en cuanto al diseño de futuras rondas de evaluación ex post. En primer lugar, teniendo en cuenta las difusas responsabilidades que hay en materia de

programación, preparación y ejecución dentro del Banco, varios participantes consideraron que la evaluación, como ejercicio de rendición de cuentas del equipo de proyecto, era contraproducente. Además, el contraste entre la evaluación reflexiva y la de efectos tratamiento planteó cuestiones en torno a la utilidad del seguimiento del desempeño como instrumento de evaluación: ¿Es inútil el seguimiento del desempeño a la luz de las conclusiones según los métodos reflexivo y de tratamiento? Se planteó la utilidad y la pertinencia normativa de analizar tanto los efectos anticipados como los no anticipados; si el proyecto excluía los efectos no intencionales en el marco lógico, ¿implicaba esto que las conclusiones de la evaluación no eran válidas o útiles? Debido a que OVE retroalimentó los datos obtenidos de terceros, los participantes consideraron que, en algunos casos, se midieron variables inadecuadas para un determinado objetivo de desarrollo, lo que dio lugar a conclusiones negativas no sorprendentes. Finalmente, los participantes cuestionaron si el análisis de proyectos terminados, para los que, en el tiempo de aprobación, no se pidió incluir resultados directos ni los datos ahora necesarios, era un uso costo-eficiente de los recursos del Banco. Estos puntos suscitaron inquietudes, que OVE comparte, en cuanto a la Política de evaluación ex post, pero también demuestra la necesidad de pasar a una evaluación determinada por la demanda y en tiempo real. Teniendo en cuenta de las lecciones aprendidas a partir del trabajo de difusión realizado en 2004, la estrategia de 2005 se modificará para asegurar la participación del personal esencial y tomar en cuenta, cuando sea pertinente, las preguntas planteadas en los seminarios técnicos.

E. Situación actual y costo

- 3.15 El documento de antecedentes de la política de evaluación ex post dispuso que en dichas evaluaciones se hiciera “una estimación de sus márgenes de costo” (párrafo 4.8), lo cual sigue siendo problemático, por numerosas razones: Primero, los proyectos seleccionados en 2004 fueron escogidos porque se encontraron datos sobre ellos o porque se pensó que podrían obtenerse a bajo costo (es decir, obtenidos de encuestas o censos de población existentes y datos administrativos del proyecto, etc.). Segundo, hay costos iniciales de indagación que vienen a ser un desperdicio si finalmente no se encuentran datos adecuados para una evaluación de calidad. Tercero, los datos presupuestarios no incorporan el gasto total de las evaluaciones individuales, ya que otros participantes involucrados absorbieron parte de los costos¹⁷.

¹⁷ En general, los costos pueden considerarse no solamente como el gasto que implica realizar evaluaciones individuales, sino también como parte de una formación de capacidad institucional, ya que esta labor también requiere financiamiento. Parte de la estrategia de evaluación ex post de OVE ha sido traspasar, en la medida de lo posible, el gasto de la evaluación de un proyecto a los participantes locales, con la idea de desarrollar la capacidad institucional. Sin embargo, incrementar la capacidad evaluativa del país compartiendo las experiencias de evaluación individual de proyectos aislados es de valor limitado, a menos que se complemente con la ejecución del “Tercer Pilar” del Plan de Acción a Mediano Plazo del Banco para mejorar la efectividad en el desarrollo.

Cuadro 3.1: Costo y situación de las evaluaciones de proyectos				
Cuadro 3.1: Costo y situación de las evaluaciones de proyectos				
Proyecto	Título	Situación	Costo total (\$)	Fondos de la Administración
PE-0037	Programa Titulación de Tierras	Eval. de calidad/Documento de trabajo	91,544	62,841
PR-0083	Consolidación Colonias Rurales	Nota técnica, alguna información para resultados	23,577	
EC-0048	Préstamo CT y Sectorial Agrícola	Nota técnica, estudios cualitativos hechos	23,777	
JA-0030	Proyecto de Titulación de Tierras	Abandonado	12,457	
BL-0007	Proy. Administración de Tierras	Nota técnica	23,577	
NI-0020	Progr. Sectorial Agrícola y CTR	Nota técnica, cálculos de costo y beneficios	24,577	
BR-0182*	Progr. Mejoramiento Urbano Río Janeiro	Eval. de calidad/Documento de trabajo	83,606	29453
EC-0138	Apoyo al Sector Habitacional	Nota técnica, no se pudo identificar beneficios	16,777	
CH-0118 & CH-0032	Mejoram. Lotes con Servicios y Vivienda Progr. y Mejor. Barrios	Abandonado	12,777	
UR-0070 (UR-0111)	Programa Desarrollo Municipal II	Nota técnica	16,777	
TT-0016	Programa Nac. Asentamientos	Abandonado	8,777	
EC-0025*	Proy. de Agua y Alcantarillado Quito *1	Eval. de calidad/Documento de trabajo	89,934	39503
PN-0054	Programa Inversión Social *2	Eval. de calidad/Documento de trabajo	27,777	
AR-0120*	Reforma Atención Primaria Salud	Nota técnica	50,337	24865
AR-0169*	Apoyo Empleabilidad Jóvenes	Eval. de calidad/Documento de trabajo	37,545	12754
ME-0186 & ME-0118	Modernización de Mercados Trabajo II *3	Nota técnica, metaevaluación	24,577	
ES0024*	Prog. Nacional Medio Ambiente	Pospuesto para 2005-2006	20,445	
PE0167*	Atención Menor Wawa Wasi	No realizado, substituido por el Programa de Becas AR-0169	9,777	
ME-0185 & ME-0213	Programa Conversión Productiva Rural y Apoyo Peq.Agric. a través Procampo	Eval. de calidad/Documento de trabajo	50,000	18103

Los estudios piloto se indican con letra cursiva. Las evaluaciones de calidad aparecen en negrita. "Fondos de la Administración" es una fuente ajena al presupuesto de OVE. Nótese que de los fondos transferidos de la Administración (\$210.000), US\$187.515 se gastaron en seis evaluaciones y el saldo se devolvió al fondo común. Las seis fueron las tres originalmente seleccionadas como estudios piloto y tres que reemplazaron a las otras originalmente seleccionadas.

