



NGÂN  
HÀNG  
THẾ  
GIỚI

NHỮNG ĐỊNH HƯỚNG TRONG PHÁT TRIỂN

# Đánh giá Tác động của các Dự án Phát triển tới Đói nghèò

*Sổ tay dành cho cán bộ thực hành*

(Sách tham khảo)

JUDY L. BAKER



NHÀ XUẤT BẢN VĂN HÓA – THÔNG TIN

**NGÂN HÀNG THẾ GIỚI**

**NHỮNG ĐỊNH HƯỚNG TRONG PHÁT TRIỂN**

**Dánh giá tác động  
của các Dự án Phát triển  
tới Đói nghèo**

**Sổ tay dành cho cán bộ thực hành**

*(Sách tham khảo)*

Judy L.Baker

*Người dịch: VŨ HOÀNG LINH*

**NHÀ XUẤT BẢN VĂN HÓA - THÔNG TIN  
Hà Nội - 2002**

Copyright © 2001  
The International Bank for Reconstruction  
and Development/ THE WORLD BANK  
1818 H Street, N.W.  
Washington, D.C.20433, U.S.A.

Giữ mọi bản quyền  
Sản xuất tại Hợp chúng quốc Hoa Kỳ  
Lần ấn bản đầu tiên tháng Năm, 2000

Các kết quả, diễn giải và kết luận trong cuốn sách này hoàn toàn là của (các) tác giả và không thuộc về Ngân hàng Thế giới, các tổ chức trực thuộc, các thành viên trong Ban Giám đốc hay các nước mà những người này đại diện. Ngân hàng Thế giới không đảm mức độ chính xác của các dữ liệu trong ấn phẩm này và không nhận trách nhiệm về hậu quả của việc sử dụng những dữ liệu này.

Tư liệu trong ấn phẩm này được giữ bản quyền. Ngân hàng Thế giới khuyến khích sự phổ biến các công trình nghiên cứu của mình và thường sẽ nhanh chóng cấp giấy phép cho việc tái chế các phần của công trình này.

Về giấy phép sao chụp một số phần để sử dụng nội bộ hay cá nhân, cho những khách hàng cụ thể, hay để sử dụng trong lớp học vì mục đích giáo dục, đề nghị trả trực tiếp mức phí phù hợp cho Trung tâm Bản quyền, Inc., 222 Rosewood Drive, Danvers, MA 01923, Mỹ; điện thoại 978-750-8400, fax 978-750-4470. Xin liên hệ với Trung tâm Bản quyền trước khi tiến hành sao chụp.

Về giấy phép in lại các bài viết hay các chương riêng, đề nghị fax bản yêu cầu với đầy đủ thông tin tới Phòng Tái xuất bản, Trung tâm Bản quyền, fax 978-750- 4470.

Tất cả các yêu cầu khác về các quyền và giấy phép cần liên hệ với Văn phòng Xuất bản, Ngân hàng Thế giới ở địa chỉ trên hoặc fax theo số 202-522-2422.

---

---

# Mục lục

Lời Nhà xuất bản .....	vii
Lời nói đầu .....	.ix
Lời cảm ơn .....	xii
1. Định nghĩa các khái niệm và kỹ thuật đánh giá tác động .....	1
2. Các bước quan trọng trong thiết kế và thực hiện đánh giá tác động .....	24
3. Áp dụng các phương pháp phân tích để đánh giá tác động: Một nghiên cứu tình huống* .....	59
4. Kinh nghiệm rút ra từ các đánh giá tác động “thành công trên thực tế”** .....	97

## Phụ lục

Phụ lục 1: Các nghiên cứu tình huống .....	132
1.1 Đánh giá lợi ích tới người nghèo từ các chương trình công việc: Chương trình TRABAJAR của Argentina... 132	
1.2 Liệu tài chính vi mô có thực sự giúp đỡ người nghèo không? Các bằng chứng mới từ những chương trình tiên phong ở Băngladét .....	142

1.3	Chương trình đổi thực phẩm lấy giáo dục ở Băngladét: Đánh giá một chương trình xã hội có định hướng khi có sự phân quyền trong tổ chức chương trình . . . . .	148
1.4	Đánh giá Quỹ đầu tư xã hội của Bôlivia . . . . .	154
1.5	Tác động của các chương trình lao động tích cực: Cộng hoà Séc . . . . .	161
1.6	Tác động của tín dụng cho giáo dục tới dinh dưỡng của bà mẹ và trẻ em: Chương trình Ngân hàng nông thôn vùng Hạ Pra ở Gana . . . . .	169
1.7	Sách giáo khoa và điểm thi: Bằng chứng từ một đánh giá tiềm năng ở Kênia . . . . .	175
1.8	Đánh giá dự án khuyến nông của Kênia . . . . .	182
1.9	Tác động của chương trình đào tạo lại của Mêhicô tới việc làm và tiền lương (PROBECAT) . . . . .	191
1.10	Mêhicô, Chương trình Giáo dục, Y tế, và Dinh dưỡng quốc gia (PROGRESA): Đề xuất đánh giá . . . . .	200
1.11	Đánh giá về cải cách trường học ở Nicaragua: Phương pháp kết hợp định lượng và định tính . . . . .	207
1.12	Cải cách giáo dục môn toán ở bậc tiểu học tại Nicaragua: nghiên cứu thử nghiệm về tác động của sách giáo khoa và Radio tới kết quả học tập . . . . .	216
1.13	Tác động của các chương trình hoàn trả chi phí khác nhau tới khả năng tiếp cận và bình đẳng ở Nigiri . . . . .	223
1.14:	Các kết quả đi học trong các trường tiểu học ở Philippin: Đánh giá tác động của bốn thử nghiệm . . . . .	229
1.15:	Đánh giá tác động tới đời nghèo của các dự án đường bộ nông thôn ở Việt Nam . . . . .	236
<b>Phụ lục 2: Mẫu điều khoản tham chiếu . . . . .</b>		242
2.1	Dự án dinh dưỡng và phát triển trẻ em ở Uganda . . . . .	242
2.2	Đánh giá tác động của đường bộ nông thôn: Điều tra cơ sở ở Việt Nam năm 1997 . . . . .	268
<b>Phụ lục 3: Một mẫu ngân sách từ một đánh giá tác động của một chương trình cho trẻ ăn tại trường . . . . .</b>		277

<b>Phụ lục 4: Các chỉ số tác động - Đánh giá Quỹ đầu tư xã hội Bôlivia .....</b>	<b>.280</b>
<b>Phụ lục 5: Mẫu về khung logarit cho tóm tắt thiết kế dự án phục vụ cho tài liệu hoàn thành dự án hay tài liệu đánh giá dự án (CAS) .....</b>	<b>.288</b>
<b>Phụ lục 6: Ma trận phân tích .....</b>	<b>.293</b>

**Hộp**

<b>Hộp 1.1</b>	Vấn đề thiên vị trong lựa chọn .....	7
<b>Hộp 1.2</b>	Tóm tắt các phương pháp định lượng nhằm đánh giá tác động của chương trình .....	9
<b>Hộp 1.3.</b>	Tóm tắt các phương pháp được sử dụng để đánh giá các chính sách điều chỉnh .....	17
<b>Hộp 2.1</b>	Các bước chủ yếu trong thiết kế và thực hiện đánh giá tác động .....	25
<b>Hộp 2.2</b>	Các điểm chủ yếu để xác định các nguồn dữ liệu cho đánh giá tác động .....	31
<b>Hộp 3.1</b>	Các bước trong phương pháp tương hợp mức độ xu hướng .....	76
<b>Hộp 3.2</b>	Các nguyên nhân gây ra sai lệch trong ước lượng đơn giản về tác động của PROSCOL .....	78
<b>Hộp 3.3</b>	Thực hiện phương pháp khác biệt kép .....	84
<b>Hộp 3.4</b>	Các thước đo nghèo đói .....	89
<b>Hộp 3.5</b>	So sánh nghèo đói khi có và không có chương trình .....	90

**Bảng**

<b>Bảng 2.1</b>	Các phương pháp đánh giá và yêu cầu dữ liệu tương ứng .....	42
<b>Bảng 2.2</b>	Các công cụ thu thập dữ liệu chủ yếu để đánh giá tác động .....	48
<b>Bảng 4.1</b>	Tóm tắt các đánh giá tác động "thành công trên thực tế" .....	100
<b>Bảng 4.2.</b>	Tóm tắt các chi phí ước tính từ các đánh giá tác động của Ngân hàng Thế giới .....	116

---

---

## Lời Nhà xuất bản

Đánh giá tác động là một phương pháp do lường kết quả của một chương trình hay dự án trên cơ sở tách biệt với các nhân tố có thể khác.

Đánh giá tác động nhằm vào việc xác định dự án có tạo ra được những tác động mong muốn, và tác động đó có phải do việc thực hiện dự án mang lại hay không? Các nhận thức thu được từ các nghiên cứu đánh giá dự án sẽ cung cấp đầu vào quan trọng cho việc thiết kế đúng đắn các dự án trong tương lai.

Cuốn "*Đánh giá tác động của các dự án phát triển tới đời nghèo – Sổ tay dành cho cán bộ thực hành*" này nhằm mục đích cung cấp cho các nhà quản lý dự án và các nhà phân tích chính sách các công cụ cần thiết cho việc đánh giá tác động của dự án.

Nội dung cuốn sách, ngoài Lời nói đầu, gồm có 4 chương và 6 phụ lục.

*Chương 1:* Trình bày một cách tổng quát về các khái niệm và phương pháp đánh giá tác động.

*Chương 2:* Nêu các bước quan trọng trong thiết kế thực hiện đánh giá tác động.

*Chương 3:* Minh họa các kỹ thuật phân tích đa dạng thông qua một nghiên cứu tình huống.

*Chương 4:* Thảo luận các bài học kinh nghiệm từ những đánh giá thành công của một số dự án về đời sống.

Trong phần Phụ lục, đề cập tới các nghiên cứu tình huống, và nêu những ví dụ tiêu biểu về các thành phần quan trọng cần

thiết cho việc hoạch định bất kỳ một đánh giá tác động nào.

Những ví dụ minh họa trong cuốn sách tập trung vào việc đánh giá tác động của các dự án hướng tới người nghèo.

Chúng tôi dịch và xuất bản cuốn sách trên với mong muốn cung cấp cho bạn đọc một tài liệu tham khảo bổ ích trong lĩnh vực này.

Trong quá trình biên dịch, mặc dù người dịch, những người biên tập đã có nhiều cố gắng, nhưng do có nhiều thuật ngữ kỹ thuật mới, chưa có những khái niệm tương đương ở nước ta nên khó tránh khỏi những chỗ khó diễn tả, thiếu sót...

Nhà xuất bản xin giới thiệu cuốn sách và mong nhận được những ý kiến đóng góp của bạn đọc.

*Hà Nội, tháng 6 năm 2002*  
NHÀ XUẤT BẢN VĂN HÓA – THÔNG TIN

---

---

## Lời nói đầu

Mặc dù đã có hàng tỷ đôla được chi tiêu để hỗ trợ phát triển mỗi năm, nhưng người ta vẫn còn biết rất ít về tác động thực sự của các dự án tới người nghèo. Có bằng chứng rộng rãi cho thấy về lợi ích của tăng trưởng kinh tế, đầu tư vào vốn con người và cung cấp các mạng lưới an toàn cho người nghèo. Nhưng đối với một chương trình hay một dự án cụ thể ở một nước nhất định, liệu nó có mang lại các lợi ích mong muốn hay không, và tác động chung của nó tới người dân là gì? Liệu chương trình hay dự án có thể được thiết kế tốt hơn nhằm đạt được các kết quả dự định không? Liệu các nguồn lực có được sử dụng một cách có hiệu quả không? Đó là các loại câu hỏi chỉ có thể trả lời thông qua đánh giá tác động, một phương pháp đo lường kết quả của một chương trình trên cơ sở tách biệt với các nhân tố có thể khác.

Nhiều chính phủ, tổ chức và nhà quản lý dự án ngần ngại trong việc tiến hành đánh giá tác động do họ cho rằng, việc đánh giá này tốn kém, mất nhiều thời gian và phức tạp về mặt kỹ thuật, và do các kết quả có thể nhạy cảm về chính trị, đặc biệt là khi các kết quả có tính tiêu cực. Nhiều đánh giá cũng bị phê phán do kết quả có được quá muộn, không trả lời đúng câu hỏi, hay không tiến hành phân tích chính xác. Một trở ngại nữa là, do hạn chế về mức độ sẵn có và chất lượng của các số liệu.

Mặc dù vậy, với việc lập kế hoạch đúng đắn và ngay từ đầu, với sự hỗ trợ của các nhà hoạch định chính sách, và một khoản đầu tư tương đối nhỏ so với chi phí toàn bộ của dự án, việc đánh giá chính xác có vai trò rất quan trọng trong việc

đánh giá sự phù hợp và hiệu quả của các chương trình phát triển. Đánh giá tác động đặc biệt quan trọng ở các nước đang phát triển, nơi mà các nguồn lực khan hiếm, và mỗi đôla chi tiêu phải nhằm tối đa hoá tác động tới việc giảm đói nghèo. Nếu các chương trình được thiết kế tồi, không đến được những đối tượng dự kiến, hay bị lãng phí, thì với thông tin phù hợp, các chương trình này có thể được thiết kế lại, cải thiện hay loại bỏ, nếu cần thiết. Những kiến thức thu nhận được từ các nghiên cứu đánh giá dự án sẽ cung cấp các đầu vào quan trọng cho việc thiết kế đúng đắn các chương trình và dự án trong tương lai.

Cuốn sổ tay này nhằm mục đích cung cấp cho các nhà quản lý dự án và các nhà phân tích chính sách các công cụ cần thiết cho việc đánh giá tác động của dự án. Nó nhằm tới các độc giả đã có những kiến thức chung về thống kê. Đối với một số phương pháp thống kê chuyên sâu hơn được thảo luận trong sách, độc giả có thể tham khảo các tài liệu kỹ thuật về chủ đề này. Chương 1 trình bày một cách tổng quát về các khái niệm và phương pháp, Chương 2 thảo luận các bước quan trọng và những vấn đề liên quan cần xem xét trong khi tiến hành, Chương 3 minh họa các kỹ thuật phân tích đa dạng thông qua một nghiên cứu tình huống, và Chương 4 thảo luận về các bài học thu được từ những kinh nghiệm phong phú trong các đánh giá "thành công trên thực tế" của các dự án về nghèo đói được xem xét trong cuốn sổ tay này. Các nghiên cứu tình huống đề cập tới trong Phụ lục 1, được lựa chọn từ những đánh giá của Ngân hàng Thế giới, các tổ chức tài trợ khác, các tổ chức nghiên cứu, và các doanh nghiệp tư vấn tư nhân. Chúng được lựa chọn do sự chính xác trong phương pháp luận, bao quát nhiều bối cảnh quốc gia, các loại dự án, và phương pháp đánh giá đa dạng. Trong phần Phụ lục cũng có những ví dụ tiêu biểu về các thành phần quan trọng cần thiết trong hoạch định bất kỳ một đánh giá tác động nào - các điều khoản tham chiếu, ngân sách, chỉ số tác động, khung logarit, và ma trận phân tích mẫu.

Mặc dù các kỹ thuật sử dụng trong đánh giá tác động

tương tự nhau giữa các khu vực và các tiểu nhóm dân cư khác nhau, nhưng các ví dụ minh họa về những phương pháp và nghiên cứu tình huống trong cuốn sổ tay này tập trung vào việc đánh giá tác động của dự án hướng tới người nghèo. Tác động tới nghèo đói có thể bao gồm rất nhiều câu hỏi về dự án và đánh giá, chẳng hạn như do lường tác động của các chương trình tín dụng vi mô tới thu nhập hộ gia đình, tác động của một chương trình đào tạo tới việc làm, tác động của một chương trình dinh dưỡng tại trường tới sự tham dự của học sinh, hay tác động của việc xây dựng đường nông thôn tới phúc lợi của các hộ gia đình.

Không phụ thuộc vào loại dự án hay câu hỏi cần trả lời, việc thiết kế trong mỗi công việc đánh giá tác động sẽ có tính đặc trưng riêng và phụ thuộc vào các nhân tố như loại dữ liệu sẵn có, năng lực của địa phương, các mối quan tâm về thời điểm và ngân sách. Cuối cùng, để các đánh giá dự án có kết quả chất lượng cao, đáng tin cậy, và có thể khai quát hoá được cho các nhà hoạch định chính sách, đòi hỏi phải có sự hỗ trợ tài chính và chính trị vững chắc; việc hoạch định từ sớm và cẩn thận; sự tham dự của các bên hữu quan trong thiết kế các mục tiêu và phương pháp nghiên cứu; dữ liệu đầy đủ, sự phối hợp phù hợp các phương pháp khác nhau, kể cả các kỹ thuật định lượng và định tính; sự ứng dụng chính xác các kỹ thuật này; và liên lạc giữa các thành viên trong nhóm trong suốt tiến trình.

---

---

## Lời cảm ơn

Việc chuẩn bị cho cuốn sách này đã nhận được những đóng góp quý giá của một nhóm cơ sở. Tôi xin cảm ơn các thành viên trong nhóm dưới đây về các tài liệu và các nhận xét có ích của họ: Gillette Hall (các nghiên cứu tình huống, các bài học), Julia Lane (các nghiên cứu tình huống, các bài học), Martin Ravallion (nghiên cứu tình huống, phương pháp phân tích), và Laura Rawlings (các vấn đề trong thực hiện, các bài học) ; và các công trình đánh giá dự án do Kene Ezemenari, Gloria Rubio, Anders Rudqvist, và K. Subbarao thực hiện. Các nghiên cứu cơ sở do Matthew Fleming và Samir Stewart tiến hành. Cuốn sách này nhận được sự đồng hỗ trợ từ Khu vực châu Mỹ Latinh và vùng Caribê và Mạng lưới Giám nghèo đói và Quản lý Kinh tế của Ngân hàng Thế giới dưới sự lãnh đạo của Norman Hicks, Guillermo Perry, và Michael Walton. Công trình này cũng nhận được sự giúp đỡ lớn từ những nhận xét của Omar Arias, Sabina Alkire, Michael Bamberger, Soniya Carvalho, Wendy Cunningham, Norman Hicks, Shahidur Khandker, Norbert Schady, và Quentin Wodon.

---

---

# Chương 1

## Định nghĩa các khái niệm và kỹ thuật đánh giá tác động

Trong các tài liệu, đánh giá toàn diện được định nghĩa là một đánh giá bao gồm việc giám sát, đánh giá quá trình, đánh giá chi phí - lợi ích, và đánh giá tác động. Tuy nhiên, các thành phần này rất khác nhau. Việc giám sát giúp đánh giá xem một chương trình có được thực hiện như kế hoạch không. Một hệ thống giám sát chương trình cho phép có sự phản hồi liên tục về tình hình thực hiện chương trình, phát hiện các vấn đề cụ thể khi chúng xuất hiện. Đánh giá quá trình liên quan tới cách thức mà chương trình hoạt động, và tập trung vào các vấn đề trong việc cung cấp dịch vụ. Đánh giá chi phí - lợi ích hay đánh giá chi phí - hiệu quả là đánh giá các chi phí của chương trình (bằng tiền tệ hay phi tiền tệ), đặc biệt là trong mối liên hệ giữa chi phí với các cách sử dụng khác đối với những nguồn lực và với các lợi ích mà chương trình tạo ra. Và cuối cùng, đánh giá tác động nhằm vào việc xác định một cách chung hơn, liệu chương trình này có tạo ra những tác động mong muốn tới các cá nhân, hộ gia đình và các thể chế, và liệu những tác động này có phải do việc thực hiện chương trình mang lại hay không. Các đánh giá tác động cũng có thể phát hiện những hậu quả không dự kiến trước, có thể là tích cực hay tiêu cực, tới những đối tượng thụ hưởng. Cuốn sổ tay này đặc biệt quan tâm tới mức độ các lợi ích của dự án đến với người nghèo và tác động của các lợi ích này tới phúc lợi của người nghèo. Một số câu hỏi được đề cập tới trong đánh giá tác động như: Dự án có tác động như thế nào tới những đối tượng thụ hưởng? Liệu một sự cải thiện nào đó có phải là kết quả trực tiếp do dự án mang lại hay không? Thiết kế của chương trình có thể được điều chỉnh để cải thiện tác động không? Liệu các chi phí có hợp lý không?

Tuy nhiên, không dễ trả lời các câu hỏi này một cách đơn giản bằng kết quả của một dự án. Có thể có những nhân tố hay sự kiện khác có tương quan với kết quả mà không phải do dự án tạo ra. Để đảm bảo tính chính xác về phương pháp luận, việc đánh giá tác động phải ước tính được trường hợp phản chứng, tức là điều gì có lẽ đã xảy ra nếu như dự án không được tiến hành. Chẳng hạn, nếu một người mới tốt nghiệp từ một chương trình đào tạo lao động nhận được việc làm, thì liệu đây có phải là kết quả trực tiếp của chương trình này không? Hay là người này dù sao cũng vẫn tìm được việc làm? Để xác định được trường hợp phản chứng, cần phải sàng lọc tác động của sự can thiệp của các nhân tố khác - một công việc tương đối phức tạp. Có thể làm được điều này nhờ sử dụng các nhóm đối chứng hay nhóm kiểm chứng (các nhóm không tham dự vào chương trình hay nhận được những lợi ích), các nhóm này sau đó sẽ được so sánh với nhóm được nghiên cứu (các cá nhân tham dự chương trình). Các nhóm kiểm chứng được lựa chọn một cách ngẫu nhiên từ cùng tổng thể như là những người tham dự chương trình, trong khi nhóm đối chứng chỉ đơn giản là nhóm không tham gia chương trình được điều tra này. Cả nhóm đối chứng và nhóm kiểm chứng đều cần phải rất giống với nhóm được nghiên cứu, và sự khác biệt duy nhất giữa các nhóm này là việc tham dự vào chương trình.

Xác định trường hợp phản chứng là trọng tâm của thiết kế đánh giá. Có thể làm được điều này nhờ sử dụng một số phương pháp gồm hai nhóm lớn: các thiết kế thử nghiệm (ngẫu nhiên) và các thiết kế bán thử nghiệm (không ngẫu nhiên). Tuy nhiên, việc sàng lọc tác động của chương trình ra khỏi các điều kiện phản chứng là rất khó khăn và có thể bị tác động của quá khứ, sự thiên lệch trong lựa chọn, và hiện tượng nhiễm (contamination). Các phương pháp định tính và tham dự có thể được sử dụng để đánh giá tác động này. Các kỹ thuật này thường cho phép hiểu rõ về quan điểm của người thụ hưởng, giá trị của chương trình đối với người thụ hưởng, các quá trình có thể tác động tới kết quả và có sự diễn giải sâu sắc hơn đối

với các kết quả quan sát trong phân tích định lượng. Ưu nhược điểm của mỗi phương pháp này sẽ được thảo luận chi tiết dưới đây. Như độc giả sẽ nhận thấy, không có kỹ thuật nào là hoàn hảo cả, và do vậy, người đánh giá phải đưa ra các quyết định về những sự đánh đổi khi lựa chọn mỗi phương pháp. Tuy nhiên, việc lập kế hoạch sớm và kỹ lưỡng, sẽ cho phép có nhiều sự lựa chọn phương pháp hơn trong việc thiết kế đánh giá.

### Các thiết kế thử nghiệm

Các thiết kế thử nghiệm, còn được gọi là sự lựa chọn ngẫu nhiên, thường được xem là phương pháp đánh giá tốt nhất. Bằng cách phân bổ ngẫu nhiên sự can thiệp (intervention) trong số những người thụ hưởng đủ tư cách, bản thân quá trình phân công này sẽ tạo ra các nhóm được nghiên cứu và nhóm kiểm chứng tương tự về mặt thống kê với nhau, với các quy mô mẫu phù hợp. Đây là một kết quả rất quan trọng, bởi vì trong lý thuyết, các nhóm kiểm chứng được tạo ra nhờ việc phân công ngẫu nhiên đóng vai trò phản chứng hoàn hảo, không gặp phải những vấn đề thiên lệch trong lựa chọn vẫn tồn tại trong tất cả các sự đánh giá. Lợi ích chính của kỹ thuật này là sự đơn giản trong diễn giải kết quả - tác động của chương trình tới kết quả được đánh giá có thể được đo lường bằng sự sai lệch giữa các giá trị trung bình của mẫu trong nhóm được nghiên cứu với nhóm kiểm chứng. Một ví dụ là, đánh giá về các sách giáo khoa ở Kenya, trong đó người đánh giá lựa chọn phân bổ ngẫu nhiên các khu vực thực hiện chương trình, tiến hành một cuộc khảo sát cơ sở; tạo ra các nhóm kiểm chứng, và sau đó tiến hành thực hiện chương trình với nhóm được nghiên cứu, trong trường hợp này là việc cung cấp sách giáo khoa. Việc có cả nhóm kiểm chứng và nhóm được nghiên cứu, cho phép người đánh giá có thể xác định một cách rõ ràng tác động của sách giáo khoa tới việc học tập của học viên.

Trong khi các thiết kế thử nghiệm được xem là phương pháp tối ưu trong ước tính tác động của dự án, thì trên thực tế có thể

có một số khó khăn. Thứ nhất, việc lựa chọn ngẫu nhiên có thể không phù hợp về đạo đức do mục đích nghiên cứu, nó phủ nhận những lợi ích hay dịch vụ mà các thành viên trong tổng thể lẽ ra có thể nhận được. Một ví dụ cực đoan là, không tiến hành điều trị y tế trong khi việc này có thể cứu sống nhiều người trong dân cư. Thứ hai, có thể gặp khó khăn về chính trị khi tiến hành chương trình với một nhóm, và không tiến hành với nhóm kia. Thứ ba, do quy mô chương trình, trong một số trường hợp sẽ không có các nhóm không được nghiên cứu như với một dự án hay một sự thay đổi chính sách có quy mô lớn ví dụ các khoản cho vay điều chỉnh hay các chương trình được thực thi ở cấp độ quốc gia. Thứ tư, các cá nhân trong các nhóm kiểm chứng có thể thay đổi một số đặc tính dễ nhận thấy trong quá trình thử nghiệm, và điều này có thể làm cho các kết quả trở nên không có giá trị hay bị sai lệch. Chẳng hạn, nếu người dân di chuyển vào hay ra khỏi khu vực tiến hành dự án thì họ có thể gia nhập hay rút lui khỏi nhóm được nghiên cứu hoặc nhóm kiểm chứng. Một cách khác, những người không nhận được lợi ích từ một chương trình có thể tìm kiếm nó thông qua những nguồn khác, hoặc những người được chương trình này phục vụ lại không muốn nhận lấy nó. Thứ năm, có thể khó đảm bảo việc phân công thực sự ngẫu nhiên. Một ví dụ có thể là, việc những người thực hiện sẽ loại trừ các đối tượng có nhiều rủi ro ra khỏi chương trình nhằm đạt được kết quả tốt hơn. Và cuối cùng, các thiết kế thử nghiệm có thể là quá tốn kém về tiền bạc và thời gian trong một số tình huống, đặc biệt là trong việc thu thập dữ liệu mới.

Với việc lập kế hoạch kỹ lưỡng, một số vấn đề này có thể được giải quyết trong khi thực thi các thiết kế thử nghiệm. Cách thứ nhất là lựa chọn đối tượng thụ hưởng một cách ngẫu nhiên. Điều này có thể cho phép có cơ chế phân công minh bạch về chính trị và là cơ sở cho sự thiết kế đánh giá tốt, bởi vì những hạn chế về ngân sách và thông tin làm cho việc xác định và phục vụ cho những người thụ hưởng xứng đáng nhất một cách chính xác trở nên không khả thi. Cách thứ hai là đưa

nhóm kiểm chứng vào trong chương trình ở giai đoạn sau, khi đã thiết kế và bắt đầu thực hiện đánh giá. Trong kỹ thuật này, sự lựa chọn ngẫu nhiên quyết định *khi nào* đổi tương thụ hưởng đủ điều kiện sẽ nhận được chương trình, chứ không phải là *nếu như* họ nhận được nó. Cách này đã được sử dụng trong việc đánh giá một chương trình dinh dưỡng ở Côn Minh, nó còn có ích hơn nữa do đã trả lời các câu hỏi liên quan tới thời gian cần thiết để chương trình trở nên hiệu quả trong việc làm giảm tình trạng suy dinh dưỡng (McKay 1978). Cuối cùng, sự lựa chọn ngẫu nhiên có thể được áp dụng trong một tập hợp nhỏ những đổi tương thụ hưởng có mức độ thích hợp như nhau, trong khi tới được với tất cả những người thích hợp nhất và không mang lại lợi ích cho những người ít thích hợp nhất, như đã được thực hiện trong các dự án giáo dục ở khu vực El Chaco trong việc đánh giá vốn xã hội của Bolivia (Pradhan, Rawlings và Ridder 1998). Tuy nhiên, nếu như thực hiện cách thứ hai, chúng ta phải luôn nhớ rằng, các kết quả có được từ đánh giá này sẽ chỉ áp dụng cho nhóm mà tại đó mẫu được lấy ra một cách ngẫu nhiên.

### Các thiết kế bán thử nghiệm

Có thể sử dụng các phương pháp bán thử nghiệm (không ngẫu nhiên) để tiến hành đánh giá khi không thể xây dựng được các nhóm được nghiên cứu và nhóm đối chứng nhờ thiết kế thử nghiệm. Các kỹ thuật này tạo ra những nhóm đối chứng tương tự nhóm được nghiên cứu, ít nhất cũng trong những đặc điểm quan sát, thông qua các phương pháp kinh tế lượng, gồm có các phương pháp tương hợp (matching methods), khác biệt kép (double difference), biến số công cụ (instrumental variables) và so sánh phản thân (reflexive comparisons) (xem Hộp 1.2). Khi các kỹ thuật này được sử dụng, các nhóm được nghiên cứu và đối chứng thường được lựa chọn sau khi thực hiện chương trình bằng cách sử dụng phương pháp không ngẫu nhiên. Do vậy, các kiểm chứng thống kê phải được áp dụng để khắc phục

những khác biệt giữa các nhóm được nghiên cứu và đối chứng, và các kỹ thuật tương hợp phải được sử dụng để tạo ra một nhóm đối chứng giống nhóm được nghiên cứu càng nhiều càng tốt. Trong một số trường hợp, một nhóm đối chứng cũng được lựa chọn trước nhóm nghiên cứu, mặc dù sự lựa chọn này không phải là ngẫu nhiên.

Lợi ích chủ yếu của các thiết kế bán thử nghiệm là chúng có thể được xây dựng từ những nguồn số liệu sẵn có, và do vậy, thường nhanh chóng và rẻ tiền hơn khi thực hiện, và chúng có thể được tiến hành sau khi chương trình đã được thực hiện, khi có đầy đủ số liệu hiện tại. Nhược điểm chính của các kỹ thuật thử nghiệm là (a) độ tin cậy của các kết quả thường giảm sút do phương pháp ít chính xác về thống kê hơn; (b) các phương pháp có thể phức tạp về thống kê; và (c) có vấn đề thiên vị trong lựa chọn (selection bias). Trong việc tạo ra một nhóm đối chứng thay vì ngẫu nhiên lựa chọn, có thể có nhiều nhân tố tác động tới độ tin cậy của các kết quả. Sự phức tạp thống kê đòi hỏi phải có kỹ năng tốt trong thiết kế đánh giá và trong việc phân tích và diễn giải các kết quả. Điều này có thể là không khả thi, đặc biệt là trong một số tình huống ở các nước đang phát triển.