3.16 En el Cuadro 3.1 se observa el costo por proyectos individuales (personal, consultores y viajes) El gasto total fue de US\$648.615. El costo promedio por proyecto fue de US\$34.000 con un elevado grado de varianza (la máxima fue de US\$91.541 y la mínima de US\$8.777). No obstante, se elaboraron solamente seis evaluaciones de calidad, es decir, las que pudieron satisfacer el estándar mínimo (con costo directo promedio de US\$63.401)¹⁸. Casi la mitad del presupuesto, US\$324.000 se gastó en intentos fallidos de realizar evaluaciones. El hecho de incorporar costos totales en forma pareja en seis productos implica un costo promedio de US\$113.694 por producto de calidad. Los costos por evaluación de calidad se sitúan en el extremo inferior de las bandas comúnmente

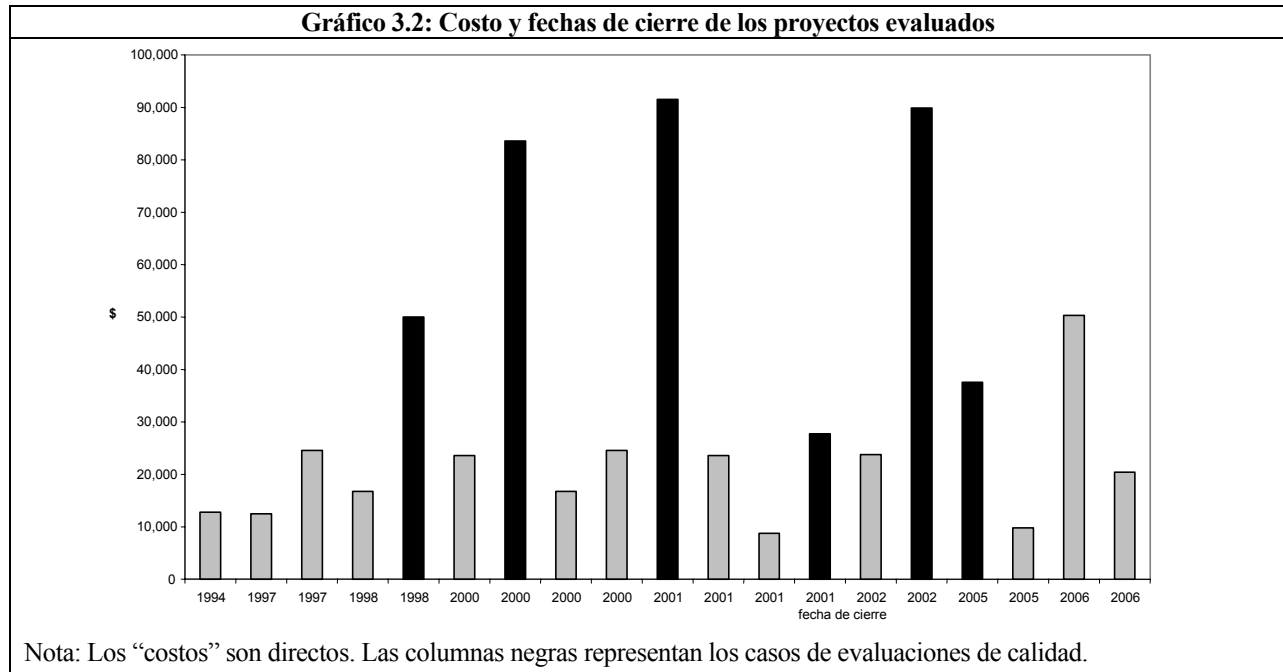
¹⁸ Esto es muy similar al costo por proyecto estimado que se propuso en el estudio experimental: US\$70.000 por evaluación. Cabe observar que la propuesta previó US\$420.000 del fondo de contingencia del Banco para seis evaluaciones. De hecho, se transfirió un total de US\$210.000 en forma pareja de los presupuestos de las Regiones 1 y 3 (ya que la Región 2 adujo no disponer de US\$70.000) y el resto de DPP. Los recursos se pusieron a disposición de OVE en octubre de 2003, dos y medio meses ante del cierre del uso de los fondos el 15 de diciembre de 2003. El resto del costo propuesto, US\$210.000, fue absorbido por OVE en su propio presupuesto.

adelantadas, tanto en valores absolutos en dólares, como en porcentajes del valor del programa en dólares. Los costos de referencia en dólares absolutos van de US\$200.000 hasta US\$900.000 y como porcentaje del programa, entre 5% y 7%¹⁹. Las seis evaluaciones de calidad (véase el Cuadro 3.2) fluctuaron entre US\$87.000 y US\$141.000 y como porcentaje del valor en dólares de los programas (suma de las operaciones de seguimiento) entre 0.01% y 0.21%.

Cuadro 3.2: Resumen de costo de las evaluaciones de calidad				
	Valor del Programa (millones de US\$)	Costo directo (US\$)	Costo total (US\$)	Costo total como porcentaje del valor del programa
PE0037/PE0107	83	91.544	141.837	0,17%
BR-0107/BR-0250	600	83.606	133.899	0,02%
EC-0025/EC-0200	280	89.934	140.227	0,05%
AR-0169	637	37.545	87.838	0,01%
ME-0185/ME0169	1881	50.000	100.293	0,01%
PN-0054	38	27.777	78.070	0,21%

- 3.17 Finalmente, a pesar de haber sesgado la selección de proyectos en favor de tasas elevadas de éxito, probable, la tasa efectiva de éxito en la consecución de las evaluaciones de calidad fue baja (véase el Gráfico 3.2). De las ocho evaluaciones ex post de resultados y sostenibilidad (ERS) propuestas, solamente tres lograron satisfacer el estándar mínimo de los efectos tratamiento. De los cinco informes de evaluación ex post de impacto (IEI) propuestos, sólo uno cumplió la norma y de los “otros” cinco (es decir, proyectos seleccionados pero que no cumplieron la regla de dos y cuatro años de haber finalizado), dos lograron satisfacer la norma. Además, no se advirtió ningún patrón claro de incremento de costo para los proyectos concluidos hace más tiempo (el supuesto originalmente planteado de que un IEI costaría al menos el doble que una ERS no se observó). Sin embargo, la tasa de éxito en la consecución de una evaluación de calidad parece ser más elevado mientras más recientemente haya terminado un proyecto o si aún se encuentra vigente. El éxito en la consecución de una evaluación de calidad para un proyecto abierto indica que el método de tratamiento puede aplicarse a los proyectos abiertos.

¹⁹ Hay considerable variación en el costo de las evaluaciones de impacto. La información procedente del Departamento de Evaluación de Operaciones indica que las evaluaciones del impacto son en general dispendiosas, pero que pueden fluctuar entre US\$200.000 y US\$900.000 según la metodología usada y la amplitud de la recopilación de datos (Banco Mundial 2002, página 21). Otra fuente indica que, como regla empírica, el costo de una evaluación se da entre aproximadamente el 5% y el 7% del costo del programa que se evalúa (Fundación W. K. Kellogg, 1998). A partir de una muestra de las evaluaciones realizadas en América Latina y el Caribe, Judy Baker encontró que el costo promedio de las evaluaciones de impacto era de US\$433.000, pero que dicha suma, como porcentaje del préstamo o crédito o costo total del proyecto, alcanzaba solamente un 0,56%. Un sondeo informal de firmas privadas que realizan evaluaciones de impacto concluyó que una evaluación rigurosa puede costar entre US\$300.000 y US\$350.000. Véase John Blomquist, *Impact Evaluations of Social Programs: A Policy Perspective*, abril de 2003, http://www1.worldbank.org/sp/safetynets/Training_Events/ImpEval_Blomquist_Paper.pdf



IV. LECCIONES APRENDIDAS EN 2004

- 4.1 Este capítulo tiene por objeto extraer lecciones de la experiencia del primer año de aplicación de la Política de evaluación ex post. Como se mencionó anteriormente, los 16 proyectos evaluados fueron considerados experimentales, ya que 2004 fue el primer año de aplicación de la Política. En términos específicos, el documento RE-284 planteó cuatro preguntas operativas, respecto de las cuales se procuró obtener más orientación a través del ejercicio experimental, que comprende los temas de coordinación, ejecución, costo, datos, resultados directos, efectividad y aprendizaje de lecciones mediante difusión.
- 4.2 La información disponible al Banco al momento del diseño de la Política sugirió que: i) la coordinación ex post entre la Administración, OVE y el país resolvería cualquier problema que surgiera en la etapa de la evaluación ex post; ii) la evaluación ex post podría ser una labor anual que empezara y finalizara dentro de un año civil; iii) habría suficiente información evaluativa como insumo para el trabajo ex post; iv) una evaluación de efectos tratamiento sería complicada de llevar a cabo y tendría un costo extremadamente elevado, y por lo tanto el esfuerzo principal del trabajo ex post sería concentrarse en evaluaciones reflexivas que presumiblemente serían sencillas de realizar y de costo relativamente bajo; v) en general, se dispondría de datos sobre los resultados de los objetivos de desarrollo (antes, durante y al final del proyecto), y normalmente no se requeriría una retroalimentación, y que la evaluación ex post podría aprovechar numerosas evaluaciones realizadas por el Banco durante la ejecución. La experiencia del primer año indica que la información disponible en ese momento dió lugar a expectativas demasiado optimistas.
- 4.3 **Coordinación.** La información disponible en 2003 que se usó como insumo en la Política de evaluación ex post indicaba que la coordinación ex post sería necesaria y suficiente para

resolver problemas que surgieran en el trabajo ex post. Sin embargo, la coordinación entre la Administración, OVE y los países prestatarios después de concluidos los proyectos se concreta demasiado tarde como para obtener la información y los datos necesarios para la tarea ex post. Si no se recopilan datos durante la preparación y ejecución de los proyectos, habrá que readecuarlos con todas las complicaciones que esto conlleva (véase el párrafo 4.8). A pesar de estas limitaciones, la calidad de la evaluación aumenta en la medida en que haya coordinación al interior del BID, es decir, que los funcionarios que hayan estado o estén involucrados en el proyecto evaluado se involucren activamente en la evaluación ex post. Teniendo en cuenta el bajo nivel de preparación de informes y mantenimiento de registros (véase más abajo), la coordinación entre la Administración y el prestatario en su responsabilidad compartida de la recopilación y análisis de datos no está dando buenos resultados para poner a disposición información de calidad sobre los resultados directos de los proyectos.

- 4.4 **Ejecución.** La información disponible en 2003 que se usó como insumo en la Política sobre Evaluación indicaba que dentro de un año civil se podría realizar un ejercicio anual de evaluación ex post. Sin embargo, el cronograma inicial establecido para el ejercicio piloto, y de hecho para las 10 evaluaciones adicionales financiadas a través del plan de trabajo de OVE, no fue realista. Incluso las evaluaciones experimentales, que se consideraron que tenían datos suficientes como resultado de un estudio documental, una vez analizados por los evaluadores, dichos datos resultaron insuficientes para satisfacer el estándar metodológico fijado y, por ende, requirieron retroalimentación. Además, los esfuerzos futuros que incluyan una recolección más intensa de datos originales llevarán más tiempo. La experiencia de 2004, con su elevada tasa de disminución, que probablemente será similar en futuros esfuerzos de evaluar un conjunto completo de proyectos, demuestra el grado de trabajo preparatorio que se requiere para la política de evaluación ex post como está concebida actualmente.
- 4.5 **Costo.** La información disponible en 2003 que se usó como insumo para la Política hacía pensar que las evaluaciones de efecto tratamiento (IEI) serían muy costosas, de aproximadamente US\$250.000, mientras que las evaluaciones reflexivas (ERS) serían de costo mucho menor, aproximadamente US\$30.000. Sin embargo—y con todas las advertencias citadas en el capítulo III—las evaluaciones ex post, o incluso las realizadas simultáneamente, parecen ser costosas en términos presupuestarios (aproximadamente el 20% del presupuesto total de OVE en 2004), aunque de mucho menor costo que el previsto en cuanto a la magnitud del préstamo original (menos de 0,5%) o en dólares, un promedio de US\$63.401, sin diferencia significativa entre el estándar de evaluación reflexivo o de efecto tratamiento. La experiencia de 2004 ilustra que existe una amplia variación en los costos. Las evaluaciones que implican la recopilación de datos originales son claramente más costosas que las que utilizan fuentes existentes de datos o que se añaden a un trabajo rutinario de recopilación de información estadística. En cuanto a la complejidad, si hay datos disponibles no se advierte una diferencia significativa entre los métodos reflexivo y de efecto tratamiento. También cabe observar que el Banco y el país normalmente deciden en conjunto cuál será la estrategia evaluativa y acuerdan los términos de referencia de la evaluación y la contratación externa para el trabajo evaluativo. La experiencia de 2004, más la calidad de las personas y las entidades que participan en EVALNET, indican que la

mayoría de los países pueden valerse de profesionales que trabajan en universidades o entidades con o sin fines de lucro que presten servicio²⁰.