Vấn đề thứ ba về sự thiên vị liên quan việc một chương trình được các nhóm nhỏ trong tổng thể mục tiêu tham dự ở các mức độ khác nhau, do đó tác động tới mẫu và tới kết quả cuối cùng. Có hai loại thiên vị: thiên vị do sự khác biệt trong những gì quan sát được hay do số liệu, và thiên vị do sai khác trong những gì không quan sát được (không ở trong số liệu), thường được gọi là sự thiên vị trong lựa chọn (Hộp 1.1). Một thiên vị quan sát được có thể bao gồm các tiêu chuẩn lựa chọn một cá nhân như vị trí địa lý, nhập học, hay tham dự vào thị trường lao động. Những nhân tố không quan sát được mà có thể tác động sai lệch tới các kết quả của chương trình có thể bao gồm, năng lực cá nhân, sự sẵn sàng làm việc, các mối quan hệ gia đình và quá trình chủ quan (thường do động lực chính trị) trong việc lựa chọn các cá nhân tham gia chương trình. Cả hai loại

### Hộp 1.1 Vấn đề thiên vị trong lựa chọn

Thiên vị trong lựa chọn liên quan tới các nhân tố không thể quan sát, có thể dẫn đến sai lệch kết quả (chẳng hạn, năng lực cá nhân, các điều kiện tồn tại từ trước). Các thử nghiệm ngẫu nhiên khắc phục vấn đề thiên vị trong lựa chọn nhờ tạo ra một nhóm kiểm chứng thử nghiệm bao gồm những người có thể tham gia chương trình nhưng đã bị từ chối một cách ngẫu nhiên sự tham gia này. Sự phân công ngẫu nhiên không làm mất đi tính thiên vị trong lựa chọn nhưng có tác dụng làm cân bằng sự thiên vị giữa các mẫu thành viên tham dự và các mẫu gồm những người không phải thành viên. Trong các thiết kế bán thử nghiệm, các mô hình thống kê (ví dụ, các mô hình tương hợp, sai lệch kép, biến công cụ) được sử dụng để mô hình hóa quá trình lựa chọn nhằm có được một giá trị ước lượng không sai lệch so với các dữ liệu phi thử nghiệm. Ý tưởng chung là so sánh giữa những người tham dự chương trình và những người không tham dự trong khi duy trì các quá trình lựa chọn không thay đổi. Giá trị của mô hình phụ thuộc vào mức độ thành công mà mô hình này được định nghĩa.

Một ví dụ là tiền lương của phụ nữ. Các số liệu đại diện cho những người phụ nữ lựa chọn làm việc. Nếu phụ nữ ra quyết định này một cách ngẫu nhiên thì chúng ta có thể bỏ qua thực tế là không phải tất cả các tiền lương đều có thể quan sát được và sử dụng phương pháp hồi quy thông thường để ước tính mô hình tiền lương. Thế nhưng quyết định đi làm của phụ nữ không được đưa ra một cách ngẫu nhiên- những phụ nữ có tiền lương thấp có thể sẽ không lựa chọn làm việc do tiền lương cá nhân dự trữ của họ cao hơn tiền lương đề nghị mà người thuê nhân công đưa ra. Do vậy, mẫu tiền lương quan sát được của phụ nữ sẽ bị sai lệch theo hướng cao hơn thực tế.

Có thể khắc phục được sự sai lệch này nếu như có một số biến khác tác động mạnh mẽ tới các cơ hội để trở thành biến quan sát (tiền lương dự trữ) nhưng không tác động tới kết quả được nghiên cứu (tiền lương đề nghị). Biến này có thể là số lượng trẻ con ở nhà.

Nguồn: Greene (1997).

thiên vị này có thể dẫn đến các kết quả không chính xác, gồm cả việc đánh giá quá thấp hay quá cao tác động thực tế của chương trình, đánh giá tác động là tiêu cực khi tác động thực tế là tích cực (và ngược lại), và đánh giá tác động là không có ý nghĩa về thống kê trong khi tác động thực tế lại có ý nghĩa thống kê, và ngược lại. (Chẳng hạn, tham khảo LaLonde 1986, Fraker và Maynard 1987, LaLonde và Maynard 1987, và Friedlander và Robins 1995.) Có thể kiểm chứng sự thiên vị này bằng các kỹ thuật thống kê như các biến tương hợp và biến công cụ, nhưng thường là rất khó để có thể hoàn toàn loại bỏ chúng, và đây là một thách thức lớn đối với những nhà nghiên cứu trong lĩnh vực phân tích tác động.

Trong số các kỹ thuật thiết kế bán thử nghiệm, các kỹ thuật so sánh tương hợp thường được xem là sự lựa chọn thay thế cho thiết kế thử nghiệm. Đa số các công trình nghiên cứu về phương pháp đánh giá tập trung xung quanh việc sử dụng cách đánh giá này. Thực tế phản ánh cả tồn xuất sử dụng so sánh tương hợp và những thách thức lớn đặt ra do sử dụng các nhóm đối chứng thiếu hoàn hảo. Trong những năm gần đây, đã có những bước tiến lớn trong các kỹ thuật tương hợp theo mức độ xu hướng (propensity score matching techniques) (Rosenbaum và Rubin 1985; Jalan và Ravallion 1998). Phương pháp này rất hấp dẫn đối với những người đánh giá bị hạn chế về thời gian và không có dữ liệu gốc, do nó có thể được sử dụng với những dữ liệu chéo (cross-section) đơn giản. Tuy nhiên, kỹ thuật này phụ thuộc vào việc phải có các dữ liệu tốt, do nó phụ thuộc vào việc chọn mẫu quá nhiều (oversampling) với những người thụ hưởng trong khảo sát thực địa của một cuộc điều tra lớn, và sau đó "làm cho chúng phù hợp" với nhóm đối chứng được lựa chọn từ mẫu cơ sở lớn hơn, thường lấy từ trong điều tra hộ gia đình toàn quốc. Với sự gia tăng việc ứng dụng các cuộc điều tra lớn tại các nước đang phát triển, chẳng hạn như Các nghiên cứu đo lường mức sống phục vụ nhiều mục đích, phương pháp đánh giá này tỏ ra đặc biệt hứa hẹn. Một ví dụ quan trọng là việc đánh giá một chương

### Hộp 1.2 Tóm tắt các phương pháp định lượng nhằm đánh giá tác động của chương trình

Các phương pháp chủ yếu dùng để đánh giá tác động được thảo luận dưới đây. Do không có phương pháp nào là hoàn hảo, nên cần luôn luôn thận trọng với các phương pháp này.

#### *Các thiết kế thử nghiệm hay kiểm chứng ngẫu nhiên*

- *Lựa chọn ngẫu nhiên*, trong đó việc lựa chọn các nhóm được nghiên cứu và nhóm kiểm chứng được tiến hành một cách ngẫu nhiên trong số những tập hợp dân cư đã được định nghĩa rõ ràng. Trong trường hợp này, cần không có sự khác biệt nào (trong kỳ vọng) giữa hai nhóm, trừ việc nhóm được nghiên cứu có thể tiếp cận chương trình. (Có thể còn có khác biệt do sai lầm mẫu; quy mô nhóm được nghiên cứu và nhóm kiểm chứng càng lớn, thì sai lầm mẫu càng nhỏ.)

#### *Các thiết kế không qua thử nghiệm hay bán thử nghiệm*

- *Các phương pháp tương hợp hay tạo kiểm chứng*, trong đó người ta cố gắng lựa chọn một nhóm đối chứng tốt nhất tương tự nhóm được nghiên cứu từ một cuộc điều tra lớn. Hình thức tương hợp được sử dụng nhiều nhất là *phương pháp tương hợp mức độ xu hướng* (*propensity score matching*), trong đó nhóm đối chứng được làm cho phù hợp với nhóm được nghiên cứu trên cơ sở một tập hợp các đặc điểm quan sát được hoặc bằng cách sử dụng “mức độ xu hướng” (xác suất tham dự dự kiến với những đặc điểm quan sát cho trước); mức độ xu hướng càng gần, thì sự tương hợp càng cao. Một nhóm

### Hộp 1.2 (Tiếp)

đối chứng tốt, xuất phát từ môi trường kinh tế tương tự và được áp dụng cùng bảng câu hỏi của những người phỏng vấn cũng được đào tạo tương tự như đối với nhóm được nghiên cứu.

- Các phương pháp khác biệt kép hay khác biệt trong khác biệt, trong đó người ta so sánh một nhóm được nghiên cứu và nhóm đối chứng với nhau (khác biệt lần đầu) trước và sau khi thực hiện chương trình (khác biệt lần hai). Cần phải loại bỏ các đơn vị so sánh khi sử dụng các mức độ xu hướng và nếu chúng có giá trị vượt quá khoảng giá trị quan sát được của nhóm được nghiên cứu.
- Phương pháp biến số công cụ hay kiểm chứng thống kê, trong đó người ta sử dụng một hoặc nhiều biến có tác động tới sự tham dự nhưng không tác động tới kết quả khi đã có sự tham dự. Phương pháp này xác định những sai lệch ngoại sinh trong kết quả là do việc thực hiện chương trình mang lại, với sự thừa nhận sự có mặt của chúng không phải là ngẫu nhiên, mà là có mục đích. “Các biến số công cụ” trước tiên được sử dụng để dự đoán sự tham dự chương trình; sau đó người ta sẽ xem xét các chỉ số kết quả khác với các giá trị dự đoán như thế nào.
- So sánh phản thân, trong đó một cuộc điều tra cơ sở đối với những người tham dự được tiến hành trước khi thực hiện chương trình và một cuộc điều tra khác tiến hành sau khi thực hiện chương trình. Cuộc điều tra cơ sở cung cấp nhóm đối chứng, và tác động được đo lường bằng sự thay đổi các chỉ số kết quả trước và sau khi thực hiện chương trình.

trình công trình công cộng, TRABAJAR, ở Áchentina (Jalan và Ravallion 1998, Phụ lục 1.1, và chương 4).

### Các phương pháp định tính

Các kỹ thuật định tính cũng được sử dụng để đánh giá tác động với mục đích xác định tác động nhờ dựa trên một cái gì khác với phản chứng để đưa ra những kết luận nhân quả (Mohr 1995). Các phương pháp định tính tập trung vào việc hiểu các quá trình, các hành vi và điều kiện như được các cá nhân hay nhóm được nghiên cứu nhận thức (Valadez và Bamberger 1994). Chẳng hạn, các phương pháp định tính và đặc biệt là quan sát sự tham gia, có thể giúp chúng ta hiểu thấu đáo việc các hộ gia đình và cộng đồng địa phương nhận định về một dự án và bị tác động của dự án như thế nào. Do việc đo lường phản chứng là tâm điểm trong các kỹ thuật phân tích tác động, nên các thiết kế định tính thường được sử dụng trên cơ sở phối hợp với các kỹ thuật đánh giá khác. Cách tiếp cận định tính sử dụng các phương pháp tương đối mở trong quá trình thiết kế, thu thập số liệu, và phân tích. Cũng có thể định lượng hóa các dữ liệu định tính. Trong số các phương pháp được sử dụng trong đánh giá tác động định tính, có các kỹ thuật được xây dựng để đánh giá khu vực nông thôn một cách nhanh chóng- những kỹ thuật này phụ thuộc vào kiến thức của những người tham dự về các điều kiện xung quanh dự án hay chương trình đang được đánh giá- hay các đánh giá sự tham gia trong đó các bên có liên quan đều tham gia tất cả các giai đoạn trong quá trình đánh giá - từ việc xác định các mục tiêu của nghiên cứu, xác định và lựa chọn các chỉ số cần được tham dự, tới việc tham gia trong quá trình thu thập và phân tích dữ liệu. Tham khảo cuộc thảo luận chi tiết về các phương pháp tham gia trong Ngân hàng Thế giới (1996), *Sách gốc về Sự tham gia của Ngân hàng Thế giới*.

Lợi ích của các đánh giá định tính là chúng linh hoạt, có thể được điều chỉnh cụ thể cho phù hợp với các mục đích của đánh giá bằng cách sử dụng các phương pháp mở, có thể được tiến hành một cách nhanh chóng thông qua các kỹ thuật xử lý nhanh và có thể cung cấp mạnh mẽ các kết quả của sự đánh giá tác động nhờ tăng cường sự hiểu biết về nhận thức và các mối ưu tiên của các bên liên quan cũng như về các điều kiện và quá trình có thể tác động tới mức độ tác động của chương trình.

Trong số các nhược điểm của phương pháp đánh giá định tính có tính chủ quan trong thu thập dữ liệu, việc thiếu một nhóm đối chứng và thiếu tính vững chắc về thống kê, do quy mô mẫu chủ yếu là nhỏ, vì vậy rất khó có thể khái quát hoá các kết quả cho tổng thể đại diện lớn hơn. Giá trị và độ tin cậy của các dữ liệu định tính phụ thuộc chặt chẽ vào các kỹ năng phương pháp, sự nhạy cảm và quá trình đào tạo của những người đánh giá. Nếu các cán bộ thực địa không nhạy cảm với các chuẩn mực và thông lệ xã hội và văn hoá cụ thể, cũng như với các quy định bất thành văn, thì những dữ liệu thu thập được có thể bị diễn giải sai lầm. Và cuối cùng, do không có một nhóm đối chứng, nên sẽ không thể xác định được các trường hợp phản chứng, và do vậy, không nhận biết được quan hệ nhân quả trong tác động của dự án.

### **Kết hợp các phương pháp định lượng và định tính**

Mặc dù có rất nhiều tài liệu nghiên cứu về các phương pháp định lượng so với các phương pháp định tính trong đánh giá tác động, nhưng người ta ngày càng thấy cần phải kết hợp hai phương pháp này với nhau. Các đánh giá tác động sử dụng dữ liệu định lượng từ những mẫu có tính đại diện thống kê phù hợp hơn trong việc đánh giá mối quan hệ nhân quả bằng các phương pháp kinh tế lượng hay để khái quát hoá các kết luận. Tuy vậy, các phương pháp định lượng cho phép có sự nghiên cứu sâu sắc một số vấn đề, trường hợp, hay sự kiện có tính chọn lọc và có thể giúp cho việc hiểu thấu đáo quan điểm của

người thụ hưởng, tính năng động của một cuộc cải cách cụ thể, hoặc các lý do dằng sau những kết quả nhất định có được trong phân tích định lượng. Cũng có sự đánh đổi đáng kể trong việc lựa chọn một kỹ thuật này so với kỹ thuật kia.

Sự phối hợp các đánh giá định lượng và định tính có thể là một công cụ tốt nhất nhằm đáp ứng nhu cầu thông tin của các dự án. Bằng cách kết hợp hai phương pháp này, chúng ta có thể sử dụng các phương pháp định tính để đưa ra thông tin về những vấn đề đánh giá tác động quan trọng, điều tra bảng câu hỏi hay sự phân lớp của mẫu định lượng, và phân tích bối cảnh xã hội, kinh tế, và chính trị của dự án. Trong khi đó, có thể sử dụng các phương pháp định lượng để đưa ra thông tin về các chiến lược thu thập dữ liệu định tính, thiết kế mẫu nhằm cho biết về mức độ có thể khái quát hoá những kết quả quan sát được trong công trình định tính cho một tổng thể lớn hơn bằng cách sử dụng một mẫu có tính đại diện thống kê, và có thể sử dụng phân tích thống kê nhằm kiểm chứng các đặc điểm của hộ gia đình và các điều kiện kinh tế xã hội của các khu vực nghiên cứu khác nhau, nhờ đó, loại bỏ các cách giải thích khác về những kết quả quan sát được.

Một số lợi ích của việc sử dụng các phương pháp phối hợp với nhau trong nghiên cứu như đã được thảo luận trong Bamberger (2000) cũng có thể áp dụng trong đánh giá tác động. Một số lợi ích đó là:

- Có thể kiểm tra tính nhất quán nhờ sử dụng các quy trình tam giác (triangulation), trong đó cho phép đưa ra hai hay nhiều hơn các ước lượng độc lập cho những biến quan trọng (ví dụ thu nhập, các quan điểm về các dự án, các lý do để sử dụng hay không sử dụng dịch vụ công cộng, và tác động cụ thể của một dự án).
- Có thể có các góc nhìn khác nhau. Ví dụ, mặc dù các nhà nghiên cứu có thể coi thu nhập hay tiêu dùng là các chỉ số chính của phúc lợi hộ gia đình, nhưng các nghiên cứu tình

- huống có thể cho thấy phụ nữ quan tâm nhiều hơn tới mức độ dễ tổn thương (được định nghĩa là việc thiếu khả năng tiếp cận các hệ thống hỗ trợ xã hội trong các trường hợp khẩn hoảng), sự bất lực và rủi ro gấp phải bạo lực.
- Có thể tiến hành phân tích ở các mức độ khác nhau. Các phương pháp điều tra có thể đưa ra các ước tính tốt về phúc lợi của các cá nhân, hộ gia đình và cộng đồng, nhưng chúng kém hiệu quả hơn nhiều trong phân tích về các tiến trình xã hội (xung đột xã hội, các lý do sử dụng hay không sử dụng các dịch vụ, và v.v) hay trong phân tích thể chế (y tế, giáo dục, tín dụng và các dịch vụ khác hoạt động hiệu quả như thế nào, và cộng đồng nhận thức về chúng như thế nào). Có nhiều phương pháp định tính được thiết kế nhằm phân tích các vấn đề như tiến trình xã hội, hành vi thể chế, kết cấu xã hội và xung đột.
  - Nó có thể tạo ra các cơ hội thu thập ý kiến phản hồi để giúp diễn giải các kết quả. Các báo cáo điều tra thường bao gồm các phần tham chiếu cho những sự thiếu nhất quán trong kết quả hay cho những khác biệt lý thú giữa các cộng đồng hay nhóm người mà không thể giải thích bằng dữ liệu. Trong hầu hết các nghiên cứu định lượng, một khi đã hoàn tất giai đoạn thu thập dữ liệu, sẽ không thể quay trở lại thực địa để kiểm chứng những câu hỏi này. Mức độ linh hoạt cao hơn của các nghiên cứu định tính có nghĩa là, thường có thể quay trở lại thực địa để thu thập thêm số liệu. Các nhà nghiên cứu điều tra cũng sử dụng các phương pháp định tính để kiểm tra các hiện tượng ngoại vi (outliers)- tức là các phản ứng chêch ra khỏi xu hướng chung. Trong nhiều trường hợp, nhà phân tích dữ liệu phải đưa ra các quyết định tuỳ ý mình là liệu hộ gia đình hay cộng đồng có các điều kiện cao hơn hay thấp hơn đáng kể so với chuẩn mực có nên bị loại ra không (trên cơ sở giả định là nó phản ánh một sai sót trong báo cáo kết quả) hay cần phải điều chỉnh số liệu. Các phương pháp định tính cho

phép sự nghiên cứu tiếp theo nhanh chóng trên thực địa để kiểm chứng các trường hợp này.

Trên thực tế, sự kết hợp giữa các phương pháp định lượng và định tính cần được tiến hành trong từng bước của quá trình đánh giá tác động. Chương 2 đề cập tới nhiều cơ hội để thực hiện công việc này. Để minh họa, Trường hợp Cải cách quyền tự chủ trong trường học của Nicaragua là một ví dụ tốt về các phương pháp phối hợp. Các phương pháp định lượng được sử dụng sau khi có thiết kế bán thử nghiệm để xác định mối quan hệ giữa sự quản lý phi tập trung và việc học tập, và khai quát hoá các kết quả cho các loại hình trường học. Để bổ sung, các kỹ thuật định tính, bao gồm một loạt các cuộc phỏng vấn quan trọng có nhiều thông tin và các cuộc thảo luận nhóm tập trung với các cán bộ trường học và phụ huynh học sinh, cũng được sử dụng để phân tích bối cảnh đưa ra cải cách, thẩm tra tính năng động trong quá trình ra quyết định ở từng trường học, và đánh giá tầm nhìn của các tác nhân trong cộng đồng trường học khác nhau đối với quá trình tự chủ (xem Phụ lục 1.11).

### **Các cách tiếp cận khác đối với đánh giá tác động**

Có hai chủ đề khác đặc biệt liên quan tới thảo luận về việc đánh giá tác động của nghèo đói tới các dự án, đó là (a) các phương pháp đo lường tác động của các chương trình điều chỉnh cơ cấu, và (b) các đánh giá dựa trên lý thuyết. Hai chủ đề này bao gồm nhiều phương pháp đã được thảo luận ở trên, nhưng mỗi chủ đề lại sử dụng một cách tiếp cận khác nhau.

**Dánh giá các chương trình điều chỉnh cơ cấu.** Có sự tranh luận đáng kể về tác động của các chương trình điều chỉnh cơ cấu tới người nghèo. Tuy nhiên, phần lớn các bằng chứng được sử dụng cho tranh luận này đều dựa vào các giả định và phương pháp yếu kém. Cũng như với các dự án khác, các thay đổi chính sách trong các dự án điều chỉnh cơ cấu cần phải (a)

được so sánh với các trường hợp phản chứng phù hợp mà sẽ phản ứng trước cùng các ràng buộc kinh tế ví mô, và (b) được phân tích trong bối cảnh cấu trúc kinh tế địa phương và dựa trên những thông tin thử nghiệm có được từ các cuộc điều tra hộ gia đình. Tuy nhiên, điều này rất khó khăn, vì ba lý do. Thứ nhất, các thay đổi chính sách có thể gây ra tác động trên toàn nền kinh tế, khiến cho không thể tìm được các nhóm đối chứng không chịu tác động. Thứ hai, do các nhân tố ngoại sinh, độ trễ, các phản hồi và thay thế, nên cần diễn giải một cách cực kỳ thận trọng đối với bất kỳ thay đổi nào trong phúc lợi của người nghèo. Và thứ ba, khó có thể dự đoán điều gì sẽ xảy ra nếu như không tiến hành điều chỉnh cơ cấu - chính phủ sẽ tiến hành những chính sách thay thế nào và tác động khi đó tới người nghèo sẽ là gì.

Trong các tài liệu nghiên cứu, người ta sử dụng một số phương pháp, và mỗi phương pháp đều có những nhược điểm của nó. Các kỹ thuật này, trong nhiều trường hợp tương tự như những kỹ thuật đã được mô tả trong Hộp 1.2, mặc dù, như chỉ ra trong Hộp 1.3, việc ước tính cho trường hợp phản chứng đòi hỏi những giả định quá rộng, có thể tác động đáng kể tới tính hợp lý của các kết quả. Công việc này có thể được giải quyết một cách khả thi bằng cách tách biệt các thay đổi chính sách có thể tác động tới dân cư, chẳng hạn chính sách tỷ giá hối đoái, chính sách thương mại, giảm chi tiêu công cộng, và giảm nhân công ở khu vực công cộng. Tuy vậy, ngay cả với phương pháp này, cũng khó mà tách biệt được tác động của các chính sách cụ thể. Chẳng hạn, xem Killick (1995), Poppele, Summarto, và Pritchett (1999), Bourguignon, de Melo, và Suwa (1991), và Sahn, Dorosh, và Younger (1996).

**Đánh giá dựa trên lý thuyết.** Cơ sở của phương pháp đánh giá dựa trên cơ sở lý thuyết là các chương trình và dự án thường dựa trên những lý thuyết rõ ràng hay ngầm định về cách thức và lý do một chương trình hoạt động. Sự đánh giá khi đó sẽ dựa trên cơ sở đánh giá mỗi lý thuyết và các giả định

### Hộp 1.3. Tóm tắt các phương pháp được sử dụng để đánh giá các chính sách điều chỉnh

#### *Các phương pháp không có phản chứng*

- Các nghiên cứu định tính đánh giá các điều kiện của dân cư (thường xác định các nhóm nhỏ dễ tổn thương) trước, trong, và sau khi tiến hành các chính sách điều chỉnh thông qua các nhóm tập trung, phỏng vấn và các kỹ thuật định tính khác.
- Phương pháp “Trước và Sau,” bao gồm việc so sánh kết quả của những biến số quan trọng trong và sau khi thực hiện chương trình với trước khi thực hiện chương trình. Phương pháp này sử dụng các kỹ thuật thống kê để đánh giá xem liệu có sự thay đổi đáng kể trong một số biến số quan trọng theo thời gian hay không. Phương pháp này thường đưa đến các kết quả có sai lệch do nó giả định là, nếu như không có chương trình thì các chỉ số kết quả sẽ nhận được các giá trị như ở thời điểm trước khủng hoảng.

#### *Các phương pháp tạo ra phản chứng bằng cách sử dụng nhiều giả định*

- Các mô hình cân bằng chung có thể tính toán (CGE), cố gắng so sánh các kết quả trong nhóm được nghiên cứu với nhóm đối chứng thông qua mô phỏng. Các mô hình này tìm cách xem xét sự hoạt động của nền kinh tế thực và thường dựa trên các ma trận kết toán xã hội chi tiết được thu thập từ những dữ liệu trong các tài khoản quốc gia, điều tra chi tiêu của hộ gia đình, và các dữ liệu điều tra khác. Các mô hình CGE đưa ra các kết quả cho trường hợp phản chứng, mặc dù sức mạnh của mô

### Hộp 1.3 (Tiếp)

hình hoàn toàn phụ thuộc vào tính hợp lý của các giả định. Phương pháp này có thể có vấn đề do các cơ sở dữ liệu thường thiếu hoàn hảo và có nhiều tham số chưa được tính của các phương pháp kinh tế lượng chính thức. Các mô hình CGE cũng đòi hỏi nhiều thời gian, khá phiền phức và tốn kém khi tạo lập.

- Phương pháp có và không có so sánh. Phương pháp này so sánh hành vi của các biến quan trọng trong một mẫu gồm các nước tham gia chương trình với hành vi của các nước không tham gia chương trình (nhóm đối chứng). Đây là một cách tiếp cận đối với phản chứng, sử dụng kinh nghiệm của nhóm đối chứng để đại diện cho những gì có lẽ đã xảy ra ở các nước tham gia chương trình. Tuy nhiên, khó có thể có được một nhóm đối chứng thực sự. Phương pháp này giả định, việc áp dụng một chương trình điều chỉnh và các môi trường bên ngoài có tác động như nhau tới cả hai nhóm.
- Các kiểm soát thống kê, bao gồm các phương trình hồi quy trong đó đã kiểm soát những khác biệt về điều kiện ban đầu và các chính sách được tiến hành ở những nước có chương trình và không có chương trình. Phương pháp này xác định những khác biệt giữa các nước tham gia và không tham gia chương trình trong thời kỳ trước khi có chương trình, và sau đó kiểm soát những khác biệt này về mặt thống kê nhằm xác định các tác động tách biệt của chương trình tới các kết quả sau cải cách.

về chương trình này trong khi thực thi hơn là tại một thời điểm ở giữa hay sau khi đã hoàn thành dự án. Trong việc thiết kế đánh giá, lý thuyết cơ sở được trình bày thành nhiều bước nhỏ, và sau đó sẽ xây dựng các phương pháp thu thập dữ liệu và phân tích nhằm theo sát tính thực tế của các giả định này. Nếu

các sự kiện không diễn ra như dự kiến, thì việc đánh giá có thể khá tự tin khi chỉ ra sai sót đã diễn ra ở đâu, tại sao, và như thế nào.

Phương pháp này nhấn mạnh đến phản ứng của người dân trước các hoạt động của chương trình. Các lý thuyết định hướng sự quan tâm của người đánh giá tới những loại tác động có thể trong ngắn hạn và dài hạn. Một trong số các lợi ích của nó là, quá trình đánh giá đã cho các tín hiệu sớm về hiệu quả của chương trình trong khi thực hiện dự án. Vì vậy, nếu có những sai sót trong quá trình thực thi thì có thể sửa chữa ngay. Thứ hai, phương pháp này giúp giải thích các tác động diễn ra như thế nào và tại sao. Nếu các sự kiện diễn ra như dự kiến, thì quá trình đánh giá có thể nêu ra với một sự tin tưởng nhất định là các tác động đã được tạo ra như thế nào. Bằng cách thực hiện tuần tự các giai đoạn, có thể theo dõi được các bước nhỏ dẫn dắt các dấu vào của chương trình cho tới các kết quả.

Nhược điểm của phương pháp này cũng tương tự như của nhiều phương pháp khác. Đặc biệt, đó là (a) việc xác định các giả định và lý thuyết có thể rất phức tạp; (b) những người đánh giá có thể gặp khó khăn trong việc đo lường mỗi bước, trừ khi có được các công cụ và dữ liệu hợp lý, (c) có thể gặp phải những vấn đề khi kiểm tra nỗ lực do những phát biểu lý thuyết có thể quá chung chung và lỏng lẻo để cho phép sự kiểm nghiệm một cách rõ ràng, và (d) có thể có những vấn đề trong diễn giải khiến cho khó có thể khái quát hoá từ các kết quả này (xem Weiss 1998).

Một ví dụ về kỹ thuật dựa vào lý thuyết hiện đang được thử nghiệm của Phòng Hoạt động và Đánh giá của Ngân hàng Thế giới nhằm đánh giá tác động của các nguồn vốn đầu tư xã hội tới các tiến trình ra quyết định, cấu trúc và quan hệ quyền lực truyền thống ở cấp cộng đồng, và năng lực, niềm tin và phúc lợi trong cộng đồng. Sự đánh giá này sẽ dựa trên lý thuyết là, các nhóm ưu tiên có thể thực hiện một cách có hiệu quả một dự án, vận hành và duy trì sự đầu tư mà dự án này tạo ra. Một tập hợp các giả định chính và giả định phụ đã được xác lập và

sẽ được kiểm tra sử dụng các dữ liệu điều tra hộ gia đình hiện có, một công cụ điều tra được thiết kế đặc biệt cho một mẫu nhỏ hơn, các nhóm tập trung và các kỹ thuật PRA khác. Thông tin từ những nguồn dữ liệu này sẽ được xử lý theo qui trình tam giác (triangulation) trong phân tích.

### **Phân tích chi phí - lợi ích hay chi phí - hiệu quả**

Trong khi loại phân tích này không liên quan chặt chẽ với việc đo lường tác động, thì nó lại cho phép các nhà hoạch định chính sách đo lường hiệu quả của chương trình bằng cách so sánh những chương trình khác nhau trên cơ sở chi phí tạo ra cùng một mức đầu ra. Phương pháp này có thể tăng cường mạnh mẽ những gợi ý chính sách của phân tích tác động, và do đó, cần được đưa vào trong thiết kế của bất kỳ một sự đánh giá tác động nào. (*Sổ tay Phân tích kinh tế trong hoạt động đầu tư*, Ngân hàng Thế giới 1996 thảo luận đầy đủ hơn về phân tích chi phí - lợi ích và chi phí - hiệu quả).

Phân tích chi phí- lợi ích cố gắng đo lường hiệu quả kinh tế của chi phí của chương trình so với lợi ích của chương trình, theo giá trị tiền tệ. Đối với nhiều dự án, đặc biệt là trong các lĩnh vực xã hội, khó có thể đo lường được tất cả các lợi ích theo giá trị tiền tệ. Chẳng hạn, lợi ích của một chương trình cung cấp đầu vào cho nhà trường (giáo trình, trang bị lớp học, các chương trình tiền học đường) sẽ là tăng cường kết quả học hành. Thay vì đo lường các kết quả bằng tiền, có thể sử dụng các điểm thành tích học tập để định lượng hoá các lợi ích này. Điều này đòi hỏi cần có phân tích chi phí - hiệu quả. Các khái niệm của cả hai loại phân tích này đều giống nhau.