4.6 **Datos: Información evaluativa.** La información disponible en 2003 que se usó como insumo para la Política de evaluación ex post partió del supuesto de que se puede edificar una evaluación ex post a partir del diseño y la información existentes, pero esta información simplemente no estuvo disponible para los proyectos estudiados. De los 16 proyectos evaluados en 2004, y tomando en cuenta que para seis de ellos se definió una estrategia de evaluación en la etapa de aprobación, antes de la conclusión del proyecto la mayoría había abandonado esa estrategia durante la ejecución, con la excepción de un proyecto que generó datos que pudieron utilizarse para calcular un efecto tratamiento, ninguno había generado información evaluativa de calidad. En los casos de abandono, la decisión fue tomada por el prestatario. Si bien para todos los proyectos estudiados se emitió un ITP, ninguno de estos informes proporcionó información adecuada sobre la evolución de los resultados directos y objetivos de desarrollo previstos como fruto del proyecto, ni una actualización con respecto a la evaluación prevista al momento de la aprobación. Finalmente, en cuanto al trabajo sectorial elaborado por el Banco, no se encontró ninguna metaevaluación de calidad sobre los grupos temáticos seleccionados ni en las áreas de proyectos aislados. Una metaevaluación de calidad se define como un informe narrativo basado en las conclusiones derivadas de las evaluaciones de los efectos tratamiento de una serie de proyectos similares del BID o externos²¹.

4.7 **Datos: Datos sobre resultados directos.** La información disponible en 2003 que se usó como insumo para la Política de evaluación ex post sugirió que los datos sobre resultados que se requería para una evaluación ex post reflexiva se generaban en general desde el diseño hasta la ejecución y hasta la terminación de todos los proyectos. Sin embargo, de los 16 proyectos evaluados, en ninguno de ellos se informó sobre sus resultados previstos y declarados en términos de desarrollo. La lección general es que el Banco carece de un sistema integral y confiable de recopilación de datos para los resultados directos en función del desarrollo de sus proyectos. Esta realidad esta en desacuerdo con el objetivo declarado por Banco de aplicar una gestión por resultados para tomar decisiones basadas en evidencia y mejorar así la efectividad en el desarrollo. Además, la experiencia recogida por OVE en 2004 también indica que los datos almacenados fuera del BID son relativamente difíciles de obtener y, en algunos casos, son inaccesibles o se han perdido. Las consecuencias inmediatas de esta falta de información es que OVE tiene que tratar de retroalimentar la información en forma retrospectiva desde una época anterior al inicio del proyecto, y hasta cuando éste concluye (labor que corresponde a la Administración),

²⁰ Entre los prestatarios del Banco hay tres modelos institucionales diferenciados de evaluación que al parecer están surgiendo, en los cuales: i) la evaluación está integrada en el sistema presupuestario del país; ii) una oficina especializada en evaluación se encarga de coordinar las evaluaciones; y iii) el Congreso, a través de una comisión especialmente creada, asume la tarea. La rapidez de esta variación entre los prestatarios implica el peligro de que la propia práctica del Banco quede a la zaga de la que apliquen algunos de sus prestatarios. Un tema que salió a menudo en las entrevistas que sirvieron de base a este informe es que si bien el Banco tenía elevadas normas técnicas y capacidad propia en los años setenta y ochenta —tal vez demasiado elevadas para la capacidad de los prestatarios— la situación actual está llegando rápidamente a ser la inversa: el Banco está probablemente quedándose rezagado.

²¹ En una publicación del Banco Mundial se informa de un metaanálisis de las evaluaciones de los fondos de inversión social realizadas en la región, varias de las cuales fueron financiadas por el Banco.

además de la tarea que le asigna la Política de Evaluación Ex Post, que consiste en recopilar información en algún momento después de la terminación del proyecto (es decir, volviendo al Gráfico 1, a partir del punto C y retroceder hasta B y luego hasta A).

- 4.8 **Datos: Retroalimentación.** La información disponible en 2003 implicaba que la necesidad de retroalimentar información constituiría la excepción y no la regla, pero se constató lo contrario. A veces se encontró que era posible llenar el vacío de información. De los 16 proyectos, siete pudieron retroalimentarse exitosamente para *algunos* indicadores de objetivos de desarrollo al menos para una evaluación reflexiva, y seis lo fueron para una evaluación de los efectos tratamiento. Sin embargo, la retroalimentación no sólo es un costo que no previó la Política, sino que también reduce el grado de confianza en la aplicabilidad o relevancia de las conclusiones de la evaluación ex post, ya que los datos restringen la gama de temas cubiertos y los métodos empleados. En términos específicos, las dificultades son las siguientes: Primero, los proyectos no ofrecen indicadores que tengan métrica claramente definida de los efectos que se prevé lograr en función del desarrollo. Esto implica que en una evaluación ex post, el evaluador tiene que readecuar una métrica, creando una fuente potencial de conflicto en cuanto a lo que los diseñadores del proyecto “quisieron proponer” realmente como objetivo de desarrollo²². Una definición ex ante por parte del equipo de proyecto de los indicadores aceptables de los resultados esperados puede ayudar a establecer los parámetros para evaluar la efectividad en el desarrollo de un proyecto. Sin embargo, es importante fijar no solamente metas en términos de productos y resultados directos, sino también identificar, durante el diseño del proyecto, las condiciones bajo las que una meta determinada sea imposible de alcanzar o las circunstancias bajo las cuales se debería rediseñar el programa durante la ejecución²³. Segundo, la

²² En el caso del proyecto EC-0138, la meta del programa fue “mejorar las condiciones habitacionales de los grupos de bajos ingresos.” Sin embargo, no se definieron indicadores para medir estas condiciones. En este caso, a partir de la información disponible en las encuestas nacionales, OVE analizó las condiciones habitacionales a base de los siguientes indicadores: materiales de construcción de los pisos, paredes y techo, conexión de agua dentro de la vivienda, existencia de una ducha y un inodoro, y de un dormitorio de uso exclusivo. El especialista destacado en la Sede estuvo de acuerdo con estos indicadores, pero indicó que el programa alcanzó otros objetivos igualmente importantes relacionados con los propósitos del programa, vale decir, “impartir mayor eficiencia e igualdad al gasto público en vivienda, y proporcionar incentivos para las instituciones financieras comerciales y la industria privada de la construcción para involucrarse en el sector habitacional”, factores para lo cual no figuraban indicadores en el documento de préstamo y para los cuales el especialista no tuvo sugerencias específicas. Otro ejemplo es el *Programa de mejoramiento urbano de Rio de Janeiro* (BR-0182). Este proyecto tenía como resultado directo declarado “reducir la incidencia de enfermedades transmitidas por vectores” (PR-2077, párrafo 2.2), es decir un indicador preciso y medible de resultados directos. Sin embargo, dado que la información no se recopiló sobre la incidencia de estas enfermedades, OVE utilizó datos sobre mortalidad causada por un saneamiento deficiente como el mejor indicador sustitutivo disponible. Un objetivo del *Proyecto de abastecimiento de agua y alcantarillado en Quito* (EC-0025) fue “mejorar la higiene y condiciones de salud de la población de la ciudad de Quito” (PR-1984, párrafo 2.1). En este caso, el indicador es impreciso y puede plantearse aplicando una variedad de métricas específicas. No habiéndose definido indicadores medibles ni recopilado datos, OVE readecuó la información usando datos del censo y utilizó la mortalidad infantil como el mejor indicador sustitutivo disponible.

²³ En los programas vigentes relativos al mercado laboral de México, una de las metas declaradas fue incrementar las tasas de colocación para los desempleados y pasar del 30% al 45% para los que reciben orientación sobre el proceso de búsqueda de empleo, y del 60% al 80% para los que reciben capacitación. Son metas que en reiterados casos no llegaron a alcanzarse. Pero la implicación no es de ofrecer recomendaciones precisas sobre cómo ajustar los programas, sino de que se necesitan más recursos financieros. Es sorprendente que el rendimiento insuficiente se haya interpretado como demostración de que era necesario ampliar el programa.

retroalimentación de la información utilizando fuentes secundarias existentes limita gravemente la viabilidad del análisis de la continuidad de los beneficios: por ejemplo, los datos del censo se producen cada diez años. Tercero, el uso de datos secundarios retroalimentados significa que las preguntas que se hacen y se contestan en las evaluaciones dependieron más de la información encontrada que de la pertinencia y la utilidad de las hipótesis que se pusieron a prueba, lo que equivale a un proceso con la lógica invertida. Cuarto, los datos retroalimentados pueden implicar que las preguntas evaluativas no necesariamente coinciden con los resultados específicos de los objetivos de desarrollo declarados por los proyectos. Un proyecto puede evaluarse utilizando efectos anticipados y no anticipados, pero al menos debe considerar como mínimo los anticipados. La retroalimentación de los datos reduce la probabilidad de hacerlo exitosamente. Quinto, los datos readecuados pueden no corresponder exactamente con los beneficiarios del proyecto²⁴, o bien, los datos retroalimentados pueden no ser aptos para identificar a los beneficiarios²⁵. Sexto, la retroalimentación limita la información del conjunto de variables de control que se pueden usar, y por lo tanto podría reducir la confiabilidad de las conclusiones obtenidas de un efecto tratamiento.

- 4.9 El panorama que forman las conclusiones sobre los datos reportados por el Banco sobre la efectividad en el desarrollo de sus proyectos se basa en una muestra utilizada en el ejercicio ex post de 2004. Las conclusiones pueden depender de la muestra y tal vez no sean aptas para una generalización. Sin embargo, un ejercicio llevado a cabo sobre todos los proyectos vigentes en julio de 2004 reveló que la falta de datos detectada en la muestra de 2004 afecta a toda la cartera del Banco: de los 522 proyectos vigentes, sólo en 99 (18%) se ratificó que se estaban recolectando datos al menos para uno de los objetivos de desarrollo del proyecto (véase el Recuadro 4.1).

²⁴ Por ejemplo, para el proyecto PR-0083 (*Consolidación de colonias rurales*), OVE logró obtener datos de panel sobre 300 familias para los años 1994, 1999 y 2002 proporcionados por académicos de universidades de Wisconsin, Washington y Berkeley. Si bien la fuente contiene información detallada sobre los objetivos de desarrollo del proyecto, no cubre la parte del país enfocada por el proyecto, lo que arroja dudas sobre la pertinencia de las conclusiones para el proyecto del BID.

²⁵ Por ejemplo, en el caso del proyecto UR-0111, todas las provincias rurales recibieron financiamiento. La información disponible no identifica a los beneficiarios dentro de la provincia. Una manera de reducir esa limitación sería obtener datos a un nivel geográfico más desglosado. Pudo haberse reconstruido la información microeconómica, con la asistencia del instituto nacional de estadística, pero las restricciones presupuestarias de 2004 lo impidieron. Las limitaciones presupuestarias también influyeron en el fracaso respecto de los proyectos CH-0118 y CH-0032. En este caso, como en el antes citado, los datos sólo pueden conseguirse fácilmente a un nivel muy agregado. Es probable que el gobierno haya podido generar datos más significativos y desglosados, pero el financiamiento necesario para obtener y convertir los datos a un formato utilizable no estuvo disponible.

Recuadro 4.1: Información reportada por el Banco sobre indicadores de desarrollo de sus proyectos activos

Información sobre resultados directos en función del desarrollo. Para determinar el cumplimiento de la promesa de la política de generar información apta para satisfacer al menos el estándar reflexivo, se llevó a cabo un censo de los proyectos vigentes a julio de 2004. La información reportada por el mismo Banco muestra que un número significativo de proyectos están cumpliendo la norma mínima de información que se fijó para el Banco, en el sentido de que sus *Informes de seguimiento del desempeño de los proyectos* afirman que existe información de línea de base, que hay un sistema de recopilación de datos establecido y que se están recolectando más datos: 33% (v.g. 177 de 522), aunque no el 100% prometido por la Política. Un sondeo de estos 177 proyectos revela un cuadro diferente.