Các bước chính trong phân tích chi phí - lợi ích và chi phí - hiệu quả là xác định tất cả các chi phí và lợi ích của dự án, và sau đó tính toán hệ số chi phí trên hiệu quả. Trong tính toán chi phí, cần đưa vào các giá trị của bản thân chương trình cũng như tất cả các chi phí khác, chẳng hạn chi phí hành chính, cung cấp, đầu tư (được chiết khấu về giá trị hiện tại ròng), giá trị

bằng tiền của các hàng hoá hay dịch vụ được cung cấp miễn phí, các chi phí xã hội như suy thoái môi trường và rủi ro tới sức khoẻ. Các lợi ích có thể được đo lường bằng tiền, chẳng hạn như sự gia tăng thu nhập, hay số lượng các đơn vị được cung cấp, các điểm bài thi, hay sự cải thiện sức khoẻ. Khi không thể định lượng hoá các lợi ích được, có thể sử dụng các chỉ số chủ quan, như các hệ thống xếp hạng và sử dụng quyền số. Tuy nhiên, phương pháp này có thể gặp khó khăn trong việc giải thích các điểm có tính chủ quan.

Một khi đã xác định được các chi phí và lợi ích, chúng ta sẽ có hệ số chi phí - hiệu quả ( $R$ ) là  $R = \text{chi phí} / \text{đơn vị}$  (hay lợi ích). Hệ số này sau đó có thể được so sánh giữa các chương trình khác nhau để đo lường hiệu quả. Về lý thuyết, phương pháp này rất dễ hiểu. Tuy vậy, trên thực tế, có thể có nhiều vấn đề trong việc xác định và định lượng hoá các chi phí và lợi ích. Cần phải bảo đảm chắc chắn là chúng ta đã lựa chọn các chỉ số phù hợp, sử dụng các phương pháp và giả định kinh tế một cách nhất quán trong các hệ số, và các hệ số thực sự có thể so sánh được với nhau. Và, cũng như với các kỹ thuật khác được sử dụng trong phân tích tác động, việc đo lường chi phí - hiệu quả có thể được tiến hành một cách tốt nhất khi nó là một phần trong thiết kế đánh giá ngay từ những giai đoạn ban đầu. Điều này cho phép thu thập được thông tin về các chi phí và lợi ích cần thiết, và đảm bảo được tính nhất quán.

### Lựa chọn một phương pháp

Với sự đa dạng trong các kiểu dự án, câu hỏi đánh giá, mức độ sẵn có của thông tin, chi phí, hạn chế về thời gian, và tình hình của mỗi nước, các nghiên cứu đánh giá tác động sẽ khác nhau và đòi hỏi phải có một sự kết hợp các phương pháp phù hợp với nhau, cả định lượng và định tính. Những người đánh giá sẽ phải khảo sát tỉ mỉ các lựa chọn phương pháp khác nhau trong thiết kế nghiên cứu, với mục tiêu tạo ra những kết quả tốt nhất có thể. Trong số các phương pháp định lượng, các thiết kế thử

nghiệm được coi là phương pháp tối ưu và phương pháp so sánh tương hợp là giải pháp tốt thứ nhì. Tuy vậy, các kỹ thuật khác cũng có thể cho các kết quả đáng tin cậy, đặc biệt là khi có sự thiết kế đánh giá tốt và các dữ liệu có chất lượng cao.

Bằng chứng từ các đánh giá “tốt nhất” được cuốn sổ tay này xem xét nhấn mạnh là sự lựa chọn các phương pháp đánh giá tác động không có tính loại trừ nhau. Trên thực tế, những đánh giá tốt thường kết hợp các phương pháp để đảm bảo tính chính xác và phòng ngừa các bất trắc trong thực hiện. Kết hợp phương pháp “có và không có” với phương pháp “trước và sau”, trong đó sử dụng các dữ liệu ban đầu và tiếp theo là một sự kết hợp rất có ích, xét từ góc độ phương pháp luận (Subbarao và các tác giả khác 1999). Việc có những dữ liệu gốc sẵn có sẽ cho phép những người đánh giá xác nhận tính nguyên vẹn của nhóm được nghiên cứu và nhóm đối chứng, đánh giá mục tiêu, và chuẩn bị cho một sự đánh giá tác động vững chắc. Điều này đúng thậm chí với cả các thiết kế kiểm chứng ngẫu nhiên. Mặc dù việc lựa chọn ngẫu nhiên đảm bảo có các nhóm được nghiên cứu và nhóm đối chứng tương đương nhau vào thời điểm lựa chọn ngẫu nhiên, nhưng đặc tính này cần không được tác động tới những người đánh giá, khiến cho họ nghĩ rằng họ không cần phải có dữ liệu ban đầu. Trên thực tế, các dữ liệu ban đầu có thể rất cần thiết cho việc tái tạo lại lý do tại sao các sự kiện nhất định lại xảy ra và kiểm chứng các sự kiện này trong đánh giá tác động.

Sự kết hợp phân tích chi phí - lợi ích hay chi phí - hiệu quả cũng được khuyến nghị sử dụng mạnh mẽ. Phương pháp này có thể cho phép các nhà hoạch định chính sách so sánh các phương án khác nhau trên cơ sở chi phí tạo ra cùng một đầu vào. Điều này đặc biệt quan trọng trong bối cảnh các nước đang phát triển, là nơi mà các nguồn lực đặc biệt bị hạn chế.

Cuối cùng, việc kết hợp các phương pháp định lượng và định tính là lý tưởng do nó sẽ cho biết tác động lượng hoá của một dự án cũng như sự giải thích về các tiến trình và hành động đưa đến các kết quả này. Mặc dù mỗi đánh giá tác động đều

có những đặc điểm riêng, đòi hỏi phải có các phương pháp tiếp cận khác nhau, nhưng một đánh giá tác động tốt thường có một số đặc điểm sau:

- Ước lượng trường hợp phản chứng bằng cách (a) sử dụng sự phân công ngẫu nhiên để tạo ra một nhóm kiểm chứng (thiết kế thử nghiệm), và (b) sử dụng một cách phù hợp và cẩn trọng các phương pháp khác, như phương pháp tương hợp, nhằm tạo ra một nhóm đối chứng (thiết kế bán thử nghiệm).
- Có các dữ liệu hợp lý được thu thập trong giai đoạn ban đầu và kế tiếp (bao gồm một khung thời gian đủ để có được tác động từ chương trình) nhằm kiểm chứng các khác biệt trước và sau chương trình trong các đối tượng tham dự, và thiết lập các tác động chương trình.
- Các nhóm được nghiên cứu và nhóm đối chứng cần có quy mô đủ để thiết lập các kết luận thống kê với mức độ suy giảm (attrition) thấp nhất.
- Phân tích chi phí- lợi ích và chi phí - hiệu quả được sử dụng để đo lường hiệu suất của dự án.
- Các kỹ thuật định tính được xử lý theo quy trình tam giác các kết quả.

---

---

---

## Chương 2

# Các bước quan trọng trong thiết kế và thực hiện đánh giá tác động\*

Việc tiến hành một nghiên cứu đánh giá tác động có thể khó khăn và tốn kém, và có thể nảy sinh nhiều vấn đề trong thực hiện ở tất cả các bước trong quá trình đánh giá. Những khó khăn này làm nổi bật tầm quan trọng của việc có một nghiên cứu được thiết kế tốt, một nhóm công tác quyết tâm và có năng lực, và sự liên lạc tốt giữa các thành viên trong nhóm công tác. Việc đưa công việc đánh giá vào trong thiết kế của dự án ngay từ những bước đầu có thể sẽ tạo ra được các kết quả kịp thời, nhờ đó, có thể sử dụng các kết quả để tiến hành điều chỉnh đối với những thành phần cụ thể trong khi dự án đang được tiến hành.

Không phụ thuộc vào quy mô, loại chương trình hay phương pháp sử dụng cho việc đánh giá, cần phải thực hiện một số bước quan trọng như được liệt kê dưới đây (Hộp 2.1). Chương này thảo luận về các bước này, cũng như về các vấn đề có thể phát sinh trong quá trình thực hiện. Trình tự của các bước này là rất quan trọng, đặc biệt trong việc đảm bảo có sự thu thập các dữ liệu cần thiết trước khi tiến hành thực hiện dự án. Việc lập kế hoạch sớm tạo ra cơ hội để lựa chọn ngẫu nhiên, xây dựng các nhóm đối chứng tương hợp từ trước, thu thập các dữ liệu ban đầu, và xác định các lần điều tra tiếp theo để có thể được sử dụng trong phương pháp tương hợp mức độ xu hướng (propensity score matching method).

---

\* Chương này chủ yếu căn cứ vào một bài viết của Laura Rawlings, Các vấn đề thực hiện trong đánh giá tác động, Đã xử lý, 1999.

### Hộp 2.1 Các bước chủ yếu trong thiết kế và thực hiện đánh giá tác động

*Trong quá trình nhận diện và chuẩn bị dự án*

1. Xác định có tiến hành đánh giá hay không
2. Làm rõ các mục tiêu của đánh giá
3. Khảo sát mức độ sẵn có của dữ liệu
4. Thiết kế việc đánh giá
5. Thành lập nhóm đánh giá
6. Nếu dữ liệu đã được thu thập thì:
  - a. Thiết kế và lựa chọn mẫu
  - b. Xây dựng công cụ thu thập dữ liệu
  - c. Tuyển dụng và đào tạo cán bộ thực địa
  - d. Kiểm tra thí điểm
  - e. Thu thập dữ liệu
  - f. Quản lý và tiếp cận dữ liệu

*Trong khi thực hiện dự án*

7. Thu thập dữ liệu hiện có
8. Phân tích dữ liệu
9. Viết các kết quả và thảo luận chúng với các nhà hoạch định chính sách và các bên hữu quan khác
10. Kết hợp các kết quả trong thiết kế dự án.

Tất cả công việc thiết kế và thu thập dữ liệu ban đầu cần được tiến hành trong giai đoạn nhận diện và chuẩn bị dự án. Trường hợp lý tưởng là khi chúng ta có một số kết quả trong quá trình thực hiện dự án để có thể sử dụng được các kết quả này nhằm cải thiện thiết kế của dự án nếu cần thiết. Một ví dụ về cách thức kết hợp các kế hoạch đánh giá trong dự án ngay từ các giai đoạn đầu tiên được minh họa trong Dự án phát triển dinh dưỡng và trẻ em Uganda (xem chương 4).

## Xác định có tiến hành đánh giá hay không

Việc đầu tiên cần quyết định là có cần phải đánh giá tác động hay không. Như đã được thảo luận ở trên, đánh giá tác động khác với những đánh giá khác do tập trung vào việc đánh giá mối quan hệ nhân quả. Do sự phức tạp và chi phí cao trong việc tiến hành đánh giá tác động, nên cần phải đánh giá các chi phí và lợi ích có liên quan, và chú ý tới việc liệu một phương pháp khác có phù hợp hơn không, chẳng hạn như việc theo dõi các chỉ số kết quả quan trọng hay đánh giá tiến trình. (Không nên coi các phương pháp này như là các phương pháp thay thế cho đánh giá tác động; trên thực tế chúng thường đóng vai trò bổ sung quan trọng cho đánh giá tác động.) Và có lẽ các đầu vào quan trọng nhất đối với quyết định có nên tiến hành đánh giá hay không là sự hỗ trợ mạnh mẽ về mặt chính trị và tài chính.

Các nỗ lực và nguồn lực bổ sung cần thiết để tiến hành đánh giá tác động được huy động một cách tốt nhất khi dự án có tính sáng tạo, có thể được lặp lại, được phân bổ nguồn lực đáng kể, và có những cách thức tác động được xác định rõ ràng. Chẳng hạn, đánh giá tác động của Quỹ Đầu tư xã hội Bôlivia đáp ứng được các tiêu chuẩn này. Thứ nhất, mô hình quỹ xã hội mới ở Bôlivia được xem có tính sáng tạo và có thể được lặp lại; thứ hai, nguồn quỹ xã hội quản lý khoảng 25 phần trăm tổng đầu tư công cộng ở Bôlivia từ khi bắt đầu quá trình đánh giá; và thứ ba, các cách thức tác động được xác định rõ ràng thông qua danh mục các dự án phụ của quỹ xã hội này.

Cũng cần ưu tiên tiến hành đánh giá tác động nếu như dự án áp dụng một cách tiếp cận mới, ví dụ một chương trình thí điểm mà sau đó có thể sẽ được mở rộng căn cứ vào các kết quả đánh giá, hay chương trình Các khoản vay cho học tập và đổi mới của Ngân hàng Thế giới. Đây là lý do khiến chương trình cải cách quyền tự chủ trường học ở Nicaragua phù hợp với sự đánh giá tác động. Nghiên cứu đánh giá đi kèm với sự thử nghiệm của chính phủ về một mô hình quản lý trường học phi tập trung từ giai đoạn thí điểm giữa thập

kỷ 1990 cho tới khi mô hình này được mở rộng cho hầu như toàn bộ các trường trung học và khoảng một nửa các trường tiểu học hiện nay. Đánh giá được một nhóm quốc tế có phối hợp chặt chẽ với nhau, bao gồm các cán bộ nội địa từ đơn vị nghiên cứu và đánh giá trực thuộc Bộ Giáo dục và các cán bộ tại Văn phòng phối hợp Dự án giáo dục tiểu học của Ngân hàng Thế giới ở Managua quản lý. Sự tham dự của những người này đảm bảo đánh giá đã cung cấp thông tin cho những quyết định chính sách quan trọng về việc điều chỉnh và mở rộng dự án thí điểm này.

Một cân nhắc quan trọng khác là phải đảm bảo phát triển đầy đủ chương trình sẽ được đánh giá để chương trình này trở thành đối tượng đánh giá tác động. Các dự án thí điểm và các cải cách mới áp dụng thường dễ điều chỉnh về nội dung cũng như về việc chúng sẽ được thực hiện như thế nào, khi nào, và do ai. Các thay đổi này có thể làm giảm sự chặt chẽ trong công việc đánh giá, đặc biệt là với các thiết kế thử nghiệm và các dạng đánh giá tiềm tàng khác, dựa trên các dữ liệu ban đầu và kế tiếp của những nhóm tham dự chương trình và nhóm kiểm chứng đã được xác định rõ ràng. Trong trường hợp vẫn chưa xác định được các chính sách sẽ đánh giá, nên tránh việc sử dụng đánh giá tác động để có thể linh hoạt trong dự án.

Việc có được sự ủng hộ từ các nhà hoạch định chính sách và các nhà tài trợ cho một đánh giá tác động có thể rất khó khăn, nhưng là một điều kiện tiên quyết để có thể tiếp tục công việc. Cần phải thuyết phục những người này rằng, đánh giá là một công việc có ích, nó giúp cho việc trả lời những câu hỏi liên quan tới quyết định về việc cần sàng lọc, mở rộng hay giảm bớt dự án đã được đánh giá. Cũng cần thuyết phục những người này về sự hợp lý của thiết kế đánh giá, và do đó, là của các kết quả thu được, đặc biệt là khi các kết quả không có tính tích cực như dự kiến.

Tài trợ cho đánh giá tác động tiếp tục là một vấn đề khó khăn đối với cả các nhà quản lý chương trình và các đối tác khách hàng. Vấn đề tài trợ trở nên phức tạp do trên thực tế

thường khó có được các dữ liệu về chi phí đánh giá. Và có lẽ vấn đề khó khăn nhất phát sinh từ giá trị hàng hoá công cộng trong đánh giá là: nếu các kết quả của đánh giá sẽ được sử dụng để cung cấp thông tin cho các chính sách được áp dụng ở bên ngoài biên giới quốc gia của nước tiến hành đánh giá như thực tế thường diễn ra- thì tại sao một nước nào đó lại phải chịu đựng các chi phí cho việc đánh giá này? Trong số các nghiên cứu tình huống có thông tin về nguồn tài trợ, các thông tin có được cho thấy, các nước thường phải chịu phần lớn, nhưng không phải toàn bộ, chi phí đánh giá. Như sẽ được thảo luận đầy đủ hơn trong chương 4, nhiều trường hợp được xem xét cho thấy, việc thực hiện thành công việc đánh giá tác động đòi hỏi không chỉ sự cam kết về nguồn lực đáng kể từ phía các nước, khách hàng, mà còn cần sự tham dự của các cán bộ Ngân hàng Thế giới, các nhà nghiên cứu và tư vấn bên ngoài, và do vậy, cần phải có các nguồn lực ngoài những nguồn lực do quốc gia đó cung cấp.

### Làm rõ các mục tiêu của đánh giá

Một khi đã xác định đánh giá tác động là phù hợp và hợp lý, thì việc xác lập các mục tiêu rõ ràng và tạo ra sự nhất trí trong các vấn đề cốt lõi của đánh giá sắp tới có vai trò lớn đối với thành công của đánh giá. Các mục tiêu rõ ràng là rất cần thiết để nhận thức được các nhu cầu thông tin, xác định các chỉ số kết quả và tác động, và xây dựng một chiến lược đánh giá chắc chắn nhằm đưa ra câu trả lời cho những câu hỏi được nêu ra. Việc sử dụng một khung logarit hợp lý cho chúng ta một công cụ tốt, khung này thường được sử dụng để xác định các mục tiêu của dự án và các nhu cầu thông tin làm căn cứ cho đánh giá.

Khung logarit hiện nay đang được sử dụng ngày càng nhiều tại Ngân hàng Thế giới, dựa trên một ma trận bốn nhân bốn. Khung này so sánh giữa thông tin về các mục tiêu dự án với việc theo dõi đầu ra của dự án bằng cách sử dụng các điểm mốc

và kế hoạch công việc như thế nào, các kết quả của dự án sẽ có tác động gì tới thể chế thụ hưởng hay hệ thống, và các tác động này sẽ được đo lường như thế nào, và đầu vào được sử dụng như thế nào để tạo ra đầu ra (xem các ví dụ trong Phụ lục 5). Nói cách khác, khung này giả định rằng, tác động dự kiến của dự án là hàm số của đầu ra dự án và một loạt các nhân tố khác. Đầu ra của dự án lại là hàm số của đầu vào dự án và các nhân tố bên ngoài dự án. Sau đó, cần phải xác định rõ các thước đo có thể lượng hoá đối với mỗi liên kết trong vòng đời dự án. Phương pháp này không loại trừ việc người đánh giá xem xét các tác động không dự kiến trước của dự án, nhưng giúp đảm bảo các mục tiêu của đánh giá rõ ràng và có tính tập trung. Các kỹ thuật định tính cũng có ích để làm rõ các mục tiêu của đánh giá và các chỉ số tác động như là các kết quả của đánh giá.

Mặc dù phát biểu về mục tiêu là một trong những phần dễ dàng nhất của tiến trình đánh giá, nhưng nó có thể trở nên đặc biệt khó khăn. Chẳng hạn, các phát biểu quá rộng sẽ không cho phép đánh giá được. Phát biểu mục tiêu đánh giá của chương trình PROBECAT ở Mêhicô (Phụ lục 1.9) là “tác động của chương trình đào tạo PROBECAT tới các kết quả trên thị trường lao động” sẽ trở nên chính xác hơn nếu được thu hẹp còn tác động của PROBECAT tới số giờ làm việc, thu nhập mỗi giờ, tiền lương hàng tháng, và thời điểm có công việc đầu tiên đối với các dạng công nhân khác nhau. Đánh giá PROGRESA của Mêhicô là một ví dụ tốt về mục tiêu đánh giá, có một đề cương rõ ràng, và việc mô tả nhiều mục tiêu từ đầu với các chi tiết riêng cho từng mục tiêu tại các khoản mục phụ (Phụ lục 1.10). Phát biểu về mục tiêu đặc biệt quan trọng, bởi vì công việc đánh giá rất phức tạp, nó phải xác định không chỉ tác động của dự án mà còn cả các khía cạnh về đối tượng và thời điểm của hoạt động chương trình.

Việc xem xét các thành phần đánh giá khác, như chi phí - hiệu quả, hay đánh giá quá trình, có thể cũng là các mục tiêu quan trọng của một nghiên cứu và có thể bổ sung cho đánh giá

tác động. Chi phí - hiệu quả có thể là mối quan tâm đặc biệt của các nhà hoạch định chính sách, những người phải ra các quyết định thu nhỏ, mở rộng, hay cải cách chương trình đang được đánh giá. Trong các vấn đề liên quan tới việc cung cấp dịch vụ, đánh giá tiến trình có thể phù hợp để đánh giá các thủ tục, sự năng động, các chuẩn mực, và các hạn chế liên quan tới một chương trình cụ thể được tiến hành.

### **Khảo sát mức độ sẵn có của dữ liệu**

Có thể sử dụng nhiều kiểu dữ liệu để tiến hành các nghiên cứu đánh giá tác động. Các dữ liệu này bao gồm dữ liệu từ các điều tra chéo hay điều tra theo nhóm mẫu (panel), tới các cuộc phỏng vấn kết thúc mở (open-ended). Điều lý tưởng là có các thông tin này ở cấp độ cá nhân để bảo đảm là có thể đánh giá được tác động thực sự. Thông tin cấp hộ gia đình có thể không cho thấy sự phân bổ nguồn lực bên trong hộ gia đình, sự phân bổ này có tác động tới phụ nữ và trẻ em do những người này thường có khả năng tiếp cận hạn chế hơn đến các nguồn lực sản xuất của hộ gia đình. Trong nhiều trường hợp, đánh giá tác động sẽ tận dụng một số hình thức dữ liệu sẵn có hay dựa vào một cuộc điều tra đang được tiến hành, nhờ đó, có thể tiết kiệm đáng kể chi phí. Tuy nhiên, với cách tiếp cận này, có thể phát sinh những vấn đề liên quan tới thời điểm thu thập dữ liệu và tính linh hoạt trong thiết kế bảng câu hỏi. Hộp 2.2 nêu bật một số điểm quan trọng cần ghi nhớ trong việc khảo sát các nguồn dữ liệu hiện có nhằm mục đích đánh giá tác động.

Với sự sáng tạo nhất định, có thể tối đa hoá những nguồn thông tin sẵn có. Một ví dụ tốt là việc đánh giá về Quỹ đầu tư xã hội Óndurat (xem chương 4). Nghiên cứu này sử dụng một môđun (module) lấy ra từ cuộc điều tra về thu nhập quốc dân và chi tiêu trong bảng câu hỏi của quỹ xã hội, nhờ đó cho phép so sánh thu nhập của những người thụ hưởng quỹ xã hội với các chỉ số toàn quốc để đánh giá mức độ định hướng tới người

### Hộp 2.2 Các điểm chủ yếu để xác định các nguồn dữ liệu cho đánh giá tác động

- Biết rõ về chương trình. Tiến hành đánh giá mà không biết rõ về các chi tiết hành chính và thể chế của chương trình là mạo hiểm; các thông tin này thường có được từ những người điều hành chương trình.
- Thu thập thông tin về “các sự kiện điển hình” có liên quan trong bối cảnh của chương trình. Các sự kiện liên quan có thể bao gồm bản đồ nghèo đói, phương thức hoạt động của thị trường lao động, sự phân chia dân tộc chủ yếu và các chương trình công cộng có liên quan.
- Cần chiết trung (eclectic) đối với các dữ liệu. Các nguồn dữ liệu có thể bao gồm các cuộc phỏng vấn không chính thức, phi cấu trúc, đối với những người tham dự chương trình và các dữ liệu định lượng từ các mẫu đại diện. Tuy nhiên, sẽ rất khó khăn khi đặt ra các câu hỏi phản chứng trong các cuộc phỏng vấn hay với các nhóm tập trung, chẳng hạn như khi cố gắng đặt câu hỏi với một người hiện đang tham dự trong một chương trình công cộng là: “Bạn đang làm gì vào thời điểm này nếu như không có chương trình này”. Việc nói chuyện với những người tham dự chương trình có thể rất có giá trị nhưng riêng mình nó thì khó có thể trở thành một đánh giá đáng tin cậy.
- Đảm bảo có dữ liệu về các chỉ số kết quả và các biến số giải thích phù hợp. Các biến số giải thích phù hợp cần khắc phục được vấn đề hiệp phương sai khác biệt trong các kết quả với điều kiện tham dự chương trình. Các kết quả có thể khác biệt phụ thuộc vào nhiều lý do, chẳng hạn việc một người nào đó có được đào tạo hay không. Chúng ta có thể không thấy được tác động của một chương trình trừ khi kiểm soát được hiện tượng hiệp

(Xem tiếp trang sau)

### Hộp 2.2 (Tiếp)

phương sai khác biệt này.

- Tuỳ thuộc vào các phương pháp được sử dụng, có thể cần đến các dữ liệu về những biến có tác động tới sự tham dự nhưng không ảnh hưởng tới các kết quả tương ứng với sự tham dự nhất định. Các biến số công cụ này có thể rất giá trị trong việc sàng lọc các tác động nhân quả có thể của các chương trình phi ngẫu nhiên (hộp 1.2).
- Dữ liệu về các kết quả và các biến số giải thích phù hợp khác có thể là định lượng hoặc định tính. Nhưng cần phải tổ chức thông tin dưới một cấu trúc dữ liệu có hệ thống nhất định. Một ví dụ đơn giản và thông dụng là trường hợp dữ liệu bao gồm giá trị của những biến khác nhau, kể cả một hay nhiều chỉ số kết quả cho những đơn vị quan sát khác nhau (các cá nhân, hộ gia đình, doanh nghiệp, cộng đồng).
- Các biến số mà chúng ta có dữ liệu và những đơn vị quan sát được sử dụng cũng thường được lựa chọn như là một phần trong phương pháp đánh giá. Các lựa chọn này cần phải dựa trên những hiểu biết từ trước về chương trình (tất nhiên là về mục tiêu của chương trình) nhưng cũng cả về cách thức vận hành của chương trình và bối cảnh trong đó chương trình này được đưa ra.
- Nguồn dữ liệu cụ thể về các kết quả và những nhân tố quyết định của chúng, kể cả sự tham dự chương trình, thường có được từ những dữ liệu điều tra nhất định. Đơn vị quan sát có thể là hộ gia đình, doanh nghiệp hay khu vực địa lý, tuỳ thuộc vào kiểu chương trình được nghiên cứu.
- Dữ liệu điều tra có thể được bổ sung bằng các dữ liệu có ích khác về chương trình (chẳng hạn từ cơ sở dữ liệu theo dõi dự án) hay từ bối cảnh (chẳng hạn từ cơ sở dữ liệu địa lý).

nghèo (Walker và các tác giả khác 1999).

Ở cấp độ cơ bản nhất, cần phải có dữ liệu tổng thể về nhóm dân cư được quan tâm, coi đó là một cơ sở để từ đó xác định quy mô mẫu, xây dựng khuôn khổ lấy mẫu, và lựa chọn mẫu. Các loại dữ liệu khác có thể sẵn có ở một nước nhất định và có thể được sử dụng trong các đánh giá tác động khác nhau, bao gồm (xem Valadez và Bamberger 1994): điều tra về thu nhập và chi tiêu của hộ gia đình; nghiên cứu đo lường mức sống dân cư (LSMS); điều tra thị trường lao động; sổ sách của các hợp tác xã, liên hiệp tín dụng, và các thẻ chế tài chính khác; sổ sách của nhà trường về tỷ lệ nhập học, lưu ban, và kết quả thi cử; sổ sách của cơ quan y tế công cộng về tỷ lệ tử vong trẻ sơ sinh, tỷ lệ mắc các căn bệnh lây nhiễm khác nhau, số phụ nữ tìm kiếm tư vấn về các biện pháp phòng tránh thai, hay mức độ tiêu thụ bao cao su; điều tra đặc biệt do các trường đại học, các tổ chức phi chính phủ (NGO), và các nhóm tư vấn tiến hành; dữ liệu giám sát từ những người điều hành chương trình; và các nghiên cứu tình huống thực tế của dự án.

**Sử dụng các dữ liệu điều tra sẵn có.** Nhiều cuộc điều tra cũng có thể ở trong giai đoạn lập kế hoạch hay đang được tiến hành. Nếu một điều tra đo lường các chỉ số cần thiết đang ở trong giai đoạn lập kế hoạch, thì để đánh giá, có thể chọn mẫu nhiều hơn (oversample) đối với nhóm dân cư được quan tâm từ một cuộc điều tra chung (chẳng hạn để sử dụng trong phương pháp tương hợp mức độ xu hướng) như đã được thực hiện trong đánh giá Quỹ đầu tư xã hội Nicaragua và đánh giá chương trình phúc lợi việc làm TRABAJAR của Áchentina (Jalan và Ravallion 1998). Trái lại, nếu một điều tra dự định sẽ bao gồm nhóm dân cư được quan tâm, thì đánh giá có thể đưa ra một câu hỏi hay một loạt các câu hỏi như là một phần trong điều tra hoặc là thêm vào một điều tra định tính để bổ sung cho các thông tin định lượng. Chẳng hạn, chương trình Tín dụng với Giáo dục ở Gana gồm có một tập hợp các phỏng vấn định lượng với những bên hữu quan trọng, cũng như với các nhóm tập trung

tham dự và không tham dự chương trình, nhờ đó, củng cố các bằng chứng định tính cho những kết quả định lượng (Phụ lục 1.6). Thẩm định này đánh giá tác động của chương trình tới tình trạng dinh dưỡng và an ninh thực phẩm của các hộ gia đình nghèo. Các dữ liệu định lượng bao gồm các câu hỏi cụ thể về thu nhập và chi tiêu của hộ gia đình và trình độ kỹ năng, trong khi đó các dữ liệu định tính tập trung vào việc trao quyền cho phụ nữ- như các vấn đề về địa vị và quyền ra quyết định trong gia đình, các mạng lưới quan hệ xã hội, sự tự tin, và những vấn đề đại loại như vậy.

### Thiết kế việc đánh giá

Một khi các mục tiêu và nguồn dữ liệu đã rõ ràng, có thể bắt đầu giai đoạn thiết kế trong nghiên cứu đánh giá tác động. Sự lựa chọn phương pháp sẽ phụ thuộc vào câu hỏi đánh giá, thời điểm, giới hạn ngân sách, và năng lực thực thi. Những ưu điểm và nhược điểm trong từng loại thiết kế đánh giá khác nhau thảo luận trong chương 1 cần được cân bằng với nhau để quyết định phương pháp nào là phù hợp nhất và có thể kết hợp các kỹ thuật định lượng và định tính như thế nào để có thể bổ sung cho nhau.

Ngay cả sau khi đã quyết định thiết kế đánh giá và đưa thiết kế này vào trong dự án, thì những người đánh giá cũng cần phải linh hoạt và có những điều chỉnh trong thiết kế khi thực hiện dự án. Hơn nữa, cũng cần có những điều khoản để theo dõi việc thực hiện dự án nếu như đánh giá bao gồm cả các dữ liệu gốc và dữ liệu tiếp theo để cho công việc đánh giá được tiến hành song song với tiến độ thực sự của dự án.

Trong định nghĩa thiết kế, cũng cần phải quyết định liệu đánh giá tác động có phù hợp với việc giám sát rộng rãi và chiến lược đánh giá được áp dụng cho dự án. Tất cả các dự án phải được giám sát để những người điều hành, người cho vay, và các nhà hoạch định chính sách có thể theo dõi được dự án khi nó được tiến hành. Công việc đánh giá, như được nêu ra ở

trên, cần phải được điều chỉnh cho phù hợp với yêu cầu thông tin của dự án.

**Câu hỏi đánh giá.** Các câu hỏi đánh giá được đặt ra có liên hệ chặt chẽ với thiết kế đánh giá về kiểu dữ liệu được thu thập, đơn vị phân tích, phương pháp sử dụng, và thời điểm của từng giai đoạn khác nhau. Ví dụ, trong đánh giá tác động của giáo trình tới kết quả học tập, cần phải điều chỉnh cách đánh giá cho phù hợp để đo lường tác động tới học sinh, lớp học và giáo viên trong một năm học nhất định. Điều này sẽ rất khác với việc đo lường tác động của các dịch vụ được cung cấp thông qua các khoản đầu tư của quỹ xã hội, trong trường hợp này sẽ cần phải có dữ liệu về trang thiết bị của cộng đồng và dữ liệu về hộ gia đình. Các nghiên cứu tình huống trong Phụ lục 1 cho thấy, các ví dụ khác về câu hỏi đánh giá có thể tác động tới thiết kế đánh giá như thế nào.

Khi làm rõ các câu hỏi đánh giá, cần phải xem xét những ngụ ý về giới của tác động của dự án. Mặc dù vậy, ở giai đoạn đầu, điều này không phải luôn luôn hiển nhiên; khi thực hiện dự án, có thể có những tác động phụ tới hộ gia đình và không phải luôn luôn có thể nắm bắt được những tác động này nếu như không có sự thu thập dữ liệu cụ thể và những nỗ lực phân tích.

**Các mối lo ngại về thời điểm và ngân sách.** Vấn đề thời điểm quan trọng nhất là liệu có thể bắt đầu thiết kế đánh giá trước khi dự án được thực hiện hay không, và khi nào thì cần đến các kết quả. Cũng cần xác định trước là tại điểm nào trong vòng đời dự án, chúng ta sẽ cần đến các thông tin từ đánh giá, nhờ đó có thể tạo ra sự liên kết giữa công việc thu thập dữ liệu và các hoạt động phân tích. Việc có các kết quả đúng lúc là rất quan trọng đối với các quyết định chính sách- ví dụ, trong khi xem xét lại dự án, xung quanh một thời kỳ bầu cử, hay khi đưa ra các quyết định liên quan tới việc tiếp tục dự án.

Một số phương pháp đòi hỏi phải có nhiều thời gian để thực

thi hơn các phương pháp khác. Việc phân công ngẫu nhiên và các phương pháp trước - và - sau (ví dụ, phương pháp so sánh phản thân) đòi hỏi nhiều thời gian thực thi hơn các phương pháp so sánh tương hợp. Khi sử dụng các phương pháp trước-và-sau, sử dụng các đánh giá cơ sở và tiếp theo, cần phải có thời gian đủ để cho thành viên cuối cùng trong nhóm tham dự chương trình nhận được tác động từ chương trình, và sau đó cần có nhiều thời gian hơn để các tác động sau khi có chương trình trở thành hiện thực và được quan sát. Grossman (1994) cho rằng, thời gian thông thường cần có trước khi kiểm tra các tác động là từ 12 đến 18 tháng sau khi chọn mẫu trong chương trình. Trong các dự án của Ngân hàng Thế giới với các dữ liệu gốc, khoảng thời gian để nhận được tác động của chương trình và hiện thực hóa các kết quả có thể mất hàng năm. Chẳng hạn, trong đánh giá của Quỹ Đầu tư xã hội Bolivia, dựa trên các dữ liệu gốc thu thập trong năm 1993, các dữ liệu tiếp theo không được thu thập cho tới năm 1998, vì cần có khoảng thời gian đủ dài để cho các chương trình (các dự án nước và vệ sinh, trạm xá y tế, và trường học) được thực hiện, và để những tác động tới sức khoẻ và giáo dục của nhóm dân cư thụ hưởng diễn ra. Việc đánh giá một dự án giáo dục tiểu học ở Pakixtan trong đó sử dụng thiết kế thử nghiệm (experimental design) với các điều tra cơ sở và tiếp theo nhằm đánh giá tác động của các trường học cộng đồng tới các kết quả của học sinh, kể cả thành tích học tập, cũng cần đến một khoảng thời gian tương tự.

Các yêu cầu về thời gian của đánh giá không thể có tác dụng bắt buộc đối với dự án được đánh giá. Do đặc tính của mình, các đánh giá cần căn cứ vào khung thời gian được xác lập do những phần còn lại của dự án. Đánh giá sẽ được tiến hành chậm hơn ở các dự án chậm chạp trong giải ngân và tạo ra tác động. Và ngay cả khi được tiến triển với tốc độ đã định, thì một số dự án có thể đòi hỏi nhiều thời gian hơn, ví dụ như các dự án cơ sở hạ tầng. Khung thời gian cho đánh giá cũng nhạy cảm với các chỉ số được lựa chọn do nhiều chỉ số, ví dụ sự thay đổi tỷ lệ sinh hay thành tích học tập, đòi hỏi

phải có nhiều thời gian hơn để thể hiện trong nhóm dân cư thụ hưởng.

**Năng lực thực hiện.** Sự cân nhắc cuối cùng về quy mô và sự phức tạp của thiết kế đánh giá là năng lực thực hiện của nhóm đánh giá. Các vấn đề thực hiện có thể rất khó khăn, đặc biệt ở các nước đang phát triển, là nơi có ít kinh nghiệm về nghiên cứu ứng dụng và đánh giá chương trình. Thành phần của nhóm đánh giá rất quan trọng, cũng như kinh nghiệm của các thành viên trong nhóm về các loại phương pháp khác nhau và năng lực của họ một cách tương đối so với các hoạt động khác mà đơn vị đánh giá tiến hành. Điều này đặc biệt quan trọng khi phải làm việc với các cơ quan nhà nước có nhiều trách nhiệm và ít nhân viên. Nhận thức về khối lượng công việc của đơn vị đóng vai trò quan trọng để đánh giá, không chỉ tác động của chúng tới chất lượng đánh giá, mà còn để biết được chi phí cơ hội của việc đánh giá trong quan hệ với các công việc khác mà đơn vị phụ trách. Có một số ví dụ về các công việc đánh giá bị sai lệch khi các cán bộ quan trọng bị yêu cầu tham gia vào các dự án khác, và do đó, không thể thu thập dữ liệu đúng kế hoạch tại những mốc thời gian quan trọng (chẳng hạn vào một thời điểm trong năm học, hay trong mùa vụ nông nghiệp). Có thể tránh khỏi những tình huống này bằng cách phối hợp với các nhà quản trị trong đơn vị phụ trách đánh giá để đảm bảo đạt được cân bằng trong thời điểm của các hoạt động khác nhau và trong sự phân công các cán bộ và nguồn lực cho những hoạt động này. Mặc khác, thuê một doanh nghiệp tư nhân tiến hành đánh giá có thể sẽ là một giải pháp tốt hơn (vấn đề này sẽ được thảo luận dưới đây).

### Thành lập nhóm đánh giá

Cần có một số các kỹ năng trong đánh giá. Chất lượng và lợi ích cuối cùng của đánh giá tác động có thể được tăng cường

mạnh mẽ nhờ sự phối hợp giữa các thành viên trong nhóm và các nhà hoạch định chính sách ngay từ giai đoạn đầu. Do vậy, cần xác định càng sớm càng tốt các thành viên trong nhóm đánh giá, nhất trí về vai trò và nhiệm vụ của các thành viên này, và thiết lập các cơ chế liên lạc trong các điểm mấu chốt của đánh giá.

Trong nhóm cốt cán phụ trách đánh giá có nhà quản lý đánh giá, các nhà phân tích (cả các nhà kinh tế và các nhà khoa học xã hội khác), và đối với thiết kế đánh giá đòi hỏi phải thu thập dữ liệu mới, thì cần có một chuyên gia chọn mẫu, chuyên gia thiết kế điều tra, người quản lý thực địa và nhóm điều tra thực địa, những người quản lý và xử lý dữ liệu (xem hướng dẫn đầy đủ cho thiết kế và thực hiện điều tra trong Grossh và Munoz 1996). Tùy thuộc vào quy mô, phạm vi, và thiết kế của nghiên cứu, một số trách nhiệm này sẽ được chia sẻ trong nhóm hoặc là bổ sung thêm nhu cầu cán bộ vào nhóm cốt cán này. Trong trường hợp những người phân tích chính sách không có kinh nghiệm phối hợp các phương pháp định lượng và định tính, thì cần phải sử dụng thời gian bổ sung ở giai đoạn xây dựng nhóm ban đầu để tăng cường độ nhạy bén và đảm bảo sự hợp tác đầy đủ của các thành viên trong nhóm. Trách nhiệm chung của các thành viên trong nhóm bao gồm:

- Người quản lý đánh giá - Người quản lý đánh giá chịu trách nhiệm xác lập các nhu cầu thông tin và các chỉ số để đánh giá (thường được xác lập cùng với khách hàng bằng cách sử dụng phương pháp khung logic), dự thảo các điều khoản tham chiếu (terms of reference) cho đánh giá, lựa chọn phương pháp đánh giá, và xác định nhóm đánh giá. Trong nhiều trường hợp, người quản lý đánh giá cũng sẽ thực hiện phân tích chính sách.
- Nhà phân tích chính sách - Cần có một nhà kinh tế để phân tích định lượng, cũng như một nhà xã hội học hoặc nhân chủng học để đảm bảo có đầu vào tham dự và phân

tích định tính ở các giai đoạn khác nhau trong đánh giá tác động. Cả hai người này đều tham gia viết báo cáo đánh giá.

- Chuyên gia chọn mẫu- Chuyên gia chọn mẫu dẫn dắt quá trình lựa chọn mẫu. Đối với các dữ liệu định lượng, chuyên gia chọn mẫu tiến hành tính toán để xác định các quy mô mẫu phù hợp tương ứng với các chỉ số đã được xác lập, lựa chọn mẫu, và đưa các trọng số của mẫu vào trong phân tích. Đối với các dữ liệu định tính, chuyên gia chọn mẫu cần dẫn dắt quá trình chọn mẫu với sự phối hợp của các nhà phân tích, đảm bảo các thủ tục được thiết lập sẽ lựa chọn được những nguồn thông tin đúng đắn. Chuyên gia chọn mẫu cũng cần chịu trách nhiệm lựa chọn các khu vực và các nhóm tiến hành kiểm tra thử và thường cần phải kết hợp với một người hợp tác cung cấp thông tin ở địa phương, người này chịu trách nhiệm thu thập cho chuyên gia chọn mẫu các dữ liệu để từ đó rút ra mẫu.
- Người thiết kế điều tra - Có thể là một người hay một nhóm, trách nhiệm của người thiết kế điều tra là thiết kế các công cụ thu thập dữ liệu, các tài liệu và sổ sách kèm theo, và phối hợp với (những) người quản lý đánh giá để đảm bảo các công cụ thu thập dữ liệu sẽ thực sự cung cấp các dữ liệu cần thiết cho phân tích. Cá nhân hay nhóm này cũng cần tham gia vào kiểm tra thử và sàng lọc bảng câu hỏi.
- Người quản lý và các cán bộ thực địa - Người quản lý chịu trách nhiệm giám sát toàn bộ quá trình thu thập dữ liệu, từ việc lập kế hoạch thu thập dữ liệu, tới việc thành lập và lên chương trình làm việc cho các nhóm công tác thực địa, thường bao gồm các giám sát viên và các cán bộ phỏng vấn. Các giám sát viên thường quản lý các cán bộ thực địa (thường là các cán bộ phỏng vấn, những người nhập số liệu và các lái xe) và chịu trách nhiệm đối với chất lượng của các dữ liệu thu thập được tại thực địa. Các cán bộ phỏng vấn

phụ trách bằng câu hỏi điều tra. Ở một số nền văn hoá, cần phải đảm bảo các cán bộ phỏng vấn nam và nữ tiến hành điều tra và họ được điều hành một cách tách biệt.

- **Những người quản lý và xử lý dữ liệu-** Các thành viên này trong nhóm thiết kế các chương trình nhập dữ liệu, tiến hành nhập dữ liệu, kiểm tra tính hợp lệ của dữ liệu, cung cấp các tài liệu về dữ liệu cần thiết, và đưa ra các kết quả cơ bản mà sau đó được các nhà phân tích dữ liệu kiểm chứng.

Trong việc xây dựng nhóm đánh giá, cũng có một số quyết định quan trọng mà người quản lý đánh giá cần phải đưa ra về năng lực của địa phương và các thể chế phù hợp nhằm đảm bảo tính vô tư và chất lượng của các kết quả đánh giá. Vấn đề thứ nhất là năng lực của địa phương trong việc thực hiện đánh giá, hoặc là các bộ phận trong việc này như thế nào, và sẽ cần có sự giám sát và hỗ trợ nào từ bên ngoài. Năng lực đánh giá khác nhau rất lớn giữa các nước, và mặc dù các hợp đồng quốc tế cho phép các doanh nghiệp ở một nước có thể tiến hành đánh giá ở một nước khác, đang ngày càng trở nên phổ biến (một ví dụ là đánh giá Tiến độ đang được Viện Nghiên cứu chính sách và Lương thực Quốc tế thực hiện), nhưng thực tế nói chung cho thấy, trong các dự án được Ngân hàng Thế giới hỗ trợ, các nhân viên trong nước tiến hành đánh giá với sự giám sát quốc tế là đáng kể. Do vậy, cần phải đánh giá thích đáng năng lực trong nước và quyết định xem ai sẽ chịu trách nhiệm về những khía cạnh nào trong công việc đánh giá. Cho dù thành phần của nhóm cuối cùng là gì đi nữa thì điều quan trọng vẫn là cần phải bổ nhiệm một người quản lý đánh giá có thể làm việc có hiệu quả với cả những người cung cấp dữ liệu cũng như với những nhà phân tích và nhà hoạch định chính sách sẽ sử dụng dữ liệu và các kết quả của đánh giá này. Nếu người này không sống ở trong nước, thì cần bổ nhiệm một nhà quản lý ở trong nước để điều phối công việc đánh giá trên cơ sở phối hợp với nhà quản lý quốc tế.

Vấn đề thứ hai là cần xác định làm việc với một doanh

nghiệp tư nhân hay một cơ quan nhà nước. Các công ty tư nhân có thể đáng tin cậy hơn trong việc cung cấp các kết quả đúng kỳ hạn, nhưng việc lựa chọn công ty tư nhân làm mất đi cơ hội phát triển năng lực trong khu vực công cộng. Hơn nữa, các doanh nghiệp tư nhân thường ít chịu trách nhiệm kết hợp các thành phần vào trong đánh giá, điều này sẽ làm cho việc đánh giá trở nên tốn kém hơn. Cho dù đối tác nào hay sự kết hợp các đối tác nào cuối cùng được tạo thành, thì việc xem xét lại các hoạt động đánh giá trong quá khứ của những người cộng tác tiềm tàng cũng rất quan trọng để có thể đưa ra một sự lựa chọn trên cơ sở đầy đủ thông tin.

Và vấn đề thứ ba là cần xác định mức độ phân tách thể chế được áp dụng giữa những người cung cấp đánh giá và những người sử dụng đánh giá. Có nhiều lợi ích từ khách quan khi việc đánh giá được tiến hành độc lập với thể chế chịu trách nhiệm cho dự án được đánh giá. Tuy nhiên, đánh giá có thể có nhiều mục tiêu, kể cả việc xây dựng năng lực đánh giá bên trong các cơ quan chính phủ và giúp những người điều hành dự án nhận thức được thực tế về dự án của họ một khi các dự án được thực hiện trên thực tế. Ở mức tối thiểu, những người sử dụng đánh giá, từ các nhà hoạch định chính sách trong các cơ quan chính phủ ở các nước khách hàng, tới các tổ chức NGO, các nhà tài trợ song phương, và các thể chế phát triển quốc tế, đều phải tham gia đầy đủ vào công việc đánh giá nhằm đảm bảo quá trình đánh giá hợp lệ và các kết quả có được phù hợp với nhu cầu thông tin của những đối tượng này. Nếu không, khó có thể sử dụng các kết quả đánh giá để cung cấp thông tin cho chính sách. Trong phân tích cuối cùng, người quản lý đánh giá và những khách hàng của anh ta cần phải đạt được sự cân bằng hợp lý giữa sự tham gia của những người sử dụng đánh giá và duy trì tính khách quan và hợp lệ của các kết quả.

### Phát triển dữ liệu

Các dữ liệu đầy đủ và đáng tin cậy là đầu vào cần thiết để

đánh giá tác động của dự án. Các dữ liệu có chất lượng cao là rất cần thiết để đảm bảo tính hợp lệ của các kết quả đánh giá. Như được thảo luận ở trên, đánh giá xem hiện đang có những dữ liệu nào là bước quan trọng đầu tiên trước khi tiến hành một nỗ lực thu thập dữ liệu mới bất kỳ. Bảng 2.1 tạo ra sự liên kết giữa các phương pháp đánh giá cơ bản nhất với các yêu cầu dữ liệu. Hầu hết các phương pháp này có thể kết hợp các kỹ thuật định tính và tham dự (participatory techniques) trong thiết kế công cụ điều tra, xác định các chỉ số, và để xác định các biến số kiểm chứng, các biến số sử dụng để tương hợp (matching) hay các biến số công cụ.

**Bảng 2.1 Các phương pháp đánh giá và yêu cầu cầu dữ liệu tương ứng**

Phương pháp	Yêu cầu dữ liệu		Sử dụng phương pháp định tính
	Tối thiểu	Lý tưởng	
Các kiểm soát thử nghiệm hay ngẫu nhiên	Dự án đơn nhất, có đại diện của cả người thụ hưởng và người không thụ hưởng	Điều tra cơ sở và tiếp theo về người thụ hưởng và người không thụ hưởng. Cho phép kiểm soát các sự kiện xảy ra đồng thời, ngoài việc kiểm soát để do lường tác động. (cho phép có ước lượng khác biệt trong khác biệt.)	<ul style="list-style-type: none"> <li>● Cung cấp thông tin về thiết kế của công cụ điều tra, chọn mẫu.</li> <li>● Xác định các chỉ số</li> <li>● Thu thập và ghi chép dữ liệu sử dụng</li> <li>- Dữ liệu dạng văn bản</li> <li>- Các cuộc phỏng vấn không chính thức hay bán cấu trúc</li> <li>- Các nhóm</li> </ul>
Các thiết kế không qua thử nghiệm	Điều tra lớn, điều tra dân số, ngân sách	Điều tra lớn, và điều tra hộ gia đình trên cơ sở dự án có quy mô nhỏ, áp dụng với hai loại điều tra	
a) Các kiểm soát được xây	quốc gia, hay		

**Bảng 2.1 (Tiếp)**

Phương pháp	Yêu cầu dữ liệu		Sử dụng phương pháp định tính
	Tối thiểu	Lý tưởng	
dựng (constructed controls) hay tương hợp (matching)	LSMS trong đó chọn mẫu nhiều hơn đối với những người thụ hưởng	mốc thời gian nhằm kiểm soát các sự kiện xảy ra đồng thời	tập trung hay các cuộc gặp gỡ cộng đồng - Quan sát trực tiếp - Các phương
b) So sánh phản thân (reflexive comparisons) và khác biệt kép	Điều tra cơ sở và tiếp theo về những người thụ hưởng	Chuỗi thời gian hay nhóm mẫu (panel) về những người thụ hưởng và những người không thụ hưởng tương đương	pháp tham dự (participatory) - Các ảnh chụp - Quy trình tam giác - Phân tích dữ liệu
c) Kiểm soát thống kê hay biến số công cụ	Dữ liệu chéo (cross-section) đại diện cho những người thụ hưởng với các biến công cụ tương ứng	Dữ liệu chéo và chuỗi thời gian đại diện cho cả những người thụ hưởng và không thụ hưởng với các biến công cụ tương ứng	

Nguồn: Điều chỉnh căn cứ vào Ezemenari, Rudqvist, và Subbarao (1999) và Bamberger

Đối với những đánh giá sẽ tự tạo ra dữ liệu, thì có những bước quan trọng trong thiết kế các công cụ thu thập dữ liệu, chọn mẫu, công việc thực địa, quản lý dữ liệu và tiếp cận dữ liệu. Phần này không phác thảo tiến trình từng bước về cách thức tiến hành một cuộc điều tra, nhưng sẽ đưa ra một thảo luận ngắn gọn về các bước này. Một số nội dung thảo luận trong phần này, nhất là về việc lấy mẫu và quản lý dữ liệu, phù

hợp hơn với các đánh giá dựa trên cơ sở thu thập và phân tích các cuộc điều tra mẫu trên quy mô lớn, sử dụng các dữ liệu định lượng, hơn là cho các đánh giá sử dụng dữ liệu định tính và có quy mô mẫu nhỏ.

**Quyết định đo lường cái gì.** Cần thiết lập đầu ra chính và các chỉ số tác động trong giai đoạn lập kế hoạch đánh giá, như là một phần trong cách tiếp cận khung logic (logical framework approach). Để đảm bảo đánh giá có thể đánh giá được các kết quả trong một thời kỳ, phù hợp với nhu cầu của người ra quyết định, thì cần phải thiết lập một thang bậc các chỉ số, từ các chỉ số tác động trong ngắn hạn, như tỷ lệ đi học, tới các chỉ số dài hạn hơn, như thành tích học tập. Điều này đảm bảo ngay cả khi các tác động cuối cùng chưa được nhận thấy tức thời thì vẫn có thể đánh giá được các đầu ra của chương trình. Hơn nữa, người đánh giá cần phải lập kế hoạch đo lường việc thực hiện chương trình cũng như tính các nhân tố ngoại sinh có thể tác động tới kết quả được quan tâm.

Các nhà quản lý đánh giá cũng có thể lập kế hoạch tiến hành đánh giá trong một vài thời kỳ, điều này cho phép ghi nhận được nhiều tác động tức thời hơn, trong khi vẫn theo dõi được các kết quả cuối cùng. Cách làm như vậy đã được thực hiện trong đánh giá Cải cách trường học ở Nicaragua, trong đó các tác động ngắn hạn của cải cách tới sự tham dự của cha mẹ học sinh và tỷ lệ tham dự của học sinh và giáo viên đã được xác định và các tác động dài hạn tới thành tích của học sinh vẫn đang được đánh giá.

Thông tin về những đặc điểm của nhóm dân cư thụ hưởng tuy không liên quan chặt chẽ tới đánh giá tác động nhưng cũng cần xem xét do cần đến trong phân tích, ví dụ như về mức độ đói nghèo, hay quan điểm của những người này về chương trình. Hơn nữa, người đánh giá cũng có thể muốn tính đến các thước do chi phí để có thể tiến hành một số phân tích chi phí - hiệu quả hay các đánh giá bổ sung khác không liên quan chặt chẽ tới đánh giá tác động.

q: Kiểu thiết kế đánh giá được lựa chọn cũng có một số yêu cầu về dữ liệu. Các yêu cầu dữ liệu phụ thuộc vào phương pháp sử dụng, nhóm dân cư được quan tâm, các thước đo tác động, và các thành phần khác của đánh giá. Chẳng hạn, nếu sử dụng phương pháp biến công cụ (một trong số các phương pháp so sánh tương hợp) thì cần phải xác định và đưa vào trong công việc thu thập dữ liệu (những) biến sẽ sử dụng như là biến công cụ, nhằm tách sự tham dự chương trình với các thước đo kết quả. Cách làm này đã được thực hiện trong đánh giá tác động của Quỹ đầu tư xã hội Bôlivia, trong đó kiến thức về quỹ xã hội và sự hiện diện của các NGO được sử dụng như là các biến công cụ để đánh giá tác động của các nội dung trong chương trình vốn xã hội.

Cũng sẽ có ích khi xây dựng một ma trận đánh giá, liệt kê các câu hỏi quan tâm, các chỉ số đầu ra sẽ được sử dụng để đánh giá kết quả, biến số, và nguồn dữ liệu cho biến số này. Sau đó có thể sử dụng ma trận này để đánh giá lại bảng câu hỏi và lập kế hoạch công việc phân tích, như đã được thực hiện trong đánh giá Quỹ đầu tư xã hội khẩn cấp của Nicaragua (xem Phụ lục 6).

**Triển khai các công cụ và phương pháp thu thập dữ liệu.** Triển khai các công cụ thu thập dữ liệu phù hợp nhằm tạo ra những dữ liệu cần thiết để trả lời các câu hỏi đánh giá là một công việc khó khăn. Công việc này đòi hỏi phải có sự tham gia của các nhà phân tích trong quá trình phát triển các câu hỏi, kiểm tra thí điểm, và xem xét lại dữ liệu thu thập được từ kiểm tra thí điểm. Sự tham dự của cả các nhà quản lý thực địa và các nhà quản lý dữ liệu trong quá trình phát triển các công cụ, cũng như của các cán bộ trong nước - nhất là của những nhà phân tích có kiến thức về cả đất nước và về chương trình - là rất quan trọng đối với chất lượng thông tin thu thập được (Grosh và Munoz 1996). Việc đảm bảo các dữ liệu thu thập được có thể tách ra theo giới tính nhằm khai thác tác động khác biệt của các chương trình và chính sách cụ thể cũng rất quan trọng.

Các đánh giá định lượng thường thu thập và ghi chép thông tin dưới dạng số hay theo các khoản mục được mã hóa từ trước. Đối với đánh giá định tính, thông tin thường được trình bày như là một văn bản mô tả có rất ít hay không có phân loại. Thông tin này có thể bao gồm phản ứng của một cá nhân đối với các câu hỏi phỏng vấn kết thúc mở, ghi chép trong nhóm tập trung, hay quan sát của người đánh giá về các sự kiện. Một số nghiên cứu định tính cũng sử dụng sự phân loại dữ liệu được mã hoá từ trước (Bamberger, 2000). Các loại công cụ thu thập dữ liệu và ưu nhược điểm của chúng được tóm tắt trong bảng 2.2- kỹ thuật được sử dụng phổ biến nhất là các bảng câu hỏi.

Các phản ứng trước các bảng câu hỏi điều tra có thể rất nhạy cảm với thiết kế của bảng câu hỏi, do vậy, điều quan trọng là đảm bảo cấu trúc và thể thức của bảng câu hỏi phải phù hợp, tốt nhất là các bảng câu hỏi cần được những cán bộ có kinh nghiệm thực hiện. Chẳng hạn, lợi ích của các dữ liệu định lượng thường bị suy giảm trầm trọng do các lý do kỹ thuật đơn giản, ví dụ việc thiếu khả năng liên kết dữ liệu giữa các nguồn với nhau. Đây là trường hợp xảy ra khi đánh giá tình trạng giáo dục quốc gia ở một nước, tại đây dữ liệu về quá trình đào tạo của người học không thể liên hệ được với kết quả điểm thi. Điều này làm cho không thể đánh giá được tác động của các đặc điểm của người học tới kết quả thi, hoặc không thể phân loại điểm thi theo tuổi tác, giới tính, tình trạng kinh tế xã hội, hay kinh nghiệm học tập của người học.

Đối với việc thu thập dữ liệu định tính và định lượng, ngay cả các cán bộ có kinh nghiệm cũng cần phải được đào tạo để thu thập các dữ liệu đặc trưng cho đánh giá, và tất cả công việc thu thập dữ liệu cần có một tập hợp các tài liệu hướng dẫn để hướng dẫn. Những tài liệu này có thể sử dụng như là định hướng trong thời gian đào tạo và như là tài liệu tham khảo trong khi thực hiện công việc thực địa. Thực tế cho thấy, tùy thuộc vào mức độ phức tạp của công việc thu thập dữ liệu, thời gian đào tạo có thể kéo dài từ ba ngày tới vài tuần.

Kiểm tra thí điểm là một bước rất quan trọng, nó sẽ cho thấy, liệu công cụ được sử dụng có thể cho các kết quả cần thiết một cách đáng tin cậy hay không, và làm thế nào để có thể thực hiện các quy trình thu thập dữ liệu trên thực tế. Kiểm tra thí điểm cần mô phỏng công việc thực tế thực địa càng giống càng tốt. Vì vậy, cần phải có sẵn các chương trình nhập dữ liệu vào thời điểm tiến hành kiểm tra thí điểm để kiểm tra hoạt động của chương trình này, cũng như để thí điểm với các nhóm dân cư và khu vực địa lý khác nhau, sẽ được tiếp cận khi tiến hành điều tra thực địa thực sự.

**Lấy mẫu.** Lấy mẫu là một nghệ thuật được thực hiện một cách tốt nhất do một chuyên gia chọn mẫu có kinh nghiệm. Thiết kế đánh giá không nhất thiết phải phức tạp nhưng cần có kinh nghiệm của chuyên gia chọn mẫu trong việc quyết định các khuôn khổ mẫu, quy mô mẫu và chiến lược lựa chọn mẫu phù hợp (việc thảo luận về chọn mẫu ở đây đề cập chủ yếu tới các vấn đề liên quan tới các đánh giá trong đó thu thập dữ liệu định lượng từ các mẫu lớn hơn và có tính đại diện về thống kê.) Chuyên gia chọn mẫu cũng cần tham gia vào tiến trình đánh giá ngay từ các giai đoạn đầu tiên để xem xét lại những thông tin sẵn có cần thiết để lựa chọn mẫu và quyết định liệu có cần công việc liệt kê (enumeration work) nào không- điều này có thể tốn nhiều thời gian.