De los 177 proyectos vigentes, el sondeo (respuestas para 129) demostró que sólo 99 de ellos (18% de los proyectos vigentes) ratificaron la existencia de información *al menos para uno de los indicadores de resultados directos de los objetivos de desarrollo*, pero ninguno de ellos la ratificó tener información para todos sus resultados declarados en función del desarrollo.

Para apreciar la facilidad de acceso a los datos, se solicitó la ubicación física de la información. De los 99 proyectos, la ubicación de los datos se distribuyó como sigue: 7 en la Representación (1% de los proyectos vigentes); 85 en el organismo ejecutor (16% de los proyectos activos), y 17 en otros lugares, de los cuales 3 eran desconocidos. Esto provocará dificultades. Si no se piden y se obtienen los datos durante la ejecución del proyecto, la probabilidad de conseguir la información una vez concluido el proyecto es reducida.

Otra dimensión de la disponibilidad de información fue si los datos se encontraban en forma electrónica. Los proyectos que contaban con datos en formato electrónico fueron 39.

La tercera dimensión fue determinar qué estándares metodológicos de evaluación podrían usarse a partir de la información recolectada: 59 (11% de los proyectos vigentes) informaron que una norma reflexiva (5 en la Representación) y 34 (7% de los proyectos vigentes) que un estándar de efecto tratamiento (3 en la Representación).

En cuanto al financiamiento del sistema de recopilación de datos, 45% estaba financiado por una cooperación técnica o un componente del préstamo, el 40% por fondos de contrapartida y el resto por “otros”.

A menudo se afirma que mientras más recientemente se haya aprobado el proyecto, más probable es que exista una estrategia de evaluación, y que haya una satisfactoria recolección de datos durante la ejecución. Los 99 proyectos organizados por año de aprobación demuestran un incremento en el curso de los años noventa, pero llegan al punto culminante en el 2000 y luego se registra una caída. Las cifras representan los límites superiores, ya que no consideran la calidad de la información recopilada, y suponen que la recolección de información no se abandona durante la ejecución de los proyectos.

4.10 Además, aun cuando los institutos estadísticos de los países tengan la capacidad para diseñar y recolectar información para las evaluaciones de los efecto tratamiento de los proyectos individuales a un precio razonable²⁶, de acuerdo con un estudio reciente de OVE²⁷, ningún organismo de estadística de la región ha estado involucrado en la generación de datos para la evaluación de algún proyecto específico del BID.

4.11 **Efectividad: Medición de la efectividad: Evaluación reflexiva frente a evaluación de efecto tratamiento/evaluación de impacto.** Una evaluación reflexiva puede arrojar conclusiones engañosas, ya sea sobre la magnitud y la orientación de los efectos en el desarrollo o sobre ambas cosas. Las evaluaciones realizadas en el ciclo de 2004 revelan que

²⁶ Los institutos de estadística indicaron que la incorporación de preguntas en las encuestas existentes, en módulos de encuestas existentes o en encuestas ad hoc costaría entre US\$15.000 y US\$30.000.

²⁷ Luis Marcano y Luis Miguel Lavista. Disponibilidad de datos de los proyectos del Banco Interamericano de Desarrollo: conclusiones de dos encuestas, OVE, 2005.

en muy pocos casos fueron equivalentes las conclusiones sobre los efectos de desarrollo entre los resultados reflexivos y los de efecto tratamiento. Dado que la orientación y la magnitud del sesgo suelen desconocerse a priori, el estándar mínimo ha de ser uno que produzca un cálculo del efecto tratamiento cada vez que sea posible. Si bien no se dispone de información completa sobre los costos, la experiencia de 2004 indica que este método puede ser de un costo asequible.

- 4.12 Sin embargo, como se observó en el capítulo II, la evaluación de efectos tratamiento responde a un conjunto de preguntas evaluativas que en el ejercicio de 2004 se enfocaron en el efecto de un programa sobre sus metas de desarrollo. No obstante, en otros tipos de análisis, como los de incidencia de los beneficios, las relaciones entre beneficios y costos y otros resultados intermedios directos son también extremadamente útiles para los diseñadores y planificadores de programas.
- 4.13 **Lecciones aprendidas, productos y difusión.** El ciclo ex post de 2004 generó 14 productos escritos y seis seminarios:
- (i) Seis documentos de trabajo elaborados para las evaluaciones que cumplieron el estándar mínimo de efecto tratamiento pasaron con éxito la prueba del sistema de revisión externa. En términos de la definición de plazos para el *Informe de evaluación ex post del impacto* y de la *Evaluación ex post de resultados y sostenibilidad*, estas evaluaciones representan dos de los primeros y tres de las segundas. El sexto documento es una evaluación coincidente de un proyecto en marcha (AR -0169).
 - (ii) Para las evaluaciones que no cumplieron el estándar mínimo de evaluación de efecto tratamiento se prepararon ocho notas técnicas, producto que no se incluyó en la Política de Evaluación Ex Post, pero que representa el mejor intento de documentar los datos sobre efectividad en el desarrollo que hubo disponibles con respecto al proyecto sometido a evaluación. Cinco de las notas se refieren a proyectos terminados más de cuatro años antes de la fecha de la evaluación, dos a proyectos cerrados dos a cuatro años antes y uno a un proyecto vigente.
 - (iii) Seis seminarios técnicos internos del BID para discutir los documentos de trabajo.
- 4.14 Con todo, una tarea que permanece pendiente es la difusión de las conclusiones entre los países prestatarios. Se hizo lo posible por asegurar que los ministerios y organismos ejecutores pertinentes recibieran copias de todos los documentos emitidos. Además, una vez completo el proceso de revisión en el Banco, todos los documentos se pondrán a disposición en el sitio virtual de la institución. En el documento RE-284 se anuncia que se realizará una conferencia con el fin de informar acerca del ciclo de 2004, conferencia que se organizará junto con la Administración, o bien se establecerán actividades alternativas de difusión.