Cũng như những phần khác trong công việc đánh giá, sự phối hợp giữa chuyên gia chọn mẫu và nhóm đánh giá là quan trọng. Sự phối hợp này trở nên đặc biệt quan trọng khi tiến hành các so sánh tương hợp, bởi vì thiết kế mẫu trở thành cơ sở cho "sự tương hợp", vốn là cốt lõi của thiết kế đánh giá và xây dựng phản chứng. Trong các trường hợp này, chuyên gia chọn mẫu phải làm việc chặt chẽ với nhóm đánh giá để xây dựng các tiêu chuẩn sẽ được áp dụng nhằm làm cho nhóm tham dự chương trình và nhóm đối chứng trở nên giống nhau. Chẳng hạn, trong đánh giá dự án cải cách quyền tự chủ trường học ở Nicaragua, các trường học tự chủ được phân tổ căn cứ vào loại hình trường, tỷ lệ nhập học, thời gian tham gia cải cách, địa

**Bảng 2.2 Các công cụ thu thập dữ liệu chủ yếu để đánh giá tác động**

Kỹ thuật	Định nghĩa và sử dụng	Ưu điểm	Nhược điểm
Nghiên cứu tinh huống	Thu thập thông tin nhờ đó hình thành một câu chuyện có tính chất mô tả hay giải thích và có thể được sử dụng để trả lời cho các câu hỏi như thế nào và tại sao	Có thể sử dụng đầy đủ các bằng chứng đã dang từ các tài liệu, phỏng vấn, quan sát - Cố gắng để thu thập năng lực giải thích khi sự tập trung nhằm vào các thể chế, tiến trình, chương trình, quyết định, và sự kiện	- Khó tiến hành các nghiên cứu tinh huống tốt - Đôi khi nghiên cứu chuyên môn và kỹ năng viết tốt - Các kết luận không thể khái quát hóa cho tổng thể - Tốn nhiều thời gian - Khó lặp lại
Nhóm tập trung	Tiến hành các cuộc thảo luận tập trung với các thành viên trong tổng thể mục tiêu, những người quen thuộc với các vấn đề liên quan trước khi soạn thảo một tập hợp các câu hỏi có câu trúc. Mục đích là so sánh quan điểm của những người thu hướng với những khái niệm trung庸 trong mục tiêu của người đánh giá	- Các lợi thế tương tự như phỏng vấn (độ cập dưới đây) - Đặc biệt có ích khi cần cỗ sự tương tác với những người tham dự chương trình - Là một cách hữu hiệu để xác định được các tác động theo thứ bậc	- Có thể tốn kém về tiền của và thời gian - Cần phải nhạy cảm với sự phản ứng các cấp độ theo thứ bậc - Không có khả năng khái quát hóa
Phỏng vấn	Người phỏng vấn đặt ra các câu hỏi với một hay nhiều người và ghi chép câu trả lời của người trả lời. Các phỏng vấn có thể chính thức hay không chính thức, trực tiếp hay qua điện thoại, các câu hỏi kết thúc đóng hoặc mở.	- Các cá nhân và thế chế có giải thích về kinh nghiệm của mình bằng lời lẽ và trong bối cảnh của bản thân họ - Linh hoạt, cho phép người phỏng vấn đặt ra những câu hỏi chưa dự kiến trước và dễ sâu vào các vấn đề	- Tốn nhiều thời gian - Có thể tốn kém - Nếu không được thực hiện đúng đắn, người phỏng vấn có thể tác động tới câu trả lời của người được phỏng vấn - Đặc biệt có ích khi gặp những khó khăn về ngôn ngữ

<p><b>Quan sát</b></p>	<p>Quan sát và ghi chép lại tình hình trong sổ sách hay nhật ký. Thông tin ghi chép gồm có: ai tham gia; điều gì diễn ra; khi nào, ở đâu, và các sự kiện diễn ra như thế nào.</p> <p>Quan sát có thể trực tiếp (người quan sát nhìn và ghi chép) hay có tính chất tham dự (người quan sát trở thành một phần của bối cảnh trong một khoang thời gian)</p>	<ul style="list-style-type: none"> <li>- Khả năng nhận được đầu vào từ các quan chức cấp cao tốt</li> <li>- Cung cấp thông tin mô tả về bối cảnh và các thay đổi được quan sát để liệu phụ thuộc lớn vào kỹ năng quan sát và ghi chép của người quan sát</li> <li>- Các kết quả có thể diễn giải theo nhiều cách khác nhau</li> <li>- Không dễ áp dụng trong khung thời gian ngắn vào sự thay đổi quá trình</li> </ul>	<ul style="list-style-type: none"> <li>- Chất lượng và tính hữu ích của dữ liệu phụ thuộc lớn vào kỹ năng quan sát và ghi chép của người quan sát</li> <li>- Chất lượng các câu trả lời phụ thuộc nhiều vào sự sáng sủa của các câu hỏi</li> <li>- Đối với khó có thể thuyết phục mọi người hoàn thành và trả lời bằng câu hỏi</li> <li>- Có thể áp dụng cho mẫu lớn một cách đồng thời</li> <li>- Cho phép người trả lời có thời gian suy nghĩ trước khi trả lời</li> <li>- Có thể trả lời một cách vô danh</li> <li>- Được sử dụng thống nhất bằng việc hỏi tất cả những người trả lời những câu hỏi giống nhau</li> <li>- Làm cho việc soạn thảo và so sánh dữ liệu trở nên dễ dàng hơn</li> <li>- Có thể áp dụng cho mẫu lớn một cách đồng thời</li> <li>- Có thể trả lời một cách vô danh</li> <li>- Được sử dụng thống nhất bằng việc hỏi tất cả những người trả lời những câu hỏi giống nhau</li> <li>- Có thể dân đến việc áp đặt các hoạt động thể chế và kinh nghiệm của mọi người vào trong các khoản mục đã được quyết định từ trước</li> <li>- Có thể nhận diện được các khó khăn để tiếp tục điều tra và cung cấp bằng chứng về hành động, sự thay đổi, và ánh hưởng nhằm hỗ trợ cho nhận thức của người trả lời</li> <li>- Có thể ít tốn kém</li> </ul>
	<p><b>Bảng câu hỏi</b></p>	<p>Phát triển một tập hợp các câu hỏi điều tra trong đó các câu trả lời có thể được mã hóa một cách nhất quán</p>	<p>Xem xét lại các tài liệu như sổ sách, cơ sở dữ liệu hành chính, các tài liệu đào tạo và thư tín</p>

### Phân tích tài liệu văn bản

Bảng câu hỏi

Nguồn: Điều chỉnh căn cứ vào Taschereau (1998).

điểm, và sau đó được so sánh với một mẫu gồm các nhà trường không tự chủ bằng cách sử dụng cùng các tổ tương tự, trừ thời gian tham gia cải cách. Công việc này do một thành viên trong nhóm thực hiện: người chịu trách nhiệm thu thập dữ liệu để hỗ trợ cho một chuyên gia chọn mẫu có được các thông tin cần thiết, kể cả dữ liệu về các chỉ số kết quả chọn lọc trong tính toán sức mạnh (power- thuật ngữ chỉ ước lượng của quy mô mẫu cần thiết để kiểm định ý nghĩa thống kê giữa hai nhóm này), một danh sách tổng thể để lựa chọn mẫu, và các đặc điểm chi tiết về nhóm tham dự và nhóm đối chứng tiềm năng quan trọng đối với tiến trình lựa chọn mẫu.

Có nhiều sự đánh đổi giữa chi phí và tính chính xác trong chọn mẫu cần phải làm rõ khi xây dựng khuôn khổ lấy mẫu. Ví dụ, việc tiến hành một mẫu trong hai hay ba giai đoạn sẽ làm giảm chi phí của cả chọn mẫu và nghiên cứu thực địa, nhưng sẽ làm tăng sai số chọn mẫu, và do đó, tác động tới tính chính xác của các ước lượng.

Một khi các biến kết quả và tổng thể được quan tâm đã được nhóm đánh giá quyết định, thì bước đầu tiên của chuyên gia chọn mẫu là quyết định tính toán sức mạnh (power- tham khảo Valadez và Bamberger 1994, trang 382-84, trong đó có thảo luận về tiến trình tính toán sức mạnh). Do việc tính toán sức mạnh có thể được thực hiện chỉ với duy nhất một thước đo kết quả, và các đánh giá thường cân nhắc một số thước đo, nên cần phải có một số quyết định chiến lược về những chỉ số kết quả nào sẽ được sử dụng khi thiết kế mẫu.

Sau khi đã xây dựng chiến lược và khuôn khổ chọn mẫu, chuyên gia chọn mẫu cần tham gia vào việc lựa chọn mẫu để điều tra và kiểm tra thử để đảm bảo việc kiểm tra thử này không được tiến hành trong một khu vực sẽ được đưa vào mẫu cho điều tra thực địa. Thông thường, công việc thực địa ban đầu được xem như là một phần trong thủ tục lựa chọn mẫu. Chẳng hạn, cần phải có một tiến trình liệt kê (enumeration) nếu không có bản đồ cập nhật về các đơn vị cần thiết cho mẫu (các hộ gia đình, trường học, v.v), hay nếu như một tổng thể được quan

Tâm nhất định, như trẻ em suy sinh dưỡng, cần được xác định từ trước để nó có thể được lựa chọn vì mục đích đánh giá.

Một khi đã kết thúc công việc thực địa, chuyên gia chọn mẫu cần phải hỗ trợ việc quyết định các trọng số mẫu để tính toán các nhân tố mở rộng và sửa chữa các sai số chọn mẫu và các trường hợp không trả lời. (Grosh và Munoz 1996 thảo luận chi tiết về các thủ tục chọn mẫu như là một phần trong điều tra hộ gia đình. Kish 1965 được xem là một trong những cuốn giáo trình chuẩn tắc trong lĩnh vực chọn mẫu). Và cuối cùng, chuyên gia chọn mẫu cần cung cấp một tài liệu chọn mẫu mô tả chi tiết chiến lược chọn mẫu, kể cả (a) từ giai đoạn thiết kế mẫu, tính toán sức mạnh, sử dụng các biến tác động, xác định các sai số chọn mẫu và quy mô mẫu, sử dụng sự phân lớp để phân tích các tổng thể quan tâm; (b) từ giai đoạn lựa chọn mẫu, phác thảo các giai đoạn lấy mẫu và thủ tục lựa chọn; (c) từ giai đoạn thực địa để chuẩn bị cho phân tích, mối liên hệ giữa quy mô của mẫu và tổng thể từ đó mẫu được lựa chọn, tỷ lệ không trả lời, và các thông tin khác được sử dụng để làm cơ sở thông tin cho các trọng số của mẫu; cũng như bất kỳ thông tin bổ sung nào mà nhà phân tích cần để cung cấp thông tin về việc sử dụng dữ liệu đánh giá. Tài liệu này có thể được sử dụng để duy trì sổ sách dự án về đánh giá và cần được phân phối kèm với dữ liệu để giúp hướng dẫn các nhà phân tích sử dụng dữ liệu đánh giá.

**Bảng câu hỏi.** Thiết kế của bảng câu hỏi có vai trò quan trọng đối với tính hợp lệ của thông tin thu thập được. Cần có bốn loại thông tin chung trong đánh giá tác động (Valadez và Bamberger 1994). Các thông tin này gồm có:

- Phân loại các dữ liệu danh nghĩa theo người phỏng vấn, tuy theo việc những người này là những người tham gia dự án hay thuộc về nhóm đối chứng;
- Các biến số được nghiên cứu không chỉ bao gồm các dịch vụ và lợi ích nhận được, mà còn cả tần số, số lượng, và chất lượng- việc đánh giá chất lượng có thể tương đối khó khăn;

- Các biến số kết quả để đo lường tác động của một dự án, kể cả các sản phẩm trung gian, đầu ra ổn định hay việc các dịch vụ được cung cấp liên tục trong một thời gian dài, và các tác động của dự án như thu nhập và việc làm được cải thiện; và
- Các biến số can thiệp có tác động tới sự tham gia dự án hay tới kiểu tác động được tạo ra, chẳng hạn, các đặc tính của cá nhân, hộ gia đình, hay cộng đồng - những biến số này có thể có vai trò quan trọng trong việc khám phá ra các sai lệch (bias).

Cách thức đặt câu hỏi, cũng như thứ tự của các câu hỏi, cũng khá quan trọng để tạo ra những thông tin đáng tin cậy. Một ví dụ phù hợp là đo lường phúc lợi, một công việc cần làm để đo lường tác động trực tiếp của một dự án tới giảm nghèo đói. Hỏi cá nhân về mức thu nhập của họ không phải lúc nào cũng cho biết chính xác về mức phúc lợi kinh tế của họ. Như được thảo luận trong các tài liệu nghiên cứu về đo lường phúc lợi, người ta thường sử dụng các câu hỏi về chi tiêu, thành phần hộ gia đình, tài sản, quà tặng và tiền gửi về gia đình, và giá trị của thực phẩm tự trồng và nhà ở của chủ nhà ở để nắm bắt được giá trị thực sự của phúc lợi hộ gia đình và cá nhân. Khung thời gian gọi nhở các khoản mục chi tiêu, hoặc thứ tự các câu hỏi được nêu ra, có thể tác động đáng kể tới tính hợp lệ của các thông tin thu thập được.

Trong số các thành phần được lưu ý của một bảng câu hỏi tốt, có việc duy trì bảng câu hỏi ngắn gọn và tập trung vào những câu hỏi quan trọng, đảm bảo những chỉ dẫn và câu hỏi rõ ràng, hạn chế chỉ trong số những câu hỏi cần thiết cho đánh giá, đưa vào lựa chọn “không có ý kiến” đối với các câu hỏi kết thúc đóng để đảm bảo có được dữ liệu đáng tin cậy, và sử dụng các thủ tục chắc chắn để thực hiện bảng câu hỏi. Những nội dung này trên thực tế có thể khác nhau đối với các cuộc điều tra định lượng và định tính.

**Các công việc thực địa.** Làm việc với các cán bộ trong nước,

những người có nhiều kinh nghiệm trong thu thập các dữ liệu tương tự với dữ liệu cần thiết cho đánh giá có thể hỗ trợ rất nhiều cho hoạt động thực địa. Các cán bộ này không chỉ cung cấp những hiểu biết cần thiết về khu vực địa lý được điều tra, mà những kiến thức của họ còn rất quan trọng để xây dựng các tiêu chuẩn sử dụng trong việc định vị và tiếp cận những người cung cấp thông tin. Có thể phát sinh vấn đề khi các cán bộ thực địa có kinh nghiệm trong một lĩnh vực khác với lĩnh vực cần đánh giá, như trường hợp đánh giá về giáo dục ở Nicaragua, sử dụng một doanh nghiệp chuyên môn hoá về trung cầu ý kiến công chúng để tiến hành điều tra các trường học và hộ gia đình. Kinh nghiệm đã có cho phép doanh nghiệp này nổi tiếng nhờ sự dự đoán chính xác các kết quả bầu cử, lại không có ích gì để có thể hiểu được cách thức tiếp cận học sinh hay kết hợp các tập hợp dữ liệu định lượng với nhau. Việc thiếu kinh nghiệm này đã gây ra các vấn đề thực hiện điều tra đáng kể, và một nhóm phối hợp giữa Bộ Giáo dục và Ngân hàng Thế giới đã phải mất nhiều tuần để sửa chữa.

Tuỳ thuộc vào mục tiêu và trọng tâm của đánh giá, loại cán bộ cần thiết cho việc thu thập dữ liệu ở thực địa sẽ khác nhau. Chẳng hạn, một đánh giá tác động định lượng đối với một chương trình dinh dưỡng có thể đòi hỏi một chuyên gia nhân trắc học trong nhóm điều tra để thu thập các chỉ số chiều cao trên cân nặng, trong khi đó đánh giá tác động của một chương trình cải cách giáo dục lại cần đến các cán bộ có chuyên môn trong việc áp dụng các kỹ kiểm tra thành tích để đo lường tác động của cải cách tới thành tích học tập. Hầu hết các điều tra định lượng sẽ cần đến ít nhất một người quản lý điều tra, một người quản lý dữ liệu, một người quản lý thực địa, các giám sát viên thực địa, các cán bộ phỏng vấn, những người nhập dữ liệu, và các lái xe. Tuỳ thuộc vào phương pháp định tính được sử dụng, các cán bộ thực địa trong điều tra định tính cũng tương tự, trừ những người nhập dữ liệu. Tuy nhiên, kỹ năng của những cán bộ phỏng vấn có thể rất khác nhau, trong đó các cán bộ phỏng vấn trong điều tra định tính cần phải được đào tạo

chuyên môn, đặc biệt là với trường hợp các nhóm tập trung quan sát trực tiếp v.v.

Có ba mối quan tâm khác cần chú ý khi lập kế hoạch cho hoạt động điều tra. Thứ nhất, cần phải tính đến các sự kiện nhạy thời có thể tác động tới thành công của công việc thực địa và tính hợp lệ bên ngoài của dữ liệu thu thập được, chẳng hạn như chương trình năm học của nhà trường, các kỳ nghỉ, các mù mưa, thời gian thu hoạch hay hình thái di cư. Thứ hai, cần phải thí điểm các công cụ thu thập dữ liệu, ngay cả khi chúng được điều chỉnh dựa trên các công cụ trước kia được sử dụng, và để kiểm tra chất lượng của công cụ trong việc tạo ra các dữ liệu được yêu cầu, vừa để những cán bộ thực địa làm quen với sự năng động trong tiến trình thu thập dữ liệu. Các kiểm tra thường đóng vai trò cơ sở để lựa chọn một nhóm cốt cán các cán bộ thực địa nhằm thực hiện điều tra thực sự. Nhiều người thường thu thập dữ liệu có kinh nghiệm sẽ bắt đầu với số cán bộ trong kiểm tra thử nhiều hơn từ 10 đến 20 phần trăm so với trên thực địa thực sự, và sau đó sẽ lựa chọn những người làm tốt nhất trong kiểm tra thử để thành lập nhóm thu thập dữ liệu thực sự. Cuối cùng, liên lạc đóng vai trò thiết yếu đối với những người điều tra thực địa. Ví dụ, nếu các điều kiện trong nước cho phép sử dụng thì có thể tăng cường hiệu quả công việc thực địa bằng việc cung cấp cho các giám sát viên các điện thoại di động mà họ có thể liên lạc với người quản lý điều tra, người quản lý thực địa, và các cán bộ khác nhằm trả lời các câu hỏi và thông báo về tiến độ thực hiện.

**Quản lý và tiếp cận dữ liệu.** Mục tiêu của một hệ thống quản lý dữ liệu tốt là phải đảm bảo sự kịp thời và chất lượng của dữ liệu thu thập. Sự kịp thời phụ thuộc vào việc, càng nhiều sự phối hợp càng tốt giữa công việc thu thập và xử lý dữ liệu để có thể xác nhận và sửa chữa các sai lầm trước khi thực công việc thực địa. Chất lượng của dữ liệu có thể đảm bảo bằng cách kiểm tra tính nhất quán để xác nhận sự hợp lệ nội bộ của dữ liệu thu thập được ở cả trong và sau

nhập dữ liệu, và bằng cách đảm bảo rằng, những nhà phân tích sử dụng dữ liệu sẽ có được các tài liệu phù hợp. Các tài liệu cần có, gồm hai dạng thông tin: (a) các thông tin cần thiết để diễn giải dữ liệu, kể cả các sổ ký hiệu (codebook), từ điển dữ liệu, hướng dẫn về các biến số được xây dựng, và các tài liệu dịch cần thiết; (b) các thông tin cần thiết để tiến hành phân tích, thường nằm trong một tài liệu thông tin cơ bản, trong đó có mô tả về trọng tâm và mục tiêu của đánh giá, các chi tiết về phương pháp đánh giá, bản tóm tắt hay bản sao các công cụ thu thập dữ liệu, thông tin về mẫu, thảo luận về công việc thực địa, và hướng dẫn sử dụng dữ liệu.

Các dữ liệu có được qua công tác đánh giá nên được cung cấp công khai do giá trị của việc đánh giá có tính chất hàng hoá công cộng và do có thể có một nhóm khác với nhóm tiến hành công việc đánh giá ban đầu thấy cần phải tiến hành các công việc bổ sung tiếp theo nhằm đánh giá những tác động lâu dài. Để hỗ trợ cho quá trình chia sẻ dữ liệu, ngay trong giai đoạn đánh giá ban đầu, cần nhất trí và ký kết về một chính sách tiếp cận dữ liệu dễ dàng, thiết lập các chuẩn mực và trách nhiệm phân phối dữ liệu. Một chính sách tiếp cận dữ liệu rộng rãi còn đòi hỏi, cần có các tài liệu về dữ liệu tốt và bảo vệ tính bảo mật của những người cung cấp thông tin. Nếu các dữ liệu lắp được thu thập từ cùng đối tượng theo thời gian do các cơ quan khác nhau, thì phải xác định được những người cung cấp thông tin để tiến hành công việc tiếp theo. Yêu cầu này cần được cân bằng với các tiêu chuẩn bảo mật thường đi kèm với bất kỳ một nghiên cứu nào trong lĩnh vực xã hội. Một giải pháp có thể là, cho phép một đơn vị ẩn danh ghi chép lại các số liệu để cung cấp cho các nhà phân tích quan tâm, nhưng phải yêu cầu các nhà nghiên cứu quan tâm tiến hành các nghiên cứu tiếp theo liên lạc với cơ quan phụ trách dữ liệu để có được danh sách các đơn vị ở trong mẫu, do vậy tạo ra cho cơ quan này cơ hội đảm bảo kiểm soát được chất lượng thông qua mối liên hệ với các nhà nghiên cứu dự định tiến hành công việc này.

### Phân tích, báo cáo, và phổ biến

Cũng như các giai đoạn khác trong tiến trình đánh giá, việc phân tích dữ liệu đánh giá, cho dù là định lượng hay định tính đòi hỏi phải có sự hợp tác giữa các nhà phân tích, những người tạo ra số liệu và các nhà hoạch định chính sách để làm rõ các câu hỏi và đảm bảo có các kết quả kịp thời và có chất lượng. Các vấn đề liên quan tới việc làm sạch và diễn giải dữ liệu hầu như chắc chắn sẽ phát sinh trong khi phân tích, và đòi hỏi phải có sự tham gia của các thành viên khác của nhóm.

Một số kỹ thuật và khó khăn trong việc tiến hành phân tích định lượng dựa vào các phương pháp thống kê sẽ được đề cập trong chương 3. Có nhiều kỹ thuật phân tích dữ liệu định tính (tham khảo Miles và Huberman 1994). Mặc dù việc thảo luận chi tiết các phương pháp này nằm ngoài phạm vi của cuốn sổ tay này, nhưng ở đây chúng tôi cũng đề cập tới hai phương pháp được sử dụng phổ biến nhất trong đánh giá tác động, đó là phân tích nội dung và phân tích tinh huống (Taschereau 1998).

Phân tích nội dung thường được sử dụng để phân tích các dữ liệu thu thập được từ phỏng vấn, quan sát và tài liệu. Khi xem xét lại các dữ liệu này, người đánh giá phát triển một hệ thống phân loại dữ liệu, tổ chức thông tin dựa trên (a) các câu hỏi đánh giá làm căn cứ thu thập thông tin; (b) các tài liệu được sử dụng như thế nào; và (c) nhu cầu tham chiếu chéo thông tin. Việc mã hoá dữ liệu có thể khá phức tạp và đòi hỏi nhiều giả định. Một khi đã tạo lập được hệ thống phân loại, thì giai đoạn phân tích được bắt đầu, và đây cũng là một quá trình khó khăn. Giai đoạn này liên quan đến việc tìm kiếm các xu hướng trong dữ liệu và đi xa hơn việc mô tả để hướng đến việc tăng cường hiểu biết về các quá trình, kết quả, và tác động của chương trình. Công việc này được tiến hành tốt nhất với sự tham gia của các thành viên trong nhóm. Hiện nay đang có các chương trình máy tính dân tộc học và ngôn ngữ học mới, được thiết kế để hỗ trợ cho việc phân tích dữ liệu định tính.

Phân tích tình huống được dựa trên các trường hợp nghiên cứu tình huống, được thiết kế để nghiên cứu kỹ lưỡng về một nhóm hay một cá nhân cụ thể. Mức độ chi tiết cao có thể cung cấp các thông tin phong phú cho việc đánh giá tác động của dự án. Các quá trình thu thập và phân tích dữ liệu được tiến hành đồng thời, khi người đánh giá quan sát cũng là lúc họ thu thập thông tin. Những người đánh giá sau đó có thể xây dựng và kiểm định các giải thích và liên kết các phần quan trọng của thông tin với nhau.

Cho dù là phân tích thông tin định lượng hay định tính thì cũng có một số bài học chung liên quan tới việc phân tích, báo cáo, và phổ biến kết quả có thể được rút ra từ các ví dụ tình huống trong Phụ lục 1.

Thứ nhất, phân tích thường kéo dài hơn dự kiến, đặc biệt nếu như dữ liệu không được làm sạch hay có thể tiếp cận được ngay từ giai đoạn đầu trong phân tích, nếu các nhà phân tích không có kinh nghiệm với dạng công việc đánh giá này, hay nếu có sự nhấn mạnh tới việc xây dựng năng lực nhờ các công việc phối hợp. Trong việc xem xét lại các nghiên cứu tình huống được sử dụng cho bài viết này, thời gian phân tích nhanh nhất là xấp xỉ một năm sau khi có dữ liệu, và thời gian dài nhất là gần hai năm. Tình huống ở trong chương 3 minh họa một số trong nhiều bước liên quan tới phân tích và giải thích tại sao công việc phân tích có thể tốn nhiều thời gian hơn dự kiến.

Thứ hai, người quản lý đánh giá cần lập kế hoạch để tạo ra một số sản phẩm như là kết quả của công việc phân tích, và luôn ghi nhớ hai điều. Thứ nhất, phải đảm bảo thời điểm có kết quả gắn với các sự kiện quan trọng khi đưa ra các quyết định về tương lai của dự án, chẳng hạn xem xét lại giữa kỳ, các kỳ bầu cử, hay thời điểm kết thúc giai đoạn thử nghiệm. Thứ hai, đối tượng của các kết quả. Các sản phẩm cần được phân biệt theo đối tượng của nó, bao gồm các nhà hoạch định chính sách của chính phủ, những người quản lý chương trình, các nhà tài trợ, công chúng nói chung, các nhà báo, và các học giả.

Thứ ba, các sản phẩm sẽ có ý nghĩa nhất về mặt chính sách

nếu chúng bao gồm các khuyến nghị rõ ràng và có tính thực tế xuất phát từ phân tích tác động. Các khuyến nghị này có thể được chia thành các ưu tiên ngắn hạn và dài hạn, và khi có thể, cần bao gồm cả các khuyến nghị về ngân sách. Các nhà hoạch định chính sách cần quan tâm tới “kết toán cuối cùng”.

Cuối cùng, cần lập kế hoạch các báo cáo như là một phần trong chiến lược phổ biến kết quả rộng rãi hơn, có thể bao gồm các lần trình bày cho những đối tượng độc giả khác nhau, các bản tin báo chí, phản hồi tới người cung cấp thông tin, và cung cấp thông tin ở trên trang Web. Cần đề cập tới một chiến lược như vậy trong các bước đầu của quá trình hoạch định để đảm bảo chiến lược này đã được tính tới trong ngân sách và các kết quả sẽ đến được với đối tượng độc giả dự định.

---

---

## Chương 3

# Áp dụng các phương pháp phân tích để đánh giá tác động: Một nghiên cứu tình huống\*

Tình huống nghiên cứu này dựa trên một chương trình chống nghèo đói giả thuyết gọi là PROSCOL, chương trình chuyển một khoản tiền cho các gia đình nghèo có trẻ em trong độ tuổi đi học ở một khu vực trong một nước đang phát triển nhất định. Tình huống nghiên cứu này nhằm minh họa cho các bước phân tích liên quan tới việc tiến hành đánh giá tác động và các lựa chọn mà một nhà phân tích có thể phải đối mặt, tiến trình này cũng có thể áp dụng cho bất kỳ một chương trình chống nghèo đói nào khác. Trong khi khám phá làm thế nào để đánh giá tác động của chương trình, nhà phân tích chính sách thường mắc một số sai lầm chung, tìm kiếm đầu vào trong những chủ đề cụ thể từ những đồng nghiệp có kỹ năng chuyên môn - có thể là từ một nhà thống kê, một nhà kinh tế, một giáo sư kinh tế lượng, và một nhà xã hội học.

Trong số các bước phân tích mà nhà phân tích cần thực hiện trong trường hợp này, có việc xác định những câu hỏi cần phải trả lời trong đánh giá tác động, đánh giá các nguồn dữ liệu, xem xét ban đầu về dữ liệu, hiểu được các sai lệch (bias), hiểu thêm về thu nhập bô lõ, bổ sung các biến số kiểm chứng, hiểu lược tầm quan trọng của các biến số ngoại sinh, khám phá các cách tốt hơn để thành lập một nhóm đối chứng (tương hợp mức xu hướng), hiểu thêm về sai lệch do những biến số không quan sát được, xem xét lại những gì có thể làm với một cuộc

---

\* Chương này căn cứ nhiều vào một tài liệu cơ sở của Martin Ravallion, *Bí  
kết của các lợi ích bị biến mất: Giới thiệu về đánh giá của nhà phân tích nhanh*,  
tài liệu nghiên cứu chính sách số, 1999.

điều tra cơ sở (khác biệt kép), sử dụng các biến số công cụ, kiểm định các phương pháp khác nhau, kết hợp các đầu vào từ thực địa, và lập kế hoạch cho công việc tương lai.

### **Mô tả về chương trình giả thuyết, PROSCOL**

Chương trình PROSCOL xác định các gia đình đủ tư cách tham dự chương trình bằng các biến số đại diện cho tình trạng nghèo đói khác nhau, bao gồm số lượng người trong gia đình, trình độ học vấn của người chủ gia đình, và các thuộc tính đa dạng của nhà ở. PROSCOL trả một khoản tiền cố định cho mỗi trẻ em trong độ tuổi đi học cho tất cả các hộ gia đình được lựa chọn, với điều kiện là trẻ em tham dự 85 phần trăm các lớp học tại trường, và sự tham dự này phải được xác nhận bằng nhận xét từ nhà trường. Các hộ gia đình phải cho trẻ em đi học cho tới khi 18 tuổi.

Chương trình này được đưa ra 12 tháng trước đây, do Ngân hàng Thế giới tài trợ và do Bộ Phát triển xã hội điều hành. Trong một nỗ lực nhằm đánh giá tác động của PROSCOL tới nghèo đói nhằm xác định liệu có nên mở rộng chương trình này để bao phủ phần còn lại của đất nước hay không từ bỏ nó, Ngân hàng Thế giới đã yêu cầu Bộ Tài chính tiến hành đánh giá tác động của chương trình. Yêu cầu này được đặt ra đối với Bộ Tài chính để đảm bảo một sự đánh giá độc lập và để giúp xây dựng năng lực cho dạng đánh giá này trong một cơ quan trung tâm của chính phủ- cơ quan gần gũi với nơi thực hiện phân bổ ngân sách.

### **Xác định các câu hỏi cần phải trả lời trong đánh giá tác động**

Bước thứ nhất đối với nhà phân tích tại Bộ Tài chính được giao nhiệm vụ tiến hành đánh giá PROSCOL là xác định cần phải chú ý tới các mục tiêu dự án nào trong đánh giá tác động. Dự án này có hai mục tiêu chính sách: chuyển một khoản tiền mặt

nhằm làm giảm tình trạng nghèo đói hiện tại; và với yêu cầu là, những người nhận được chuyển khoản phải cho con cái đi học, chương trình này nhằm mục tiêu làm giảm tình trạng nghèo đói trong tương lai bằng cách nâng cao trình độ học vấn trong số các trẻ em nghèo hiện tại. Do vậy cần có hai phần thông tin để đánh giá tác động. Thứ nhất là, liệu các khoản chuyển tiền mặt có chủ yếu nhắm vào các hộ gia đình có thu nhập thấp hay không? Và thứ hai, chương trình này làm tăng tỷ lệ nhập học lên bao nhiêu?

### Đánh giá các nguồn dữ liệu

Để tiến hành đánh giá, nhà phân tích có hai nguồn thông tin chính. Nguồn thứ nhất là báo cáo dựa trên các phỏng vấn định tính với những người thực hiện chương trình và với các nhóm tập trung, gồm những người tham dự chương trình. Tuy vậy, chúng ta vẫn không rõ liệu những người được phỏng vấn có tính đại diện cho những người tham dự PROSCOL hay không, hay những người này có mức độ nghèo đói tương đối như thế nào so với những người không được lựa chọn tham gia chương trình, và do đó, không được phỏng vấn. Báo cáo này cho biết, trẻ em đã đi học, nhưng có thể chúng cũng vẫn sẽ đi học cho dù không có chương trình này. Mặc dù báo cáo này là một bước khởi đầu quan trọng, nhưng nó không cho nhà phân tích biết mức độ nghèo đói của những người tham dự PROSCOL và liệu chương trình này có tác động gì tới việc đi học không. Nguồn thông tin thứ hai là điều tra hộ gia đình toàn quốc được tiến hành độc lập gần đây của Cơ quan Thống kê của quốc gia, được gọi là Điều tra Mức sống (LSS). LSS gồm một mẫu ngẫu nhiên có 10.000 hộ gia đình và đặt ra các câu hỏi về thu nhập hộ gia đình chia theo nguồn gốc, việc làm, chi tiêu, tình trạng sức khoẻ, trình độ học vấn, nhân khẩu học và các thuộc tính khác của gia đình. Điều tra này có một câu hỏi là, hộ gia đình nằm trong mẫu có tham gia PROSCOL hay không, và nó có một

dòng khoản mục để theo dõi số tiền các hộ gia đình nhận được từ PROSCOL trong phần liệt kê các nguồn thu nhập.

### Xem xét ban đầu về dữ liệu

Nhà phân tích sau đó sẽ tiếp tục với những dữ liệu LSS thô để tập trung vào việc đánh giá xem những ai sẽ được hưởng lợi từ chương trình. Bà ta sử dụng một phần mềm thống kê như SPSS hay SAS để tạo ra một bảng chéo (cross-tabulation) của số tiền trung bình nhận được từ PROSCOL từ những nhóm phần mươi (decile) các hộ gia đình, trong đó những nhóm phần mươi này được tạo nên bằng cách sắp xếp tất cả các hộ gia đình trong mẫu căn cứ theo thu nhập trên đầu người của họ. Để tính toán thu nhập trên đầu người, nhà phân tích quyết định trừ đi bất kỳ số tiền nào nhận được từ PROSCOL như là một thước đo thu nhập hợp lý trong trường hợp không có chương trình, với mục đích xác định xem những ai có lợi từ chương trình, tuỳ thuộc vào thu nhập từ trước khi có chương trình của người đó.

Cách lập bảng chéo cho thấy, sự chuyển nhượng tiền mặt trong chương trình được định hướng tương đối tốt tới người nghèo. Theo ngưỡng nghèo chính thức của quốc gia, có khoảng 30 phần trăm dân số sống ở vùng Tây Bắc là người nghèo. Từ bảng này, các tính toán cho thấy, 30 phần trăm số nghèo nhất trong mẫu điều tra nhận được 70 phần trăm các chuyển nhượng của PROSCOL. Thoạt nhìn thì dường như đây là một kết quả tích cực.

Câu hỏi tiếp theo là tác động tới việc đi học. Có thể thấy tác động này thông qua bảng chéo về tỷ lệ nhập học trung bình của các nhóm tuổi khác nhau của các gia đình tham gia PROSCOL so với các gia đình không tham gia PROSCOL. Kết quả cho thấy, gần như không có sự khác biệt giữa hai nhóm này: tỷ lệ nhập học trung bình của trẻ em lứa tuổi từ 6 đến 18 xếp xỉ bằng 80 phần trăm trong cả hai trường hợp. Nhà phân tích sau đó tính toán số năm đi học trung bình ở từng độ tuổi, và các kết quả được vẽ thành đồ thị điểm (plotted) một cách

riêng biệt cho các gia đình tham gia PROSCOL và các gia đình không tham gia. Kết quả cho thấy, cả hai đồ thị không giống nhau hoàn toàn, nhưng rất gần với nhau. Ở giai đoạn này, nhà phân tích tự hỏi, có phải chương trình không hề có tác động gì tới việc đi học, hay là có sai lầm trong phương pháp sử dụng.

### Hiểu được các sai lệch

Với sự không chắc chắn này, nhà phân tích sau đó tìm kiếm sự hỗ trợ từ một chuyên gia thống kê cao cấp để khám phá ra tại sao các kết quả cho thấy trẻ em tham gia PROSCOL không có xu hướng đi học nhiều hơn các trẻ em không tham gia PROSCOL. Nhà thống kê đưa ra giả thuyết là, các kết quả này có một sai lệch nghiêm trọng. Để đánh giá tác động của chương trình, chúng ta cần phải biết những gì có lẽ đã xảy ra nếu như không có chương trình. Tuy vậy, nhà phân tích đã chưa tính đến điểm này; thay vào đó, các gia đình không tham gia PROSCOL được sử dụng như là nhóm đối chứng để rút ra kết luận về việc học hành của những người tham dự PROSCOL sẽ ra sao nếu như không có chương trình này.

Nói cách khác,  $P_i$  đại diện cho sự tham dự PROSCOL về đứa trẻ thứ  $i$ . Biến này có thể nhận hai giá trị, cụ thể là  $P_i = 1$  nếu đứa trẻ tham gia PROSCOL và  $P_i = 0$  nếu như nó không tham gia. Nếu đứa trẻ thứ  $i$  không tham gia thì mức độ học hành của nó sẽ bằng  $S_{0i}$ , viết tắt cho mức độ học hành  $S$  của đứa trẻ thứ  $i$  khi  $P = 0$ . Nếu đứa trẻ tham gia PROSCOL thì mức độ học hành bằng  $S_{1i}$ . Lợi ích từ việc học hành do PROSCOL là  $S_{1i} - S_{0i}$ . Lợi ích của đứa trẻ thứ  $i$  tham gia chương trình ( $P=1$ ) khi đó bằng:

$$G_i = S_{1i} - S_{0i} | P_i = 1.$$

Ký hiệu đại diện cho “với” hay “với điều kiện là” và cần thiết để làm rõ kết quả tính toán chính là lợi ích của đứa trẻ thực sự tham gia chương trình. Nếu chúng ta muốn biết về lợi ích trung bình, thì lợi ích này chỉ đơn giản bằng trung bình tất

cả các giá trị  $G$ , và đây chính là lợi ích đi học của trung bình mẫu trong số tất cả những người tham gia PROSCOL. Chứng nào giá trị trung bình này được tính toán chính xác (sử dụng các trọng số mẫu phù hợp từ điều tra), nó sẽ cung cấp một ước lượng không chênh của lợi ích trung bình thực sự. Giá trị này là “giá trị kỳ vọng” của  $G$ , và có thể viết nó như sau:

$$G = E(S_{1i} - S_{0i} | P_i = 1).$$

Đây là một cách nói khác thay cho từ “trung bình”. Tuy vậy, nó không hoàn toàn bằng với giá trị trung bình được tính từ dữ liệu của mẫu, do sẽ có một số sai lầm chọn mẫu. Trong các tài liệu đánh giá,  $G = E(S_{1i} - S_{0i} | P_i = 1)$  đôi khi được gọi là “tác động nghiên cứu” hay “tác động nghiên cứu trung bình tới các đối tượng tham dự.” Trong trường hợp này, PROSCOL được xem là chương trình tác động.

Nhà thống kê chỉ ra cho nhà phân tích là bà ta đã không tính đến  $G$ , chính là sự khác biệt giữa mức độ đi học trung bình giữa các trẻ em trong các gia đình tham dự PROSCOL và những trẻ em trong các gia đình không tham dự. Giá trị này là ước lượng mẫu của

$$D = E(S_{1i} | P = 1) - E(S_{0i} | P = 0).$$

Có một đồng nhất thức đơn giản liên kết  $D$  và  $G$ , cụ thể:

$$D = G + B.$$

Thuật ngữ “ $B$ ” là sai lệch trong ước lượng, và nó bằng

$$B = E(S_{0i} | P_i = 1) - E(S_{0i} | P_i = 0).$$

Nói cách khác, sai lệch chính bằng chênh lệch dự kiến trong mức độ đi học không có sai lệch PROSCOL giữa những trẻ em thực tế tham dự chương trình với những trẻ em không tham dự.

Sai lệch này có thể được sửa chữa nếu như biết  $E(S_{0i} | P_i = 1)$ , nhưng chúng ta không thể có được ước lượng mẫu của giá trị này. Chúng ta không thể quan sát được mức độ đi học của những trẻ em thực sự tham dự PROSCOL nếu như chúng không tham dự; đây là các dữ liệu bị bỏ sót - và các dữ liệu này cũng được gọi là trung bình của "phản chứng".

Các sai lệch này gây ra một mối lo ngại quan trọng. Trong trường hợp không có chương trình, các bậc cha mẹ của những gia đình tham dự PROSCOL có thể sẽ gửi trẻ em tới trường ít hơn các bậc cha mẹ khác. Nếu như vậy thì đã xảy ra có sai lệch trong tính toán. Quay trở lại các câu hỏi đánh giá ban đầu, chúng ta quan tâm tới mức độ đi học tăng lên do PROSCOL. Có lẽ PROSCOL chỉ tác động tới những gia đình thực sự tham dự chương trình. Nói cách khác, chúng ta cần biết mức độ đi học sẽ thấp hơn bao nhiêu nếu như không có chương trình. Nếu không có sai lệch thì mức độ đi học tăng lên theo chương trình bằng chênh lệch giữa mức độ đi học trung bình giữa những người tham dự và những người không tham dự chương trình. Do vậy, sai lệch này sinh nếu như có khác biệt trong mức độ đi học trung bình giữa các bậc cha mẹ tham gia PROSCOL và các bậc cha mẹ không tham gia PROSCOL trong trường hợp không có chương trình.

Để loại trừ sai lệch này, phương pháp tốt nhất là phân công chương trình một cách ngẫu nhiên. Sau đó, những người tham dự và những người không tham dự sẽ có cùng mức độ đi học dự kiến trong trường hợp không có chương trình, tức là  $E(S_{0i} | P_i = 1) = E(S_{0i} | P_i = 0)$ . Mức độ đi học của những gia đình không tham dự sau đó sẽ cho biết một cách chính xác về phản chứng, tức là mức độ đi học mà chúng ta lẽ ra quan sát được ở các đối tượng tham dự nếu như họ không tham dự chương trình. Trên thực tế, sự phân công ngẫu nhiên sẽ bình quân hoá toàn bộ phân phối, chứ không chỉ các giá trị trung bình. Sẽ vẫn có sai lệch do sai lầm chọn mẫu, nhưng với các mẫu đủ lớn, chúng ta có thể giả định một cách an toàn là bất kỳ sai khác có ý nghĩa thống kê nào trong phân phối mức độ đi học giữa những người

tham dự và không tham dự chương trình đều có nguyên nhân từ chương trình.

Trong phạm vi thiết kế hiện tại của chương trình, rõ ràng sự tham dự không có tính ngẫu nhiên. Trên thực tế, đây sẽ là một chỉ trích nghiêm trọng về PROSCOL nếu như chương trình này lựa chọn sự tham dự một cách ngẫu nhiên. Thực tế là chương trình này nhằm mục đích hướng tới các hộ gia đình nghèo, những người được xem là có xu hướng gửi con cái đi học ít hơn, và vì vậy, gây ra sai lệch.

Do vậy, nảy sinh câu hỏi là, nếu như PROSCOL hoạt động tốt thì chúng ta cần dự kiến những người tham dự sẽ có mức độ đi học thấp hơn nếu như không có chương trình. Khi đó thì  $E(S_{0i} | P_i = 1) < E(S_{0i} | P_i = 0)$  và tính toán ban đầu của nhà phân tích sẽ ước lượng thấp lợi ích nhận được từ chương trình. Chúng ta có thể tìm thấy rất ít hay không tìm thấy lợi ích nào cho dù chương trình thực sự làm việc tốt.

Bây giờ, nhà phân tích nhận thấy mức độ của sai lệch này có thể rất lớn. Giả định là các gia đình nghèo gửi con đi làm thay vì tới trường; vì họ nghèo và không thể vay mượn một cách dễ dàng nên họ cần có thêm tiền ngay. Các hộ gia đình không nghèo thì gửi con tới trường. Chương trình lựa chọn các gia đình nghèo, những người này sau đó gửi con đến trường. Chúng ta quan sát thấy sự khác biệt không đáng kể trong mức độ đi học trung bình giữa các gia đình tham gia PROSCOL và các gia đình không tham gia; trên thực tế  $E(S_{1i} | P_i = 1) = E(S_{0i} | P_i = 0)$  theo dự kiến. Nhưng tác động của chương trình có tính tích cực và bằng  $E(S_{0i} | P_i = 0) - E(S_{0i} | P_i = 1)$ . Từ sự so sánh của nhà phân tích về mức độ đi học trung bình giữa các gia đình tham gia PROSCOL và các gia đình không tham gia, chúng ta thấy việc không tính đến mục tiêu định hướng tới người nghèo của chương trình đã làm suy giảm đáng kể lợi ích của PROSCOL.

### Hiểu thêm về thu nhập bỏ lỡ

Nhà phân tích sau đó chỉ các kết quả trong bảng chéo về số

tiền nhận được từ PROSCOL so với thu nhập cho một đồng nghiệp khác là một nhà kinh tế học công tác tại Bộ tài chính. Nhà kinh tế học nêu ra mối quan tâm chủ yếu- đó là các lợi ích mà người nghèo nhận được từ PROSCOL rõ ràng đã được ước lượng cao quá do bỏ qua thu nhập bỏ lỡ. Trẻ em phải đến trường nếu các gia đình muôn nhận được chuyển nhượng từ PROSOL; do vậy chúng sẽ không thể làm việc trong công việc kinh doanh của gia đình hay ở trên thị trường lao động. Ví dụ, trẻ em ở độ tuổi từ 15 đến 18 có thể nhận được hai phần ba hay nhiều hơn tiền công của người lớn làm việc trong lĩnh vực nông nghiệp và xây dựng. Các gia đình tham gia PROSCOL sẽ bị mất các khoản thu nhập thu được từ việc làm của trẻ em. Khoản thu nhập tất nhiên này cần phải được tính đến khi tính toán các lợi ích thu nhập ròng nhận được từ chương trình. Và cần xác định được mức thu nhập trước khi có chương trình bằng cách khấu trừ lợi ích thu nhập ròng, chứ không phải là khoản chuyển nhượng tổng cộng. Điều này cũng quan trọng trong việc quyết định hộ gia đình sẽ nghèo như thế nào nếu như không có chuyển nhượng PROSCOL. Do vậy, bảng hiện tại có thể phóng đại các lợi ích của chương trình đối với người nghèo.

Nhà phân tích tự hỏi tại sao cần phải quan tâm tới thu nhập tất nhiên từ lao động trẻ em, một khi giả định là càng có ít lao động trẻ em càng tốt. Nhà kinh tế nhấn mạnh cần quan tâm tới các lợi ích nhận được từ việc giảm lao động trẻ em, trong đó lợi ích chính là tăng mức độ đi học, và do vậy thu nhập tương lai cao hơn, đối với các hộ gia đình hiện đang nghèo. Nhà phân tích đã tạo ra các bảng phản ánh hai cách thức chính theo đó PROSCOL làm giảm nghèo đói: đó là tăng thu nhập hiện tại của người nghèo và tăng thu nhập trong tương lai của họ. Tác động tới lao động trẻ em quan trọng đối với cả hai cách thức này, nhưng theo những định hướng khác nhau; do vậy, có một sự đánh đổi trong PROSCOL.

Điều này nhấn mạnh tại sao cần phải có một ước tính tốt về tác động tới việc đi học: chỉ khi đó mới có thể xác định được thu nhập bỏ lỡ. Chẳng hạn, đó có thể là thời gian tăng thêm ở

nà trường nhờ thời gian không làm việc.

Về bảng chéo thứ hai, mỗi quan tâm chủ yếu mà nhà kinh tế nêu ra là nó không đề cập tới các nhân tố quyết định việc đi học khác, trừ việc tham gia vào PROSCOL. Nhà kinh tế đề nghị chạy một hồi quy các năm đi học theo một tập hợp các biến kiểm chứng cũng như một biến phản ánh gia đình đứa trẻ có tham gia PROSCOL hay không. Đối với đứa trẻ thứ  $i$  trong mẫu, ta có

$$S_i = a + bP_i + cX_i + \varepsilon_i .$$

Ở đây  $a$ ,  $b$ , và  $c$  là các tham số;  $X$  đại diện cho các biến kiểm chứng, chẳng hạn tuổi của đứa trẻ, trình độ học vấn của mẹ và cha, quy mô và thành phần nhân khẩu học của hộ gia đình, và các đặc điểm của nhà trường; và là phần dư (residual) bao gồm các nhân tố quyết định khác tới việc đi học và các sai số do lường. Giá trị ước lượng của  $b$  cho chúng ta biết tác động của PROSCOL tới việc đi học.

Cần chú ý là, nếu như gia đình của đứa trẻ thứ  $i$  tham dự vào PROSCOL thì  $P = 1$ , và mức độ đi học của đứa trẻ sẽ bằng  $a + b + cX_i + \varepsilon_i$ . Nếu đứa trẻ không tham dự chương trình thì  $P = 0$  và mức độ đi học sẽ bằng  $a + cX_i + \varepsilon_i$ . Chênh lệch giữa hai kết quả này là lợi ích đi học do chương trình này, và lợi ích này đúng bằng  $b$ .

### Bổ sung các biến số đối chứng

Với sự gợi ý này, nhà phân tích liền chạy một phương trình hồi quy có và không có các biến đối chứng. Khi phương trình hồi quy không có biến đối chứng, các kết quả cho thấy giá trị ước lượng của  $b$  không khác biệt về mặt thống kê với không (sử dụng kiểm định  $t$  chuẩn do phần mềm thống kê tính toán). Các kết quả này rất giống với các kết quả ban đầu, với khác biệt trong giá trị trung bình giữa những người tham dự và những người không tham dự - điều này cho thấy PROSCOL không có

tác động gì tới việc đi học. Tuy nhiên, khi có một số biến đổi chứng trong phương trình hồi quy, thì hệ số của sự tham dự PROSCOL đạt giá trị dương và có ý nghĩa. Kết quả tính toán cho thấy, cho đến độ tuổi 18, chương trình làm tăng thời gian đi học lên hai năm.

Nhà phân tích tự hỏi, tại sao các biến đổi chứng lại tạo ra sự khác biệt này? Và liệu đã sử dụng đúng các biến đổi chứng chưa? Bà ta bèn đến thăm vị giáo sư dạy kinh tế lượng trước đây của mình và chỉ cho ông thấy phương trình hồi quy của bà. Mỗi lo ngại đầu tiên của vị giáo sư này về phương trình hồi quy việc đi học theo  $P$  và  $X$  là nó không cho phép tác động của chương trình thay đổi theo  $X$ ; tác động này giống nhau đối với mọi người, và điều này có vẻ như không đúng. Các bậc cha mẹ được học hành nhiều hơn, sẽ có xu hướng gửi con đi học nhiều hơn, và lợi ích của PROSCOL với những người này sẽ thấp hơn. Để cho các lợi ích này thay đổi theo  $X$ , nếu gọi mức độ đi học trung bình của những người không tham dự là  $a_0 + c_0X_i$  và của những người tham dự là  $a_1 + c_1X_i$ , thì mức độ đi học quan sát được là:

$$S_i = (a_1 + c_1X_i + \varepsilon_{1i})P_i + (a_0 + c_0X_i + \varepsilon_{0i})(1 - P_i)$$

Trong đó  $\varepsilon_0$  và  $\varepsilon_1$  là các sai lầm ngẫu nhiên, mỗi sai lầm có trung bình bằng không và không tương quan với  $X$ . Để ước lượng mô hình này, cần bổ sung một đại lượng để phản ánh các tác động tương tác giữa sự tham dự chương trình và các đặc điểm quan sát được, vào phương trình hồi quy đã được sử dụng này. Do đó, phương trình hồi quy gia tăng (augmented regression) là:

$$S_i = a_0 + (a_1 - a_0)P_i + c_0X_i + (c_1 - c_0)P_iX_i + \varepsilon_i$$

trong đó  $\varepsilon_i = \varepsilon_{1i}P_i + \varepsilon_{0i}(1 - P_i)$ . Khi đó  $(a_1 - a_0) + (c_1 - c_0)X$  là tác động trung bình của chương trình tương ứng với một giá trị nhất định của  $X$ . Nếu giá trị  $X$  trung bình trong mẫu những

người tham dự được sử dụng, thì nó sẽ cho biết lợi ích trung bình nhận được từ chương trình.

### **Hiểu được tầm quan trọng của các biến số ngoại sinh**

Sự lo ngại thứ hai được giáo sư kinh tế lượng nêu ra là, làm thế nào để ước lượng phương trình hồi quy. Việc sử dụng lệnh hồi quy trong phần mềm thống kê, phương pháp bình phương thông thường nhỏ nhất (OLS), gây lo ngại, do các ước lượng theo OLS của các tham số sẽ bị sai lệch ngay cả trong các mẫu lớn, trừ khi các biến ở phía tay phải là ngoại sinh. Gọi là ngoại sinh vì các biến phía tay phải được quyết định độc lập với các lựa chọn đi học, và vì vậy, chúng không có tương quan với hạng tử biểu diễn sai lầm trong phương trình hồi quy về việc đi học. Do sự tham dự chương trình được định hướng có mục đích, nên sự tham dự PROSCOL không phải là một biến số ngoại sinh. Điều này có thể tác động tới việc tính toán tác động của chương trình như sau: phương trình số năm đi học là:

$$S_i = a + bP_i + cX_i + \varepsilon_i$$

Giá trị của  $a + b + cX_i + \varepsilon_i$  được sử dụng như là ước lượng cho việc đi học của hộ gia đình thứ  $i$  khi hộ gia đình này tham dự vào PROSCOL, trong khi  $a + cX_i + \varepsilon_i$  được sử dụng để ước lượng việc đi học nếu như hộ gia đình không tham dự. Do đó, chênh lệch  $b$ , là lợi ích thu được từ chương trình. Tuy nhiên, khi tính toán, có một giả định ngầm là  $i$  giống nhau trong cả hai trường hợp. Nói cách khác, giả định này là độc lập với  $P$ , và điều này sẽ tác động tới việc tính toán tác động của chương trình.

Thực tế này nhấn mạnh tới sai lệch do việc áp dụng chương trình không ngẫu nhiên, điều này cũng có thể tác động tới ước lượng dựa trên mô hình hồi quy được nhà kinh tế học gợi ý trước đây ( $S_i = a + bP_i + cX_i + \varepsilon_i$ ). Tuy nhiên, như vậy không có nghĩa là các kết quả sẽ hoàn toàn sai.

Vì giáo sư kinh tế lượng làm rõ điều này bằng một phương trình rõ ràng cho  $P$ , cụ thể là

$$P_i = d + eZ_i + v_i$$

trong đó  $Z$  là một số biến số bao gồm tất cả các "biến đại diện cho nghèo đói" được sử dụng vì mục đích định hướng của PROSCOL. Cũng sẽ có một số hạng tử sai số hoàn toàn ngẫu nhiên có tác động tới sự tham dự; có các biến đại diện cho nghèo đói không có trong dữ liệu, và cũng sẽ có những sai lầm trong lựa chọn các thành viên, dẫn đến hạng tử  $v$ . Phương trình này là tuyến tính, tuy nhiên  $P$  chỉ có thể nhận hai giá trị khả dĩ là 0 và 1. Các giá trị dự đoán nằm giữa 0 và 1 là có thể chấp nhận được, nhưng một mô hình tuyến tính không thể loại trừ khả năng có những giá trị dự đoán âm, hay các giá trị cao hơn 1. Có các mô hình phi tuyến tính để giải quyết vấn đề này, nhưng để đơn giản hóa trong thảo luận, việc giới hạn sự quan tâm vào các mô hình tuyến tính sẽ là dễ dàng nhất.

Có một trường hợp đặc biệt, trong đó phương trình hồi quy theo OLS của  $S$  theo  $P$  và  $X$  sẽ cho một ước lượng không chênh của  $b$ . Đây là khi  $X$  bao gồm tất cả các biến trong  $Z$  cũng có tác động tới việc đi học, và hạng tử sai số  $v$  không có tương quan với hạng tử sai số trong phương trình hồi quy về việc đi học. Trường hợp này đôi khi được gọi là "lựa chọn trên các biến quan sát được" trong các tài liệu về đánh giá.

Giả định là, các biến đối chứng  $X$  trong phương trình hồi quy trước về việc đi học bao gồm tất cả các biến quan sát được  $Z$ , có tác động tới sự tham dự  $P$  và  $v$  không tương quan với  $e$  (vì vậy các biến không quan sát tác động tới việc áp dụng chương trình không tác động tới việc đi học phụ thuộc điều kiện  $X$ ). Khi đó, bất kỳ một khả năng  $P$  có tương quan với  $e$  đều bị loại trừ. Biến số này bây giờ sẽ trở thành ngoại sinh trong phương trình hồi quy về đi học. Nói cách khác, ý tưởng cơ bản của lựa chọn các biến quan sát được là có một số biến quan sát được  $X$  nào đó sao cho sự sai lệch sẽ biến mất tùy vào điều kiện của  $X$ .

Bổ sung các biến đối chứng vào phương trình hồi quy đi học theo sự tham gia PROSCOL sẽ tạo ra sự khác biệt lớn do X phải bao gồm các biến trong số các biến đại diện cho nghèo đói được sử dụng để định hướng, hoặc có tương quan với các biến này, và chúng phải là những biến cũng có tác động tới việc đi học. Tuy nhiên, phương pháp này chỉ có hiệu quả nếu các giả định hợp lệ. Cần phải nhận thức được hai vấn đề. Thứ nhất, phương pháp trên sẽ không phù hợp nếu không có các nhân tố quyết định sự tham dự mà không thể quan sát được; nói cách khác, nếu như hạng tử sai số  $v$  có phương sai bằng không, và tất cả các nhân tố quyết định sự tham dự cũng có tác động tới việc đi học. Khi đó, sẽ không có sự biến thiên độc lập trong việc tham dự chương trình để cho phép chúng ta nhận ra tác động của nó tới việc đi học; có thể dự đoán  $P$  hoàn toàn từ  $X$ , và vì thế sẽ không ước lượng được phương trình hồi quy. Vấn đề này khó có thể phát sinh thường xuyên, bởi vì hầu như luôn luôn có những nhân tố quyết định việc áp dụng chương trình không thể quan sát.

Vấn đề thứ hai phổ biến hơn, và đáng lo ngại hơn trong trường hợp này. Hạng tử sai số  $\epsilon$  trong phương trình hồi quy đi học có lẽ bao gồm cả những biến không được phát hiện trong điều tra nhưng có thể tác động lớn tới sự tham dự chương trình, và vì thế, chúng có thể tương quan với hạng tử sai số trong phương trình tham dự. Nếu đây là trường hợp xảy ra thì  $E(\epsilon|X,P) \neq 0$ , và các phương pháp hồi quy thông thường sẽ vẫn bị sai lệch khi các phương trình hồi quy về đi học được ước lượng. Do vậy, vấn đề chính là mức độ tương quan giữa hạng tử sai số trong phương trình tham dự và trong phương trình đi học.

### **Khám phá các cách tốt hơn để thành lập một nhóm đối chứng tương hợp mức độ xu hướng**

Với sự hỗ trợ thêm từ vị giáo sư, nhà phân tích biết được có những cách tốt hơn để thành lập một nhóm đối chứng. Mục

tiêu là, so sánh trình độ học tập căn cứ vào các đặc điểm quan sát được. Nếu các nhóm mẫu được chia thành nhóm các hộ gia đình có giá trị X giống nhau hay tương tự nhau, chúng ta sẽ so sánh các giá trị trung bình có điều kiện của các hộ gia đình tham gia và không tham gia PROSCOL. Nếu việc đi học khi không có chương trình độc lập với sự tham dự, thì với X đã cho, so sánh sẽ cho ta một ước lượng không chênh về tác động của PROSCOL. Đôi khi trường hợp này được gọi là “độc lập có điều kiện,” và đây chính là giả định quan trọng của tất cả các phương pháp sử dụng nhóm đối chứng.

Do vậy, một cách tốt hơn để lựa chọn một nhóm đối chứng, với các dữ liệu có sẵn, là sử dụng nhóm này như một hình thức kiểm chứng đối với mỗi người tham dự bằng một người không tham dự cũng có những đặc điểm quan sát được. Tuy nhiên, công việc này có thể rất khó khăn, bởi vì tập hợp dữ liệu có thể có rất nhiều biến. Có thể là không có người nào trong số những người không tham dự có cùng giá trị trong tất cả các đặc điểm quan sát được giống như một người nào đó tham dự PROSCOL.

Một phương pháp thống kê có tên là tương hợp mức độ xu hướng, cung cấp các kỹ thuật để làm đơn giản hóa vấn đề một cách đáng kể. Thay vì nhầm vào việc đảm bảo đối tượng kiểm chứng tương hợp của mỗi thành viên tham dự đều có cùng giá trị X với thành viên này, thì có thể đạt được kết quả này bằng cách làm cho phù hợp trong giá trị dự đoán được của  $P$ , với X cho trước, mà đôi khi được gọi là mức độ xu hướng của X. Rosenbaum và Rubin (1983) chỉ ra rằng, nếu như (trong trường hợp này) việc đi học mà không có PROSCOL độc lập với sự tham dự, với X cho trước, thì những người tham dự cũng độc lập với sự tham dự, với mức độ xu hướng của X nhất định. Bởi vì mức độ xu hướng chỉ là một con số, nên sẽ dễ kiểm chứng hơn là X- có thể bao gồm nhiều biến số - rất nhiều. Và mặc dù vậy, phương pháp tương hợp mức độ xu hướng đủ khả năng loại trừ được sai lệch, miễn là có sự độc lập có điều kiện với X cho trước.

Nói cách khác, chúng ta trước hết sẽ hồi quy  $P$  theo X để

nhận được giá trị dự đoán của  $P$  với mỗi giá trị có thể của  $X$ , sau đó sẽ ước lượng giá trị này với toàn bộ mẫu. Đối với mỗi người tham dự, chúng ta cần phát hiện được những người không tham dự có giá trị xác suất dự đoán được gần nhất. Sự khác biệt trong tỷ lệ đi học khi đó bằng lợi ích ước lượng được từ chương trình đối với người tham dự này.

Sau đó, chúng ta có thể lấy trung bình của tất cả các chênh lệch này để ước lượng tác động. Hoặc lấy trung bình đối với các nhóm thu nhập khác nhau. Tuy nhiên, cần phải có sự thận trọng trong việc ước lượng mô hình tham dự. Một mô hình tuyến tính có thể cho các xác suất dự đoán được trở nên bất thường, cao hơn một hoặc âm. Tốt hơn là nên sử dụng lệnh LOGIT trong phần mềm thống kê. Lệnh này giả định rằng, hạng tử sai số  $v$  trong phương trình tham dự có phân phối logarit, và ước lượng các tham số nhất quán với giả định này do các phương pháp khả năng tối đa (maximum likelihood methods). Việc ước lượng này dựa trên các nguyên tắc ước lượng khả năng tối đa của các mô hình phản ứng nhị phân (binary response models).

Một vấn đề nữa cần phải nhận thức được là việc một số người không tham dự có thể bị loại trừ như là những đối tượng tương hợp tiềm tàng ngay từ đầu. Trên thực tế, một số kết quả gần đây trong các tài liệu nghiên cứu về kinh tế lượng chỉ ra rằng, thất bại trong so sánh những người tham dự và kiểm chứng theo các giá trị thông thường đối với các biến tương hợp là một nguồn gốc chính gây ra sai lệch trong đánh giá (xem Heckman và các tác giả khác 1998).

Theo trực giác, chúng ta muốn nhóm đối chứng giống nhóm tham dự càng nhiều càng tốt về đặc điểm của các biến quan sát được, như được tóm tắt lại trong mức độ xu hướng. Chúng ta có thể nhận thấy, một số mẫu không tham dự có mức độ xu hướng thấp hơn bất kỳ người tham dự nào trong mẫu tham dự. Thực tế này đôi khi được gọi là hiện tượng "thiếu sự hỗ trợ chung". Khi thành lập nhóm đối chứng, chúng ta cần phải loại các quan sát này trong tập hợp những người không tham dự

chương trình để đảm bảo chỉ có những lợi ích nằm trong phạm vi các mức độ xu hướng tương tự được sử dụng để so sánh. Chúng ta cũng cần loại khỏi nhóm không tham dự những người có xác suất tham dự bằng không. Có lẽ nên lọc ra một tỷ lệ nhỏ trong mẫu, chẳng hạn 2 phần trăm, những người không tham dự có mức độ xu hướng cao nhất và thấp nhất. Một khi đã xác định được những người tham dự và không tham dự trên một khu vực tương hợp chung, thì cần lấy trung bình của (ví dụ) khoảng năm đối tượng gần nhất căn cứ vào chênh lệch tuyệt đối trong mức độ xu hướng (hộp 3.1).

Tiếp theo, cần đưa vào tất cả các biến số trong bộ dữ liệu, là hoặc có thể, đại diện cho các chỉ số nghèo đói được sử dụng cho việc lựa chọn các thành viên tham gia PROSCOL. Một lần nữa, X cần phải có các biến trong Z. Tuy nhiên, điều này tạo ra một điểm yếu trong phương pháp tương hợp mức độ xu hướng. Với phương pháp tương hợp, một X khác biệt có thể cho một ước lượng tác động khác. Với phương pháp lựa chọn ngẫu nhiên và là trường hợp thử nghiệm lý tưởng, các kết quả không phụ thuộc vào việc chọn X nào. Lựa chọn ngẫu nhiên cũng không đòi hỏi chúng ta phải nêu rõ một mô hình tham dự, cho dù mô hình này có phải là logit hay là một dạng mô hình khác. Hộp 3.1 tóm tắt các bước để thực hiện tương hợp mức độ xu hướng.