V. CONCLUSIONES Y RECOMENDACIONES

5.1 Las conclusiones del ciclo de 2004 pueden sintetizarse en seis puntos principales:

- La clave para una exitosa evaluación ex post es contar con un diseño evaluativo de calidad al momento de la aprobación y con recopilación de datos durante la ejecución del proyecto. Se advirtió que la definición de resultados directos de objetivos de desarrollo durante la preparación y la recopilación de datos durante la ejecución son la excepción y no la regla, con la consecuencia de que ello obliga a realizar una evaluación retroalimentada. El actual caudal de datos que guarda el Banco no sirvió para realizar una evaluación reflexiva ni de efecto tratamiento²⁸. Lo que sí puede realizarse es una evaluación retroalimentada, pero no exenta de problemas. Antes que no tener nada, este procedimiento incrementa los costos de la evaluación y parte de la base de datos disponibles de terceros, en lugar de usar los datos que evidencien más exactamente los objetivos de desarrollo de un proyecto.
- Para responder a las preguntas evaluativas planteadas en la Política de evaluación ex post, las evaluaciones reflexivas probablemente arrojarán conclusiones engañosas. Puede considerarse que una evaluación de efecto tratamiento que establezca vínculos causales entre las acciones y los resultados directos sea la norma deseada para proyectos públicos de cobertura parcial, complementada por análisis de costos, incidencia de beneficios y efectos institucionales, entre otros elementos. El costo de estas evaluaciones depende de los puntos antes citados, pero son marginales con respecto a la magnitud del préstamo.
- Las evaluaciones de proyectos aislados son difíciles de valorar; no es sencillo separar los efectos que se derivan de la idiosincrasia o el contexto de ese proyecto frente al resultado sistemático del diseño del proyecto. Este problema puede abordarse evaluando un grupo de proyectos similares.
- El éxito de la agenda de los resultados y el papel de la evaluación ex post dentro de la misma depende del desarrollo que haga el país prestatario de su propia agenda de gestión por resultados y de evaluación, y no solamente de la capacidad interna del

²⁸ Si se aplican eficazmente las recientes iniciativas para implantar la visión de la Política de evaluación ex post en cuanto a evaluabilidad, teniendo en cuenta un período promedio de desembolso de diez años, el sistema en “tiempo real” prometido por dicha Política estará funcionando a plenitud en 2014.

Banco. El sistema actual de “responsabilidad compartida” no está funcionando bien ni para el seguimiento concurrente ni para la evaluabilidad ex post de un proyecto²⁹.

- Una mejor y más pronta colaboración entre la Administración, OVE y los prestatarios en la planificación y ejecución de las evaluaciones conduce a la preparación de evaluaciones de mejor calidad y mayor relevancia. Además, el valor agregado de la evaluación concurrente y colaborativa indica que tal vez los plazos que exige la Política con respecto a los informes de evaluación ex post del impacto y las evaluaciones ex post de resultados y sostenibilidad sean innecesarios.
- Las seis evaluaciones de efecto tratamiento realizadas durante 2004 demuestran que las intervenciones del Banco tienen un significativo efecto en al menos un objetivo de desarrollo declarado. Estas conclusiones indican que posiblemente actualmente el Banco esté subestimando su contribución al desarrollo.

- 5.2 Durante el ciclo 2005 puntos que no fueron abordados de frente serán tratados, tales como: (i) la aplicabilidad del análisis costo-beneficio, (ii) mayor énfasis en análisis institucionales y de sostenibilidad en las evaluaciones ex-post de sostenibilidad de 2005, (iii) analizar las consecuencias para el Banco y para el País de contar con estimaciones de impacto engañosas. Las siguientes recomendaciones tentativas plantean sugerencias para el ajuste de la práctica de evaluación del Banco (coincidente con el proyecto o ex post). Son tentativas porque OVE considera que se requiere más información, que se obtendrá de la experiencia de al menos un año más, antes de que pueda abordarse un conjunto definitivo de recomendaciones.
- 5.3 Para responder a las preguntas evaluativas planteadas en la Política, tanto para el informe de evaluación ex post de impacto como para la evaluación ex post de resultados y sostenibilidad, el estándar mínimo de evaluación de proyectos con cobertura parcial es probablemente la de la evaluación de efecto tratamiento. La veracidad de esta afirmación se verificará en el ciclo de 2005. Probablemente deberían eliminarse los plazos exigidos para la evaluación y diferenciarse mejor los productos de evaluación en términos de las preguntas evaluativas que se hayan de responder. Las normas de evaluación para los proyectos que inciden en toda la economía deben desarrollarse en futuros ciclos de evaluación, pero dependerán probablemente de técnicas de econometría estructural (véase el párrafo 2.7) y otros métodos, incluidos el análisis de costo-beneficios.
- 5.4 El ciclo de 2005 debe enfocar su labor de evaluación en un grupo de proyectos similares, que sean de mayor relevancia para la programación y la cartera. Se han seleccionado

²⁹ Pocos proyectos del BID empiezan con evaluaciones de calidad como parte de su diseño original, menos aún las llevan a cabo y casi ninguno queda a disposición del público. Si bien algunos observadores señalan que el prestatario es responsable de la falta de una evaluación durante el diseño o de abandonarla durante la ejecución del proyecto, los resultados de evaluación en el ejercicio de 2004 estuvieron principalmente basados en información recopilada a través de los sistemas permanentes de información del prestatario. Muy pocos resultados dependieron de la información contenida en los sistemas de información del Banco. De hecho, considerando que los sistemas del prestatario aportaron la mayor parte de los datos cuantitativos sobre resultados de desarrollo, la principal contribución de los sistemas del Banco —en el ciclo de 2004— fue ya sea la información cuantitativa sobre productos o los datos anecdóticos y de opinión respecto de los resultados de desarrollo.

proyectos que pueden agruparse en tres temas, a saber: i) programas de capacitación laboral, ii) caminos rurales, y iii) ciencia y tecnología (el Anexo 3 contiene una lista y una breve descripción de los proyectos). Hasta la fecha, la propuesta incluye 48 proyectos: 11 sobre capacitación laboral (en los que Argentina y México tienen operaciones consecutivas); 16 sobre caminos rurales (en los que Ecuador, El Salvador, Guatemala, Honduras, Nicaragua y Perú tienen operaciones consecutivas); y 19 sobre ciencia y tecnología (en los que Argentina, Brasil, Chile, Colombia, México, Uruguay y Venezuela tienen operaciones consecutivas). La lista de proyectos cubre 16 países. Para obtener el número de evaluaciones necesarias para una metaevaluación, se realiza un sobremuestreo de los proyectos al principio del ciclo de evaluación³⁰.