### Học hỏi về các sai lệch do các biến số không quan sát được.

Thậm chí sau khi đã tạo được nhóm đối chứng, nhà phân tích vẫn không thể chắc chắn là việc này sẽ cho phép có một ước lượng tốt hơn nhiều về tác động của chương trình. Các phương pháp được mô tả ở trên sẽ chỉ loại trừ được sai lệch nếu có sự độc lập có điều kiện, làm cho các nhân tố quyết định không quan sát được của việc đi học - không được đưa vào trong tập hợp các biến đối chứng X- không tương quan với sự tham gia chương trình. Có hai nguyên nhân sai lệch khác nhau: sự khác

### Hộp 3.1 Các bước trong phương pháp tương hợp mức độ xu hướng

Mục tiêu của tương hợp là tìm ra nhóm đối chứng gần nhất từ một mẫu gồm những người không tham dự với mẫu gồm những người tham dự chương trình. Khái niệm “gần nhất” được đo lường theo các đặc điểm tiềm tàng. Các bước chính trong tương hợp dựa trên các mức độ xu hướng như sau:

**Bước 1:** Bạn cần có một điều tra chọn mẫu có tính đại diện về những người không tham dự đủ tiêu chuẩn, cũng như cho những người tham dự. Mẫu bao gồm những người không tham dự càng lớn thì càng tốt, do khả năng tương hợp sẽ dễ dàng hơn. Nếu hai mẫu được lấy từ các cuộc điều tra khác nhau, thì chúng cần phải là các cuộc điều tra có thể so sánh được (cùng bảng câu hỏi, cùng người phỏng vấn, hay cùng chương trình đào tạo phỏng vấn, cùng thời kỳ phỏng vấn, vv).

**Bước 2:** Gộp hai mẫu với nhau và ước lượng một mô hình logit về sự tham dự chương trình như là một hàm của tất cả các biến trong dữ liệu có thể quyết định sự tham dự này.

**Bước 3:** Tạo ra các giá trị có thể dự đoán về xác suất tham dự từ phương trình hồi quy logit; chúng được gọi là “các mức độ xu hướng.” Bạn sẽ có một mức độ xu hướng cho bất kỳ một người tham dự hay không tham dự được chọn mẫu.

**Bước 4:** Một số người trong mẫu, gồm những người không tham dự có thể bị loại ngay từ đầu, do họ có mức độ xu hướng nằm ngoài phạm vi (thường là quá thấp) của mẫu gồm những đối tượng tham dự chương trình. Phạm vi của các mức độ xu hướng được ước lượng cho nhóm tham dự cần tương ứng chặt chẽ với nhóm phụ còn lại trong số những người không tham dự. Bạn cũng có thể

(Xem tiếp trang sau)

### Hộp 3.1 (Tiếp)

hạn chế cặp tương hợp tiềm tàng theo những cách khác, tùy thuộc vào bối cảnh. Ví dụ, bạn có thể muốn chỉ cho phép có các cặp tương hợp trong cùng một khu vực địa lý để đảm bảo các cặp tương hợp này xuất phát từ cùng một môi trường kinh tế.

**Bước 5:** Đối với mỗi cá nhân trong mẫu tham dự, bây giờ bạn muốn tìm ra những quan sát trong nhóm không tham dự mà có mức độ xu hướng gần nhất, như được đo lường bằng sự chênh lệch tuyệt đối về mức độ xu hướng. Quan sát này được gọi là “hàng xóm gần nhất.” Bạn có thể tìm thấy năm (giả dụ như vậy) hàng xóm gần nhất.

**Bước 6:** Tính toán giá trị trung bình của chỉ số kết quả (hay của mỗi chỉ số nếu như có nhiều hơn một chỉ số) cho năm hàng xóm gần nhất. Sự chênh lệch giữa trung bình và giá trị thực tế của quan sát tham dự chương trình là ước lượng lợi ích từ chương trình của quan sát đó.

**Bước 7:** Tính toán giá trị trung bình của những lợi ích cá nhân để có được lợi ích toàn bộ trung bình. Có thể phân lớp mức lợi ích này theo một số biến quan tâm, như thu nhập, trong mẫu không tham dự.

Đây là hình thức tương hợp mức độ xu hướng đơn giản nhất. Có thể phát sinh những vấn đề phức tạp trên thực tế. Chẳng hạn, nếu như có hiện tượng chọn mẫu quá nhiều đối với những người tham dự, bạn có thể sử dụng các phương pháp chọn mẫu dựa trên sự lựa chọn để sửa chữa (Manski và Lerman 1977); mặt khác, bạn có thể sử dụng hệ số cơ hội xảy ra (bằng  $p/(1-p)$ , trong đó  $p$  là mức độ xu hướng) để thực hiện tương hợp. Thay vì dựa vào hàng xóm gần nhất, bạn có thể sử dụng tất cả những người không tham dự như là các đối tượng tương hợp tiềm tàng, nhưng cần đặt các trọng số khác nhau, tùy theo mức độ gần gũi của các đối tượng này (Heckman và các tác giả khác 1998).

bíệt trong các biến quan sát và khác biệt trong các biến không quan sát; trường hợp sau thường được gọi là “sai lệch lựa chọn.” Hộp 3.2 mô tả chi tiết về sự khác nhau này.

### Hộp 3.2 Các nguyên nhân gây ra sai lệch trong ước lượng đơn giản về tác động của PROSCOL

Sự sai lệch được nhà thống kê mô tả là sự chênh lệch dự kiến trong tỷ lệ đi học khi không có PROSCOL giữa các gia đình được lựa chọn tham dự chương trình và những người không được lựa chọn. Sự sai lệch này có thể do hai loại nguyên nhân khác nhau:

- Sai lệch do khác biệt trong các đặc điểm quan sát được. Sai lệch này có thể xảy ra do hai nguyên nhân. Thứ nhất, có thể không có sự hỗ trợ chung. “Sự hỗ trợ” là tập hợp các giá trị của các biến đối chứng, mà nhờ đó, các kết quả và sự tham dự chương trình được quan sát. Nếu sự hỗ trợ giữa mẫu tham dự và nhóm đối chứng khác nhau thì khi đó sẽ có sự sai lệch trong các kết quả. Thứ hai, ngay cả với sự hỗ trợ chung, thì sự phân phối các đặc điểm quan sát được có thể khác nhau trong khu vực có chung sự hỗ trợ; đó là khi trên thực tế, dữ liệu của nhóm đối chứng bị áp dụng các trọng số sai. Sự lựa chọn kỹ lưỡng nhóm đối chứng có thể loại trừ được nguyên nhân sai lệch này.
- Sai lệch do khác biệt trong các biến không quan sát được. Thuật ngữ sai lệch do lựa chọn đôi khi chỉ được giới hạn cho nguyên nhân sai lệch này (mặc dù một số tác giả sử dụng thuật ngữ này để chỉ sai lệch toàn bộ trong một đánh giá không qua thử nghiệm). Nguyên nhân của sai lệch này nảy sinh khi, với giá trị X nhất định, có mối liên

### Hộp 3.2 (Tiếp)

hệ thống giữa sự tham dự chương trình và các kết quả khi không có chương trình. Nói cách khác, đây là những biến không quan sát mà tác động đồng thời cả sự di học và sự tham dự chương trình, căn cứ vào các biến quan sát trong dữ liệu.

Không có gì đảm bảo hai nguyên nhân của sai lệch này sẽ xảy ra cùng chiều. Do đó, việc loại trừ chỉ một trong 2 nguyên nhân không có nghĩa là sai lệch tổng thể được giảm xuống về giá trị tuyệt đối. Đây là một vấn đề thực nghiệm. Tại một trong số ít các nghiên cứu nhằm trả lời câu hỏi này, tác động thực sự, được đo lường bằng một thử nghiệm được thiết kế kỹ, được so sánh với các ước lượng không qua thử nghiệm (Heckman và các tác giả khác 1998). Sai lệch trong ước lượng đơn giản là rất lớn, nhưng việc tiến hành tương hợp một cách kỹ lưỡng nhóm đối chứng dựa trên các biến quan sát đã làm giảm mạnh sai lệch này.

Quay trở lại phương trình cuối cùng của vị giáo sư ta thấy, sự độc lập theo điều kiện sẽ chỉ có được nếu  $P$  là biến ngoại sinh, vì khi đó  $E(\varepsilon_i | X_i, P_i) = 0$ . Tuy nhiên, việc tham dự chương trình có tính nội sinh do mục tiêu định hướng dựa trên các biến không quan sát sẽ vẫn tạo ra sai lệch. Công việc này đôi khi được gọi là lựa chọn các biến quan sát. Do đó, điều kiện cần thiết để giải thích cho phương pháp được nhà kinh tế nêu lên trước đây không kém kém hạn chế hơn các điều kiện cần thiết để lý giải cho một dạng của phương pháp thử nhất, phương pháp này căn cứ trên việc so sánh các gia đình tham dự PROSCOL với các gia đình không tham gia PROSCOL đối với các hộ gia đình có giá trị  $X$  tương đương. Cả hai phương pháp này đều dựa trên niềm tin là các biến không quan sát

không có tác động tới cả việc đi học và việc tham dự chương trình, với điều kiện X nhất định.

Một cách trực giác, chúng ta có thể thấy việc tiến hành tương hợp một cách kỹ lưỡng làm giảm sai lệch, nhưng không nhất thiết là như vậy. Tương hợp loại trừ một phần sai lệch trong ước tính đơn giản ban đầu về tác động của PROSCOL. Nhưng vẫn còn có sai lệch do các biến không quan sát gây ra. Tuy nhiên, hai nguyên nhân sai lệch này có thể bù trừ cho nhau - một tích cực và một tiêu cực. Heckman và các tác giả khác (1998) đã nêu ra điểm này. Vì vậy, ước lượng tương hợp có thể có nhiều sai lệch hơn là ước lượng đơn giản. Chúng ta không thể nào biết trước được liệu một người nào đó sẽ giàu có hơn bao nhiêu, thậm chí cả với một nhóm đối chứng được lựa chọn kỹ lưỡng, vì đây là một vấn đề thực nghiệm.

### **Xem xét lại có thể làm gì với điều tra cơ sở- các ước lượng khác biệt kép**

Nhà phân tích sau đó đặt ra câu hỏi, liệu có một phương pháp khác ngoài lựa chọn ngẫu nhiên mà cũng có tính vững vàng trước các biến không quan sát gây nhiều phiền toái này hay không. Điều này đòi hỏi phải thu thập các dữ liệu gốc về cả những người tham dự và những người không tham dự từ trước khi bắt đầu PROSCOL. Ý tưởng của phương pháp này là cần phải thu thập dữ liệu về các kết quả và những nhân tố quyết định của chúng cả trước và sau khi chương trình được đưa ra, cùng với dữ liệu cho một nhóm đối chứng không tham dự, cũng như cho nhóm tham dự chương trình. Sau đó có thể lấy chênh lệch về mức độ đi học của các thành viên và của nhóm đối chứng sau khi chương trình được triển khai trừ đi chênh lệch này trước khi bắt đầu chương trình. Ước lượng này được gọi là ước lượng “khác biệt kép” hay “khác biệt trong khác biệt.” Nó sẽ giải quyết vấn đề với các biến không thể quan sát, miễn là các biến này không thay đổi theo thời gian.

Có thể giải thích điều này bằng cách thêm chỉ số con vào phương trình trên để đưa thêm vào biến số về tỷ lệ đi học sau khi có chương trình:

$$S_{ia} = a + bP_i + cX_{ia} + \varepsilon_{ia}$$

Trước khi có chương trình, trong điều tra cơ sở, tỷ lệ đi học là

$$S_{ib} = a + cX_{ib} + \varepsilon_{ib}$$

(Tất nhiên  $P = 0$  trước khi chương trình được áp dụng.) Hạng tử sai số bao gồm thêm một tác động cố định theo thời gian, do đó chúng ta có thể viết là:

$$\varepsilon_{it} = \eta_i + \mu_{it} \text{ (với } t = a, b)$$

trong đó  $i$  là tác động không đổi theo thời gian, có tương quan với  $P_i$ , và  $\mu_{it}$  là sai số do đổi mới (innovation error), không có tương quan với  $P_i$  (hay  $X_i$ ).

Ý tưởng cơ bản ở đây là sử dụng dữ liệu gốc để biết về các biến không quan sát này. Cần chú ý, vì điều tra cơ sở cũng được tiến hành cho cùng các hộ gia đình mà chúng ta hiện có dữ liệu, nên hộ gia đình thứ  $i$  trong phương trình  $S_{ia}$  cũng chính là hộ gia đình thứ  $i$  trong phương trình  $S_{ib}$ . Chúng ta có thể tính chênh lệch giữa phương trình "sau" và phương trình "trước":

$$S_{ia} - S_{ib} = bP_i + c(X_{ia} - X_{ib}) + \mu_{ia} - \mu_{ib}.$$

Bây giờ chúng ta có thể hồi quy thay đổi trong tỷ lệ đi học theo sự tham dự chương trình và thay đổi trong  $X$ . Phương pháp OLS sẽ cho chúng ta một ước lượng không chênh của tác động của chương trình. Các biến không quan sát - các biến có tương quan với sự tham dự chương trình - đã được triệt tiêu.

Như vậy, nếu sự tham dự chương trình chỉ dựa trên các biến,

cả quan sát và không quan sát, được biết vào thời điểm điều tra cơ sở, thì có thể sẽ hợp lý khi giả định là các  $\eta$  không thay đổi giữa hai lần điều tra. Điều này sẽ đúng chừng nào các biến không quan sát nhiều phiền phức không thay đổi theo thời gian. Thay đổi trong tỷ lệ đi học theo thời gian của nhóm đối chứng sẽ cho thấy điều gì lẽ ra đã xảy ra với nhóm tham dự nếu như không có chương trình.

Điều này sẽ đòi hỏi phải hiểu kỹ lưỡng về chương trình và khả năng xác định thời điểm của các điều tra đánh giá để có thể phối hợp với chương trình. Nếu không, chắc chắn sẽ có những thay đổi không quan sát được sau khi tiến hành điều tra cơ sở, và do vậy, tác động tới những người tham gia chương trình. Điều này sẽ tạo ra các giá trị thay đổi giữa hai kỳ điều tra.

Phương trình cuối cùng có thể được diễn giải với ý nghĩa đứa trẻ và các đặc điểm của hộ gia đình trong X không liên quan gì với sự thay đổi trong tỷ lệ đi học, nếu như các đặc điểm này không thay đổi theo thời gian. Nhưng sự gia tăng trong tỷ lệ đi học có thể phụ thuộc vào trình độ học vấn của cha mẹ (và không chỉ vào bất kỳ sự thay đổi nào trong trình độ học vấn của họ), và có thể vào việc gia đình sống ở đâu, bởi vì điều này sẽ quyết định sự tiếp cận tới trường học. Có thể cũng có những tình huống trong đó những thay đổi theo thời gian trong chỉ số kết quả bị tác động của các điều kiện bên ngoài. Khi đó, chúng ta sẽ muốn kiểm chứng đổi với các khác biệt trong điều kiện ban đầu. Có thể làm được điều này chỉ đơn giản bằng cách thêm  $X_a$  và  $X_b$  trong phương trình hồi quy một cách riêng biệt, nhờ đó phương trình hồi quy có dạng sau:

$$S_{ia} - S_{ib} = bP_i + c_a X_{ia} + c_b X_{ib} + \mu_{ia} - \mu_{ib}.$$

Ngay cả khi một số (hoặc tất cả) các biến trong X không thay đổi theo thời gian, chúng ta vẫn có thể cho phép X tác động tới các thay đổi theo thời gian trong tỷ lệ đi học.

Phương pháp tương hợp mức độ xu hướng được mô tả ở trên

có thể giúp cho việc đảm bảo nhóm đối chứng tương tự với nhóm tham dự chương trình trước khi tiến hành phương pháp khác biệt kép. Trong một nghiên cứu thú vị về một chương trình việc làm của Mỹ, người ta nhận thấy việc không so sánh được trong một khu vực có sự hỗ trợ chung, là nguyên nhân chính gây ra sai lệch trong ước lượng khác biệt kép khi so sánh với một nhóm kiểm chứng được lựa chọn ngẫu nhiên. Tuy nhiên, ở trong khu vực có sự hỗ trợ chung, thì sự sai lệch với điều kiện của X không thay đổi nhiều lăm theo thời gian. Do vậy, sau khi đã tiến hành phương pháp tương hợp, việc áp dụng phương pháp khác biệt kép là hợp lý (tham khảo Heckman và các tác giả khác (1998)).

Tuy nhiên, trên thực tế, việc tiếp tục sử dụng các hộ gia đình trong các lần điều tra là khó khăn. Có thể không dễ tìm được tất cả các hộ gia đình ban đầu được đưa vào trong điều tra cơ sở. Một số người tham gia điều tra cơ sở không muốn bị phỏng vấn lần nữa, hoặc họ đã di chuyển tới một khu vực không được biết.

Nếu các trường hợp bị loại bỏ khỏi mẫu là hoàn toàn ngẫu nhiên, thì điều tra tiếp theo sẽ vẫn có tính đại diện cho cùng tổng thể như trong điều tra cơ sở. Tuy vậy, nếu có một số xu hướng có tính hệ thống trong việc loại bỏ những cá nhân có những đặc điểm nhất định ra khỏi mẫu, thì sẽ xuất hiện một vấn đề được gọi là “sai lệch do suy giảm” (attrition bias). Chẳng hạn, PROSCOL có thể giúp một số gia đình nghèo di chuyển tới các nhà ở tốt hơn. Và thậm chí, khi sự lựa chọn người tham dự hoàn toàn dựa trên các thông tin sẵn có vào thời điểm điều tra cơ sở (tác động không đổi theo thời gian 0i), thì những thành viên được lựa chọn có thể tự nguyện rút ra khỏi mẫu do các thay đổi xảy ra sau thời điểm này. Sự suy giảm trong nhóm tham dự rõ ràng sẽ làm sai lệch ước lượng khác biệt kép về tác động của chương trình. Hộp 3.3 nêu các bước để có được một ước lượng kép.

### Hộp 3.3 Thực hiện phương pháp khác biệt kép

Phương pháp khác biệt kép bao gồm việc so sánh một nhóm tham dự với một nhóm đối chứng (có thể được xác định một cách tốt nhất bằng phương pháp tương hợp như trong hộp 3.2) cả trước và sau khi có chương trình. Các bước chính trong phương pháp này như sau:

**Bước 1:** Bạn cần có điều tra gốc trước khi tiến hành chương trình, và điều tra này phải bao gồm cả những người không tham dự và người tham dự. Nếu bạn không biết những người nào sẽ tham dự, bạn phải dự đoán trên cơ sở thông tin có được. Hãy nói chuyện với những người phụ trách chương trình.

**Bước 2:** Bạn cần một hay hai kỳ điều tra tiếp theo sau khi chương trình đã được tiến hành. Các kỳ điều tra này phải có thể so sánh được với điều tra gốc (về bảng câu hỏi, phỏng vấn, v.v.). Tốt nhất là các lần điều tra tiếp theo cần sử dụng cùng những quan sát được chọn mẫu giống như điều tra gốc. Nếu điều này là không thể, thì các quan sát này cần ở trong cùng khu vực địa lý hay cùng lớp căn cứ theo một số biến khác.

**Bước 3:** Tính toán chênh lệch trung bình giữa các giá trị sau và trước của chỉ số kết quả đối với nhóm tham dự và nhóm đối chứng.

**Bước 4:** Tính toán chênh lệch giữa hai mức chênh lệch trung bình. Đây chính là ước lượng của bạn về tác động của chương trình.

Đây là dạng đơn giản nhất trong phương pháp khác biệt kép. Bạn cũng có thể muốn kiểm chứng các khác biệt trong điều kiện ban đầu nội sinh hay những thay đổi trong các biến ngoại sinh, có thể cho phép có các tác động tương tác với chương trình (nhờ thế mà lợi ích từ chương trình là hàm số của các biến quan sát). Một mô hình hồi quy phù hợp có thể cho phép có những dạng như vậy của phương pháp này.

## Sử dụng các biến số công cụ

Khi không có điều tra cơ sở về cùng những hộ gia đình để thực hiện phương pháp khác biệt kép, thì vị giáo sư khuyến nghị sử dụng một phương pháp khác để có một ước lượng vững vàng trước các biến không quan sát gây phiền phức- đó là phương pháp “biến số công cụ”.

Biến số công cụ là giải pháp cổ điển đối với vấn đề biến số nội sinh. Biến số công cụ là một biến số quan sát được sử dụng để tạo ra biến số ngoại sinh trong sự tham dự chương trình. Nói cách khác, nó tương quan với  $P$  nhưng không có sẵn trong phương trình hồi quy về tỷ lệ đi học và không tương quan với hạng tử sai số  $\epsilon$  trong phương trình đi học. Vì vậy, chúng ta phải có ít nhất một biến số trong  $Z$  mà không ở trong  $X$ , và không tương quan với  $\epsilon$ . Sau đó các biến số công cụ sẽ ước lượng tác động của chương trình bằng cách thay  $P$  bằng giá trị dự đoán của nó với điều kiện của  $Z$ . Do giá trị dự đoán này hoàn toàn phụ thuộc vào  $Z$  (là ngoại sinh) và  $Z$  không tương quan với  $\epsilon$ , nên bây giờ việc sử dụng OLS cho phương trình hồi quy mới hợp lý.

Do giá trị dự đoán chỉ phụ thuộc vào biến số ngoại sinh do biến số công cụ và các biến số ngoại sinh khác, nên các biến số không quan sát sẽ không còn là vấn đề, bởi vì chúng sẽ không có tương quan với hạng tử sai số trong phương trình hồi quy. Điều này cũng cho thấy, có một biện pháp khác hiệu quả hơn để giải quyết vấn đề này. Chúng ta nhớ rằng, nguyên nhân sai lệch trong ước tính trước đây về tác động của chương trình là sự tương quan giữa hạng tử sai số trong phương trình đi học và trong phương trình tham dự. Điều này tạo ra mối tương quan giữa mức độ tham dự và hạng tử sai số trong phương trình đi học. Do đó, một cách tự nhiên để tránh vấn đề này một khi có biến số công cụ, là bổ sung các phần dư từ phương trình tham dự giai đoạn một vào phương trình đi học nhưng vẫn duy trì sự tham dự thực tế trong phương trình hồi quy đi học. Tuy

nhiên, vì bây giờ chúng ta đã bổ sung vào phương trình hồi quy đi học giá trị ước lượng của hạng tử sai số từ phương trình tham dự, nên có thể xem sự tham dự như một biến số ngoại sinh và chạy hồi quy OLS. Cách thức này sẽ có kết quả, nếu có một công cụ hợp lý. Nếu không, phương trình hồi quy sẽ không ước lượng được bởi vì hoàn toàn có thể dự đoán được phần dư của sự tham dự từ sự tham dự thực tế và  $X$ , trong một mô hình tuyến tính.

Một biến số công cụ cũng có thể có ích nếu như xảy ra sai số đo lường đáng kể trong dữ liệu về sự tham dự chương trình - một nguyên nhân gây ra sai lệch khác. Sai số đo lường có nghĩa là có khả năng sự tham dự chương trình thay đổi nhiều hơn trên thực tế. Sự ước lượng quá cao đối với phương sai của  $P$  tất nhiên sẽ dẫn đến ước lượng thấp quá hệ số  $b$  của phương trình. Sai lệch này được gọi là sai lệch suy giảm (attenuation bias), do nó làm suy giảm hệ số hồi quy ước lượng được.

Mặc dù một biến số công cụ có thể là rất có ích, nhưng cần phải thận trọng trong việc sử dụng nó trên thực tế. Khi sự tham dự thực tế được thay thế bằng giá trị dự đoán và khi chạy OLS, thì các kết quả độ lệch chuẩn sẽ không chính xác, vì máy tính sẽ không biết được phải sử dụng các tham số ước lượng trước đây để có được các giá trị dự đoán. Cần có biện pháp sửa chữa đối với các độ lệch chuẩn được ước lượng theo OLS, mặc dù có các phần mềm thống kê cho phép chúng ta làm được điều này một cách dễ dàng, ít nhất là với các mô hình tuyến tính.

Tuy vậy, nếu như có một biến số phụ thuộc mà chỉ có thể nhận hai giá trị, chẳng hạn đi học hay không đi học, thì chúng ta cần sử dụng một mô hình phản ứng nhị phân phi tuyến tính, chẳng hạn mô hình logit hay probit. Nguyên tắc kiểm định tính ngoại sinh của sự tham dự chương trình cũng tương tự trong trường hợp này. Có một bài viết của Rivers và Vuong (1988) thảo luận về vấn đề đối với các mô hình này; Blundell và Smith (1993) cung cấp một tổng quan thiết thực về các mô hình phi tuyến tính khác nhau, trong đó có một biến số nội sinh.

## Kiểm định các phương pháp

Khi nhà phân tích bắt đầu nghĩ đến việc xác định một biến số công cụ, sẽ nhận thấy đây không là phải một tiến trình đơn giản. Các biến số trong X có thể nhận bất kỳ một khả năng nào. Vấn đề là, tìm ra được một "sự hạn chế có tính loại trừ" hợp lệ để có thể đưa một số biến số vào trong phương trình tham dự, nhưng không có mặt trong phương trình đi học.

Nhà phân tích quyết định sử dụng phương pháp tương hợp mức độ xu hướng. Mô hình tham dự có dạng logit tương đối hợp lý, và cho thấy PROSCOL được định hướng tốt. Hầu như tất cả các biến số mà chúng ta dự kiến sẽ đi kèm với nghèo đói đều có hệ số dương và có ý nghĩa. Nhà phân tích sau đó sử dụng phương pháp tương hợp mức độ xu hướng. Khi so sánh các tỷ lệ nhập học trung bình, các kết quả cho thấy, trẻ em từ nhóm đối chứng tương hợp có tỷ lệ nhập học bằng 60 phần trăm so với con số 80 phần trăm của các gia đình tham dự PROSCOL.

Để giải thích cho vấn đề thu nhập dự kiến, nhà phân tích căn cứ vào một điều tra hiện tại về lao động trẻ em, trong đó có các câu hỏi về thu nhập. (Ở nước đang phát triển này, có luật chính thức cấm trẻ em làm việc khi chưa đến 16 tuổi, nhưng chính phủ gặp nhiều khó khăn trong khi thực hiện luật này; tuy vậy, tiền lương của trẻ em vẫn là một vấn đề nhạy cảm.) Từ điều tra này, có thể xác định được thu nhập mà một đứa trẻ lẻ ra sẽ có nếu nó không đi học.

Sau đó chúng ta có thể trừ khoản thu nhập bỏ lõi ra khỏi khoản thanh toán tiền mặt của PROSCOL cho những người tham dự, và nhờ đó, tính được giá trị chuyển nhượng thu nhập ròng. Trừ giá trị chuyển nhượng ròng này khỏi tổng thu nhập, chúng ta có thể xác định được những người tham dự PROSCOL xuất phát từ đâu trong hàm phân phối thu nhập trước khi tham dự chương trình. Những người này không hoàn toàn nghèo như đã tưởng tượng ban đầu (không tính đến thu nhập bỏ lõi), nhưng họ vẫn là người nghèo; chẳng hạn, hai phần ba trong số

họ nằm dưới ngưỡng nghèo chính thức của đất nước.

Sau khi đã tính được lợi ích thu nhập ròng đối với tất cả những người tham dự, bây giờ có thể tính được tỷ lệ nghèo đói khi có và không có PROSCOL. Tỷ lệ nghèo đói sau khi can thiệp (có chương trình) chỉ đơn giản là tỷ lệ dân cư sống trong các hộ gia đình có thu nhập trên đầu người thấp hơn ngưỡng nghèo, trong đó “thu nhập” là thu nhập quan sát được, “kể cả các khoản chuyển nhượng gộp nhận được từ PROSCOL”. Có thể tính một cách trực tiếp tỷ lệ này từ điều tra hộ gia đình. Bằng cách lấy toàn bộ thu nhập quan sát được, trừ đi lợi ích thu nhập ròng (thanh toán tiền mặt từ PROSCOL trừ đi thu nhập bỏ lỡ từ công việc của trẻ em) do PROSCOL mang lại, các kết quả cho thấy một sự phân phối mới của các thu nhập trước chương trình. Tỷ lệ nghèo đói không có chương trình khi đó là tỷ lệ những người sống trong các hộ gia đình nghèo, dựa trên sự phân phối mới này. Nhà phân tích nhận thấy, tỷ lệ nghèo đói quan sát được ở Tây Bắc bằng 32 phần trăm so với con số 36 phần trăm, nếu như không có PROSCOL. Chương trình này đã giúp cho 4 phần trăm dân cư thoát khỏi cảnh nghèo ở thời điểm hiện tại. Lợi ích từ di học có nghĩa là, sẽ có cả những lợi ích bằng tiền và không bằng tiền cho người nghèo trong tương lai. Trong quá trình đo lường nghèo đói, nhà phân tích nhớ lại rằng tỷ lệ người nghèo sống dưới ngưỡng nghèo chỉ là một thước đo cơ bản, bởi vì nó không cho chúng ta biết gì về những thay đổi ở dưới ngưỡng này (xem Hộp 3.4), do đó ta bèn tính cả chỉ số khoảng nghèo (poverty gap index) và chỉ số khoảng nghèo bình phương, và các kết quả cho thấy, các chỉ số này cũng giảm nhờ PROSCOL.

Trong tính toán này, nhà phân tích cũng thừa nhận có một số sự thiếu chắc chắn trong ngưỡng nghèo của quốc gia. Để kiểm định các kết quả này, nhà phân tích lặp lại sự tính toán trên nhiều ngưỡng nghèo, và nhận thấy, tại một ngưỡng nghèo trong đó 50 phần trăm dân số là người nghèo căn cứ theo thu nhập quan sát được, sau khi thực hiện chương trình, tỷ lệ này sẽ bằng 52 phần trăm nếu như không có PROSCOL. Ở một

### Hộp 3.4 Các thước đo nghèo đói

Thước đo nghèo đói đơn giản và phổ biến nhất là chỉ số đếm đầu (headcount index). Trong trường hợp này, đây là tỷ lệ dân số sống tại các hộ gia đình có thu nhập trên đầu người thấp hơn ngưỡng nghèo. (Ở các nước khác, chỉ số này là một thước đo dựa trên cơ sở tiêu dùng, thước đo trên cơ sở tiêu dùng có một số ưu điểm; tham khảo thảo luận và tài liệu tham khảo trong Ravallion 1994.)

Chỉ số đếm đầu không cho chúng ta biết điều gì về phân phối thu nhập ở dưới ngưỡng nghèo: một người nghèo có thể nghèo hơn nhưng chỉ số đếm đầu sẽ không thay đổi; nó cũng không phản ánh các lợi ích mà người nghèo nhận được, trừ khi họ vượt qua được ngưỡng nghèo.

Một thước đo khác được sử dụng rộng rãi thay cho chỉ số đếm đầu là chỉ số khoảng nghèo (PG). Khoảng nghèo đói cho mỗi hộ gia đình là sự chênh lệch giữa ngưỡng nghèo và thu nhập của hộ gia đình; đối với những người ở trên ngưỡng nghèo, khoảng nghèo này bằng không. Khi khoảng nghèo được so với ngưỡng nghèo và chúng ta tính toán giá trị trung bình trên tất cả các hộ gia đình (cho dù nghèo hay không nghèo), thì chúng ta sẽ có được chỉ số khoảng nghèo.