- 5.5 Es preciso establecer una coordinación institucional de mayor magnitud y más formal con los fines de: i) llevar a cabo evaluaciones ex post del conjunto actual de proyectos del BID y ii) mejorar la probabilidad de realizar evaluaciones ex post de calidad en el futuro. Para el ciclo de 2005, se han formado grupos consultivos temáticos integrados por profesionales del Banco, nombrados por la Administración, que participaron en los proyectos que se estaban revisando o que son especialistas en la materia. Se espera lograr una mayor colaboración con funcionarios del Banco, a fin de mejorar la relevancia de las preguntas evaluativas que se formulen y el sentido de propiedad de los resultados. Los prestatarios, por intermedio de los organismos ejecutores, entidades de evaluación y ministerios de hacienda y planificación, también participarán en las evaluaciones de 2005. La incorporación de autoridades nacionales tiene igualmente el objetivo de incrementar el sentido de propiedad de la evaluación. El desarrollo de relaciones de trabajo con las autoridades del país se formalizará en el futuro a medida que entre en vigor la iniciativa PRODEV del Banco, que tiene como una de sus tareas la de intensificar “... *la planificación y evaluación de los proyectos de inversión* ...” (párrafo 9)³¹.
- 5.6 Para elevar la probabilidad de realizar evaluaciones ex post de calidad en el futuro, OVE comenzará a evaluar la calidad del diseño evaluativo al momento de la aprobación y recopilación de datos durante la ejecución del proyecto. En 2005, OVE diseñará e implantará un sistema de información, en forma conjunta con la Administración, con el fin de cumplir la sugerencia de la Política de Evaluación de que “*En el caso de estas [evaluaciones de impacto] ... OVE procurará identificar operaciones que estarán sujetas a este tipo de evaluación lo más pronto posible durante el ciclo de proyectos (v.g., durante la programación y el diseño de las operaciones)*” (párrafo 3.2).

³⁰ En el ciclo de 2005 habrá sobremuestreo de los proyectos contenidos en cada grupo; en la lista de la primera ronda se incluyen 48 proyectos. Esta estrategia parte de la base de las conclusiones contenidas en el ciclo de 2004: i) reducida probabilidad de encontrar datos adecuados en el sistema de seguimiento y evaluación del Banco; ii) baja tasa de retroalimentación exitosa de fuentes secundarias, y iii) los resultados de desarrollo creíbles que sean atribuibles al proyecto requieren como mínimo una evaluación de efecto tratamiento. En su conjunto, implican la necesidad de una estrategia de sobremuestreo para llegar tanto al número requerido (unos 15 proyectos individuales evaluados al año) y uno en el cual el estándar mínimo sea apto para preparar estimaciones de efecto tratamiento y reflexivas. Se prevé una tasa de disminución de aproximadamente 65%.

³¹ Véase el documento GN-2346-2.

- 5.7 El éxito final de la gestión por resultados depende de la adopción de un modelo decisorio basado en evidencias entre sus prestatarios. Pero esto no ocurrirá de la noche a la mañana. En el entretanto, el propio sistema del Banco revela un vacío que es inversamente proporcional a su ambición estratégica, cuya corrección requiere la adopción de medidas inmediatas. A la luz de lo anterior, se hacen las siguientes tres recomendaciones.
- 5.8 Primero, adoptar la evaluación ex ante como una actividad de rutina en la preparación del proyecto y la evaluación de efecto tratamiento como el estándar preferido. Es preciso identificar claramente la métrica para los objetivos de desarrollo que se prevean para el proyecto; es necesario llenar el marco lógico del proyecto con cifras pero eso no es suficiente. El Banco debe incorporar la evaluación ex ante durante la fase de preparación y asegurar que durante la etapa de aprobación del proyecto se diseñe una estrategia de evaluación que genere datos adecuados y que los datos se recopilen durante la ejecución. El Banco debe adoptar la norma de explorar la posibilidad de incorporar el sistema estadístico del país lo más pronto posible y, para facilitar la incorporación de información en el sistema estadístico del país, considerar los beneficios de conceder una excepción global en materia de adquisiciones a fin de autorizar la contratación directa por los organismos ejecutores de los institutos estadísticos que funcionen en el país.
- 5.9 Segundo, reformular la base institucional para generar datos adecuados para las evaluaciones de efecto tratamiento en todos sus proyectos durante la ejecución. El trabajo reciente se ha concentrado en mejorar el diseño de los informes de seguimiento y en adosar al sistema existente los nuevos requisitos de información. Hasta el momento, los resultados han sido decepcionantes. Esto era de esperar; el sistema actual fue diseñado para generar la información (financiera y sobre productos) requerida para las evaluaciones de desempeño y operación, y en la práctica continúa siéndolo. La agenda de efectividad en el desarrollo crea necesidades de datos que son significativamente diferentes y sustancialmente más amplios. La recopilación de datos es menos frecuente, más extensa (de beneficiarios, no beneficiarios y variables de control) y requiere una combinación de aptitudes completamente diferente, a la vez que una estructura institucional de contrapartida para el actual sistema de seguimiento del desempeño.
- 5.10 Tercero, es preciso que el Banco defina oficialmente e institucionalice mecanismos que aseguren que las lecciones extraídas de las evaluaciones ex post se conviertan en lecciones aprendidas, es decir, la base institucional para la adopción de decisiones basadas en evidencias.
- 5.11 OVE sugiere que durante el año 2005 la Administración y la Oficina trabajen en forma coordinada para preparar un plan de acción que establezca un marco que permita conjugar su ambición estratégica con una efectiva capacidad y práctica institucional (el marco entonces debería contener una descripción detallada de los estándares metodológicos, la combinación de aptitudes requerida, la base institucional y los requisitos presupuestarios) .