Chỉ số khoảng nghèo sẽ cho chúng ta biết mức độ tác động của một chương trình tới độ sâu của nghèo đói, nhưng nó sẽ không phản ánh bất kỳ thay đổi nào mà chương trình này gây ra trong phân phối giữa những người nghèo. Chẳng hạn, nếu chương trình gồm có một lợi ích nhỏ cho một người nghèo, có thu nhập cao hơn thu nhập trung bình của những người nghèo, trong khi lại làm cho một người nào đó ở mức thu nhập thấp hơn thu nhập trung bình bị thiệt hại thì PG sẽ không thay đổi.

Có nhiều thước đo “nhạy cảm với thu nhập” sẽ phản

### Hộp 3.4 (Tiếp)

ánh những thay đổi này trong phân phối thu nhập giữa những người nghèo. Một thước đo như vậy là “khoảng nghèo bình phương” (Foster, Greer, và Thorbecke 1984). Thước đo này được tính toán tương tự như PG trừ việc các khoảng nghèo riêng lẻ tính theo tỷ lệ với ngưỡng nghèo được bình phương trước khi tính giá trị trung bình (đối với cả người nghèo và không nghèo). Một ví dụ khác về thước đo nghèo đói nhạy cảm với phân phối là chỉ số Watt. Chỉ số này bằng trung bình logarit của tỷ lệ ngưỡng nghèo trên thu nhập, trong đó tỷ lệ này được đặt bằng một đối với những người không nghèo. Atkinson (1987) mô tả các ví dụ khác trong bài viết của mình.

ngưỡng nghèo khác, có 15 phần trăm số dân nằm dưới ngưỡng nghèo khi có chương trình và 19 phần trăm nếu như không có chương trình. Bằng cách lặp lại các tính toán này trên những mức thu nhập khác nhau, toàn bộ “các đường cong diện nghèo đói” được xác định khi có và khi không có chương trình. Hình này cũng được gọi là “hàm phân phối tích luỹ” (xem Hộp 3.5).

### Hộp 3.5 So sánh nghèo đói khi có và không có chương trình

Sử dụng các phương pháp được mô tả trong nội dung chính và trong các hộp ở trên, chúng ta có được một ước lượng về lợi ích đối với mỗi hộ gia đình. Trong các đánh giá đơn giản nhất, lợi ích này chỉ là một con số. Nhưng có lẽ để cho lợi ích này thay đổi theo những đặc điểm của hộ gia đình sẽ tốt hơn. Khi đó chúng ta có thể tóm tắt thông tin này dưới

(Xem tiếp trang sau)

### Hộp 3.5 (Tiếp)

hình thức đường cong diện nghèo đói (PIC), khi có và không có chương trình.

**Bước 1:** Cần phải có từ trước dữ liệu về thu nhập sau khi có chương trình (hay chỉ số lợi ích khác) của mỗi hộ gia đình trong toàn bộ mẫu (so sánh cả những người tham dự và những người không tham dự). Chúng ta cũng biết số người ở trong mỗi gia đình. Và, tất nhiên, chúng ta biết tổng số người trong mẫu ( $N$ ; hoặc đây có thể là quy mô tổng thể ước tính, nếu như các tỷ lệ mẫu đảo ngược đã được sử dụng để “sử dụng hết” mỗi quan sát mẫu).

**Bước 2:** Chúng ta có thể vẽ dưới hình thức một đường cong diện nghèo đói trên cơ sở thông tin này. Đường cong này cho chúng ta (tính theo trực tung) tỷ lệ phần trăm dân cư sống trong các hộ gia đình có thu nhập thấp hơn hoặc bằng giá trị ở trên trực hoành. Để vẽ đồ thị này, chúng ta có thể bắt đầu bằng hộ gia đình nghèo nhất, đánh dấu thu nhập của hộ này trên trực hoành, và sau đó dếm trên trực tung giá trị bằng số người sống trong hộ gia đình này, nhân với 100, và chia cho  $N$ . Điểm tiếp theo là tỷ lệ những người sống trong hai hộ gia đình nghèo nhất, và cứ thế tiếp tục. Kết quả sẽ cho là một PIC sau khi có chương trình.

**Bước 3:** Bây giờ chúng ta tính toán sự phân phối thu nhập trước khi có chương trình. Để làm được điều này, bạn trừ đi lợi ích ước tính của mỗi hộ gia đình khỏi thu nhập sau chương trình của hộ này. Bạn sẽ có một danh sách các thu nhập sau chương trình tương ứng với mỗi hộ gia đình được chọn mẫu. Sau đó lặp lại bước 2, bạn sẽ có được PIC trước chương trình.

Nếu chúng ta nghĩ đến mức thu nhập nhất định ở trực tung như là ngưỡng nghèo, thì sự chênh lệch giữa hai PIC ở điểm này cho chúng ta biết tác động của chỉ số đếm đầu tương ứng với ngưỡng nghèo này (hộp 3.4). Mặt khác,

### Hộp 3.5 (Tiếp)

nhìn vào trục hoành, chúng ta sẽ biết lợi ích thu nhập cho nhóm phân trăm này. Nếu không có mức lợi ích nào nhận giá trị âm, thì đường PIC sau chương trình phải nằm ở dưới đường trước chương trình. Mức nghèo đói sẽ giảm xuống cho dù có sử dụng ngưỡng nghèo nào. Trên thực tế, điều này cũng đúng với rất nhiều thước đo nghèo đói khác; xem Atkinson (1987). Nếu một số lợi ích có dấu âm thì các PIC sẽ cắt nhau. Sự so sánh nghèo đói khi đó sẽ không rõ ràng; câu trả lời sẽ phụ thuộc vào ngưỡng nghèo và các thước đo nghèo đói nào được sử dụng. Chẳng hạn, bạn có thể tin rằng ngưỡng nghèo không vượt quá một số giá trị tối đa, và nếu các đường này cắt nhau ở trên giá trị này, thì việc so sánh nghèo đói trở nên khó khăn. Nếu điểm cắt nhau (và có thể có nhiều điểm như vậy) nằm dưới ngưỡng nghèo cao nhất có thể chấp nhận, thì chỉ có thể có một sự so sánh nghèo đói đáng tin cậy với một số các thước đo nghèo đói nhất định. Để kiểm tra mức độ hạn chế của những thước đo này, cần tính toán được đường cong độ sâu nghèo đói (PDC). Có thể làm được điều này chỉ đơn giản bằng cách cộng tích luỹ tới mỗi điểm ở trên PIC. (Điểm thứ hai trên PDC là điểm thứ nhất trên PIC cộng với điểm thứ hai, v.v.).

Nếu các PDC không cắt nhau, thì tác động của chương trình tới nghèo đói là không rõ ràng chừng nào chúng ta còn hạn chế sự quan tâm của mình dành cho chỉ số khoảng nghèo hay bất kỳ các thước đo nghèo đói nhạy cảm với phân phối nào như mô tả trong hộp 3.4. Nếu các PDC cắt nhau thì bạn có thể tính toán “các đường cong mức độ nghiêm trọng của nghèo khổ” khi có và không có chương trình bằng cách cộng tích luỹ dưới các PDC. Nếu các đường này không cắt nhau ở trên những ngưỡng nghèo có thể chấp nhận được, thì tác động tới bất kỳ một thước đo nghèo đói nào nhạy cảm với thu nhập trong hộp 3.4 cũng là không rõ ràng.

### Kết hợp đầu vào tại thực địa

Trong việc thực hiện bất kỳ một chương trình nào, sự sáng suốt của người hưởng lợi và người phụ trách chương trình có thể không được phản ánh trong dữ liệu của chương trình. Ví dụ, trong trường hợp này, nhận thức của những người làm việc tại thực địa là hầu hết những gia đình tham gia PROSCOL đều nghèo, và chương trình này thực sự đã giúp đỡ các gia đình đó. Khi thảo luận điều này với một nhà xã hội học đã làm việc với chương trình, nhà phân tích biết được một số sự thiếu chắc chắn trên thực tế về số thu nhập dự kiến và vấn đề việc làm. Nhà xã hội học cho rằng người ta quan sát thấy nhiều trẻ em từ các gia đình nghèo vừa làm việc lại vừa đi học, và một số trẻ em ít tuổi hơn không đi học nhưng cũng không làm việc. Nhà phân tích nhận thấy điều này đòi hỏi phải kiểm tra, liệu có sự khác biệt trong số lượng lao động trẻ em của các trẻ em trong chương trình PROSCOL so với trong nhóm đối chứng tương hợp hay không. Tuy vậy, dữ liệu này không có sẵn trong điều tra hộ gia đình, mặc dù có thể trình bày các kết quả khi tính đến và không tính đến việc trừ thu nhập bỏ lỡ.

Nhà xã hội học cũng nhận thấy, với một gia đình nghèo tham gia PROSCOL, việc gia đình này sống trong khu vực nhà trường (SBA) nào rất quan trọng. Tất cả các SBA đều nhận được phân bổ của chương trình PROSCOL từ trung tâm, ngay cả với các SBA có rất ít các gia đình nghèo. Nếu một gia đình nghèo nhưng lại sống trong một SBA giàu có, họ có nhiều khả năng nhận được sự giúp đỡ từ PROSCOL hơn là trường hợp họ sống trong một SBA nghèo. Điều thực sự quan trọng là khi đó, mức độ nghèo đói tương đối- tương đối so với những người khác trong khu vực mà gia đình đó sống- có vai trò quan trọng hơn nhiều mức sống tuyệt đối.

Sự phân bổ này sẽ tác động tới việc tham dự PROSCOL, nhưng chúng ta không kỳ vọng nó tác động tới việc đi học, việc này phụ thuộc nhiều hơn vào mức sống tuyệt đối, tình hình gia đình và các đặc điểm của nhà trường. Do đó, sự phân bổ ngân

sách của PROSCOL cho các SBA có thể được sử dụng như các biến số công cụ nhằm triệt tiêu sai lệch trong ước lượng về tác động của chương trình.

Chúng ta có thông tin về SBA của mỗi gia đình trong điều tra hộ gia đình, các quy tắc mà chương trình sử dụng để phân bổ số tiền trong chương trình cho các SBA và lượng tiền được phân bổ cho mỗi SBA. Sự phân bổ này dựa trên số lượng trẻ em ở độ tuổi đi học, với một "nhân tố điều chỉnh" theo mức độ nghèo đói của SBA. Tuy vậy, quy tắc này đôi khi cũng không rõ ràng.

Nhà phân tích cố gắng cân nhắc đến các điểm này, và sau đó chạy lại phương trình hồi quy đi học, nhưng thay thế sự tham dự PROSCOL thực tế bằng giá trị dự đoán (mức độ xu hướng) từ phương trình hồi quy tham dự- phương trình này bây giờ bao gồm cả sự phân bổ ngân sách cho SBA. Công việc này giúp cho việc bổ sung các đặc điểm của trường học vào phương trình hồi quy đi học. Mặc dù các đặc điểm của nhà trường không có tác động một cách chính thức tới cách thức phân bổ các nguồn lực của PROSCOL, nhưng bất kỳ những đặc điểm nào của nhà trường bị bỏ qua đều sẽ tác động đồng thời tới sự phân bổ PROSCOL theo SBA và kết quả học tập cá nhân, đều sẽ gây ra sự sai lệch trong các ước lượng biến số công cụ của nhà phân tích. Mặc dù luôn có khả năng xảy ra sai lệch, nhưng với nhiều biến đổi chứng theo khu vực địa lý hơn, thì phương pháp này ít nhất cũng cần tạo ra được một kết quả so sánh đáng tin cậy so với ước lượng tương hợp.

Từ các kết quả này, chúng ta thấy sự phân bổ ngân sách cho SBA thực sự có hệ số dương, có ý nghĩa trong phương trình hồi quy logit về sự tham dự PROSCOL. Bây giờ sự tham dự PROSCOL (dự đoán) có ý nghĩa trong phương trình hồi quy đi học, trong đó tất cả các biến số tương tự từ hồi quy logit đều được đưa vào trừ việc phân bổ ngân sách cho SBA. Hệ số này phản ánh tỷ lệ đi học cao hơn 15 điểm phần trăm với những người tham dự PROSCOL. Nhà phân tích cũng tiến hành chạy một phương trình hồi quy số năm đi học cho các trẻ em trai và

gái một cách riêng biệt. Với cả trẻ em trai và em gái ở độ tuổi 18, các kết quả cho thấy, những người này sẽ rời khỏi nhà trường sớm hơn hai năm nếu không có PROSCOL. Tuy vậy, phương trình hồi quy này vẫn có những vấn đề- chẳng hạn liệu nó có sử dụng các độ lệch chuẩn đúng đắn không, và liệu có nên sử dụng mô hình tuyến tính hay không.

### Lập kế hoạch cho công việc tương lai

Cuối cùng, nhà phân tích đã sẵn sàng để báo cáo các kết quả đánh giá. Các kết quả này cho thấy, chương trình PROSCOL đã có kết quả tốt, và do vậy, các nhà hoạch định chính sách tỏ ra quan tâm tới việc mở rộng chương trình. Căn cứ vào tiến trình thực hiện đánh giá, nhà phân tích có một số quan sát quan trọng:

- Đánh giá tác động có thể phức tạp hơn nhiều so với dự đoán ban đầu
- Có thể có nhiều ước lượng đến mức đáng lo ngại, tùy thuộc vào các phương pháp cụ thể được sử dụng
- Cần sử dụng các phương pháp thay thế trong những tình huống thường xuyên khi dữ liệu không hoàn toàn tốt, mặc dù mỗi phương pháp đều có các nhược điểm; và
- Chúng ta cần phải chiết trung đối với các dữ liệu.

Ngoài các bài học này, nhà phân tích còn có một số khuyến nghị quan trọng đối với việc đánh giá tương lai về PROSCOL. Thứ nhất, cần phải loại bỏ một cách ngẫu nhiên một số gia đình trong PROSCOL ở phần còn lại của đất nước, và sau đó điều tra tiếp theo với cả những người tham dự thực tế và những người bị loại bỏ ngẫu nhiên không được tham dự. Khi đó chúng ta sẽ có một ước tính lợi ích chính xác hơn. Tuy nhiên, việc loại bỏ một số đối tượng có thể nhạy cảm về mặt chính trị. Mặc dù vậy, nếu chương trình không có đủ nguồn lực để thực hiện ngay lập tức trong cả nước, và nếu chương trình phải lựa chọn về những người nào sẽ được tham dự lần

đầu, thì cần phải lựa chọn một cách ngẫu nhiên trong số những người đủ tư cách tham dự. Mặt khác, cũng có thể lựa chọn các trường học hay khu vực nhà trường một cách ngẫu nhiên. Điều này sẽ đảm bảo cho việc lựa chọn nhà trường hay khu vực nhà trường trở thành một biến số công cụ tốt đối với sự tham dự chương trình cá nhân.

Thứ hai, nếu như điều này là không thể, thì cần phải tiến hành điều tra cơ sở ở các khu vực có độ tập trung những người tham dự PROSCOL cao trước khi bắt đầu chương trình ở miền Nam. Có thể làm được điều này đồng thời với giai đoạn tiếp theo trong lần điều tra toàn quốc từng được sử dụng để đánh giá chương trình PROSCOL. Cũng nên bổ sung một số câu hỏi vào điều tra, chẳng hạn như câu hỏi trẻ em có làm một công việc được trả công nào không.

Và thứ ba, việc kết hợp các công việc định tính sẽ rất có ích, góp phần hình thành các giả thuyết cần được kiểm định và giúp cho việc đánh giá tính hợp lý của các giả định quan trọng trong phân tích định tính.

---

---

## Chương 4

# Kinh nghiệm rút ra từ các đánh giá tác động “thành công trên thực tế”\*

Các chương trước đã trình bày các phương pháp, vấn đề, và thách thức chủ yếu có thể nảy sinh trong đánh giá tác động của dự án. Trong việc xem xét lại các nghiên cứu tình huống nêu ra trong bảng 4.1, cho thấy có nhiều ví dụ minh họa về các phương pháp thú vị trong thiết kế, sử dụng dữ liệu, lựa chọn và áp dụng các phương pháp phân tích, và xây dựng năng lực ở trong nước. Các ví dụ này, cũng như cuộc thảo luận về chi phí của đánh giá và các vấn đề kinh tế chính trị có thể phát sinh khi thực thi, sẽ được nêu bật trong phần dưới đây.

Mười lăm nghiên cứu tình huống nêu trong phần đánh giá được lựa chọn từ những đánh giá do Ngân hàng Thế giới, các cơ quan tài trợ khác, các viện nghiên cứu, và các hãng tư vấn tư nhân. Các nghiên cứu tình huống này được lựa chọn như là một mẫu bao gồm các tình huống “thành công trên thực tế” nhờ tính vững vàng về phương pháp luận, với sự cố gắng phản ánh những ví dụ đa dạng tại các lĩnh vực và khu vực khác nhau. Mỗi đánh giá tác động đều có ưu điểm và nhược điểm riêng, tuy nhiên các bài học rút ra từ những kinh nghiệm này sẽ giúp ích cho các nhà quản lý dự án hay nhà phân tích chính sách có ý định thiết kế và thực thi công việc trong tương lai.

---

\* Chương này dựa trên các nghiên cứu tình huống tốt nhất trên thực tế ở phụ lục 1 và các phần tổng quan được Gillette Hall và Julia Lane, và Subbarao và các tác giả khác (1999) chuẩn bị.

## Lập kế hoạch thiết kế đánh giá sớm và kỹ lưỡng

Sự chuẩn bị đầy đủ trong các giai đoạn bắt đầu định hình dự án sẽ đảm bảo cho việc thu thập được các thông tin cần thiết và có thể sử dụng các kết quả cho các điều chỉnh về thành phần của dự án khi dự án đang được thực hiện. Với việc lập kế hoạch sớm và kỹ lưỡng, có thể kết hợp được tất cả các thành phần để góp phần làm tăng độ chắc chắn trong đánh giá tác động, chẳng hạn tiến hành một cuộc điều tra cơ sở với nhóm đối chứng (control group) được lựa chọn ngẫu nhiên, và dữ liệu định tính về các quá trình tác động tới tác động của dự án.

**Dự án dinh dưỡng và phát triển trẻ em ở Uganda.** Đánh giá này, tuy vẫn chưa được thực thi, nhưng đã cho chúng ta một ví dụ xuất sắc về việc lập kế hoạch sớm và kỹ lưỡng (xem Ngân hàng Thế giới 1998a; Garcia, Alderman và Rudqvist 1999). Dự án tập trung vào việc nâng cao năng lực của các bậc cha mẹ và cộng đồng trong việc chăm sóc trẻ em bằng cách cung cấp cho họ những kiến thức chăm sóc trẻ em tốt hơn và tăng cường cơ hội để nâng cao thu nhập. Dự án này dựa trên cộng đồng và do một mạng lưới các tổ chức phi chính phủ (NGO) thực hiện. Thành phần đánh giá, được kết hợp trong chu kỳ dự án từ ngày một, là rất hoàn hảo nếu căn cứ theo thiết kế đánh giá. Thứ nhất, đánh giá này tạo ra các dữ liệu điều tra cơ sở và điều tra tiếp theo, cùng với một nhóm đối chứng được lựa chọn ngẫu nhiên, nhờ đó tác động của chương trình tới những người thụ hưởng được đánh giá một cách rõ ràng. Thứ hai, nó tăng cường thành phần định lượng bằng một quá trình giám sát và đánh giá (M&E) có sự tham dự rộng rãi (thành phần định tính).

Về mặt định lượng, dự án được thiết kế nhằm cho phép một thiết kế nghiên cứu thử nghiệm trong đó các xã được phân công ngẫu nhiên vào trong các nhóm tham dự và nhóm đối chứng. Sau đó, người ta sử dụng các thẻ sức khoẻ để ghi chép dữ liệu về cân nặng của trẻ em tại các xã thuộc hai nhóm này. Hơn nữa, điều tra hộ gia đình cơ sở được tiến hành trước khi cung cấp

dịch vụ cho cộng đồng, kỳ điều tra tiếp theo được tiến hành hai năm sau đó, với chính những hộ gia đình này. Việc đánh giá lại các dữ liệu này một cách nhanh chóng, dự kiến sẽ cung cấp thông tin cho các quyết định nhằm mở rộng một số thành phần của chương trình trong đợt đánh giá giữa kỳ của dự án. Sự phân tích sâu sắc hơn về dữ liệu vào cuối kỳ của dự án sẽ hướng dẫn cho thiết kế giai đoạn thứ hai của dự án.

**Dự án tín dụng cùng với giáo dục ở Gana.** Đánh giá dự án này rất phức tạp, bao gồm nhiều bước trung gian. Dự án kết hợp các thành phần của một chương trình cho vay theo nhóm cùng với giáo dục, trên cơ sở sức khoẻ, dinh dưỡng, thời điểm sinh và khoảng cách sinh, cùng với kỹ năng kinh doanh nhỏ. Đánh giá chung tập trung vào tình trạng dinh dưỡng của trẻ em, năng lực kinh tế của phụ nữ trong việc đầu tư vào lương thực và y tế, kiến thức của phụ nữ và việc áp dụng nuôi con bằng sữa mẹ và cai sữa. Dự án bắt đầu bằng một khung khái niệm rất rõ ràng, sẽ được minh họa dưới đây. Lược đồ này xác định rõ các đầu vào, lợi ích trung gian, và các lợi ích lâu dài, nhờ đó giúp cho việc xây dựng và diễn giải một số mô hình. Bằng cách lập kế hoạch đánh giá kỹ lưỡng và sử dụng một lược đồ ở giai đoạn ban đầu, có thể làm rõ nhiều điểm trong thiết kế tương đối phức tạp này (xem phụ lục 1.6).

### Các phương pháp đánh giá khi không có dữ liệu gốc

Trên thực tế, nhiều đánh giá không có đầy đủ dữ liệu. Các đánh giá thường được bổ sung vào sau khi có thể tiến hành điều tra cơ sở hay khi không có nhóm đối chứng. Một số ví dụ về các đánh giá này là đánh giá các dự án Đổi thực phẩm lấy giáo dục của Băngladét, PROBECAT của Mêhicô, Các chương trình lao động tích cực của Séc, và TRABAJAR của Áchentina. Do không có dữ liệu gốc, nên cần phải tiến hành so sánh kết quả bằng các phương pháp tương hợp (matching methods) như đã thảo luận trong các chương trước. Tuy nhiên, công việc này

**Bảng 4.1 Tóm tắt các đánh giá tác động "thành công trên thực tế"**

Đề án	Tên nước	Đóng cơ sở dữ liệu	Đơn vị phân tích	Phương pháp kinh tế luồng					
				Lựa chọn	Số sản xuất	Tiếng phản hồi	Biến	Định h	Ưu điểm
Giáo dục				ngẫu nhiên	khác	công cụ	tinh		
Radio ở Nicaragua	Nicaragua	Điều tra cơ sở và sau khi có chương trình	Học sinh và các lớp học	Kết quả điểm kiểm tra	Có	Không	Có	Không	Không
Cải cách tư duy trường học	Nicaragua	Điều tra nhóm mẫu và đánh giá định hình	Học sinh, cha mẹ, giáo viên, hiệu trưởng	Kết quả điểm kiểm tra, mức độ ra quyết định của địa phương	Có	Không	Có	Không	Không
Sách giáo khoa	Kenya	Điều tra cơ sở và sau khi có chương trình	Học sinh, lớp học, giáo viên	Kết quả điểm kiểm tra	Có	Không	Có	Có	Không
Tình trạng bỏ học	Philippin	Điều tra cơ sở và sau khi có chương trình	Học sinh, lớp học, giáo viên	Kết quả điểm kiểm tra và tình trạng bỏ học	Có	Không	Có	Có	Không
<b>Chương trình lao động TRABAJAR Argentina</b>				Người lao động, hộ gia đình	Thu nhập, định hướng đối tượng, chi phí	Không	Có	Không	Có

(Xem tiếp trang sau)

Bảng 4.1 (Tiếp)

Điểm	Tên nước	Đang cơ sở để liệu	Đơn vị phân tích	Phương pháp kinh tế lượng					
				Lựa chọn ngẫu nhiên	Do luồng kết quả	Tuồng hợp	Số sẵn phản thần	Khác biệt	Định công cụ
PROBECAT	Méxicô	Điều tra hỏi tướng và điều tra lực lượng lao động	Người lao động	Thu nhập và việc làm	Không	Có	Không	Không	Không
Chương trình	Công hòa lao động Séc	Điều tra thuế tin hồi tưởng tích cực	Người lao động	Thu nhập và việc làm	Không	Có	Không	Không	Kỹ thuật tương hợp
Tài chính	Bangladesh	Điều tra sau khi có chương trình và số sánh hành chính	Hộ gia đình và giáo dục	Tiêu dùng và giáo dục	Có	Có	Không	Có	Không
Tin dụng cho giáo dục	Gana	Điều tra cơ sở và sau khi có chương trình	Các cấp mẹ và con	Thu nhập, sức khỏe và quyền lực	Có	Có	Không	Không	Có
Tài trợ cho sức khoẻ	Nigeria	Điều tra cơ sở và sau khi có chương trình, và số sách hành chính	Hộ gia đình và trung tâm chi phí và y tế	Sự hoàn lại khá rieng tiếp cận	Không (tại các huyệ)	Có (tai các huyệ)	Không	Không	Sử dụng các định lượng để liệu hành chinh

(Xem tiếp trang sau)

Bảng 4.1 (Tiếp)

Địa điểm	Tên nước	Dạng cơ sở đữ liệu	Đơn vị phân tích	Đo lường kết quả nhiên	Phương pháp kinh tế lượng				Ưu điểm
					Lựa chọn	So sánh	Khác	Biến	
<b>Thực phẩm và định đường</b>									
Đối thực phẩm lấy giáo dục	Bangladesh	Điều tra chỉ tiêu hộ gia đình	Hộ gia đình và cộng đồng	Tỷ lệ di hoc	Không	Không	Không	Có	Không
Y tế, giáo dục và định đường	Méxicô	Điều tra cơ sở và sau khi có chuỗi trình	Hộ gia đình	Sức khỏe, giáo dục và kết quả định hướng	Có	Có	Không	Không rõ	Không khái niệm rõ ràng; phân tích các nhẫn tổng hợp
<b>Cơ sở hạ tầng</b>									
Quỹ đầu tư xã hội	Bolivia	Điều tra cơ sở và sau khi có chuỗi trình	Hộ gia đình, dự án	Các chỉ số giáo dục và y tế	Có	Có	Có	Có	Sử dụng nhieu phương pháp đánh giá
Đường bộ nông thôn	Việt Nam	Điều tra cơ sở và sau khi có chuỗi trình	Hộ gia đình, cộng đồng	Các chỉ số phúc lợi tại cấp hộ giá đình và xã	Có	Có	Có	Có	Đo lường các kết quả phức lạ
<b>Nông nghiệp</b>									
Dự án khuyến nông quốc gia	Kenya	Số liệu lấp danh giá người thu hưởng	Hộ gia đình, nông trại	Năng suất và hiệu quả của các nông trại	Không	Không	Có	Không	Sử hợp lý về mặt chính sách của các kết quả

tương đối khó khăn. Một ví dụ tốt là kỹ thuật tương hợp mức độ xu hướng như đã sử dụng trong dự án TRABAJAR của Argentina nhằm xây dựng một nhóm đối chứng với dữ liệu chéo về các thành viên tham dự chương trình và những người không tham dự.

**Dự án TRABAJAR II ở Argentina.** Dự án này tập trung vào việc cung cấp việc làm với mức lương thấp cho các dự án phụ về cơ sở hạ tầng xã hội và kinh tế có quy mô nhỏ do các nhóm cộng đồng lựa chọn. Đánh giá tác động của chương trình được thiết kế để đánh giá xem thu nhập của những người tham gia chương trình có cao hơn so với trường hợp không có chương trình hay không. Các phương pháp thường được áp dụng để ước lượng thu nhập hộ gia đình khi không có chương trình đã không thể áp dụng được trong trường hợp này do không có sự lựa chọn ngẫu nhiên khi xây dựng nhóm đối chứng để so sánh với thu nhập của những người thụ hưởng dự án, và cũng không có điều tra cơ sở nên không thể tiến hành đánh giá theo kiểu so sánh trước và sau được.

Đánh giá TRABAJAR sử dụng dữ liệu có sẵn để tạo ra một nhóm đối chứng bằng cách xác định những người không tham dự được lấy từ dân số toàn quốc có những đặc điểm gần gũi với những người tham dự chương trình dựa trên một tập hợp các biến số kinh tế xã hội như giáo dục, giới tính, nhà ở, nhận thức chủ quan về phúc lợi, tư cách thành viên trong các đảng chính trị và các hiệp hội khu vực thông qua một kỹ thuật gọi là mức độ xu hướng. Nghiên cứu này chứng tỏ khả năng sử dụng phong phú của dữ liệu điều tra hộ gia đình trên toàn quốc hiện có - điều tra này có tên là Encuesta de Desarrollo Social (EDS)- trong việc tạo ra nhóm đối chứng. Điều tra này được kết hợp với một đợt điều tra có quy mô nhỏ hơn về những người tham dự TRABAJAR được tiến hành nhằm phục vụ riêng cho đánh giá. Đợt điều tra quy mô nhỏ này được thiết kế kỹ lưỡng để sử dụng cùng bảng câu hỏi như EDS, và cũng được thực hiện gần như đồng thời với EDS bằng chính những nhóm

phỏng vấn trong EDS để đảm bảo thực hiện kỹ thuật tương hợp thành công. Kỹ thuật này tỏ ra khả thi trong trường hợp của TRABAJAR bởi vì hiện nay, người ta đang thảo luận về một cuộc điều tra hộ gia đình trên toàn quốc. Các nhà đánh giá có thể sử dụng chính cuộc điều tra này để lấy mẫu nhiều hơn đối với các thành viên của TRABAJAR. Các nhóm phỏng vấn giống nhau được sử dụng cho cả điều tra toàn quốc và điều tra dự án, nhờ đó làm tăng hiệu quả thu thập dữ liệu (xem phụ lục 1.1).

**Đánh giá chương trình thị trường lao động ở Séc.** Đánh giá này cố gắng bao gồm năm chương trình lao động đang hoạt động nhằm, (a) xác định liệu các thành viên trong các chương trình khác nhau có thành công hơn khi tái gia nhập thị trường lao động so với những người không phải là thành viên không, và liệu điều này có khác nhau giữa các nhóm nhỏ và với các điều kiện nhất định của thị trường lao động hay không; và (b) xác định hiệu quả chi phí của mỗi chương trình và đưa ra các gợi ý nhằm cải thiện. Đánh giá này sử dụng một kỹ thuật tương hợp do không thu thập được dữ liệu gốc. Người đánh giá điều tra đối với những người tham dự chương trình và sau đó lựa chọn một mẫu ngẫu nhiên bao gồm những người không tham dự. Vì những người không tham dự nhiều tuổi hơn và ít học vấn hơn một cách hệ thống, nên người đánh giá cần xây dựng một nhóm đối chứng hợp lý cho mỗi chương trình. Điều này được thực hiện nhờ lựa chọn từng người tham dự và so sánh những người này với các cá nhân trong số những người không tham dự trên cơ sở bảy đặc điểm: tuổi tác, giới tính, học vấn, số tháng đi làm trước khi đăng ký với chương trình, quy mô thành phố, tình trạng hôn nhân, và loại công việc cuối cùng. Những cá nhân gần gũi nhất sẽ được đưa vào trong nhóm đối chứng. Mặc dù phương pháp này đơn giản nhưng vẫn có khả năng xảy ra sai lệch - đó là khi nhóm không tham dự khác biệt một cách có ý nghĩa với nhóm tham dự trong các đặc điểm không quan sát được (phụ lục 1.5).

## **Khắc phục khó khăn trong việc phát triển các nhóm đối chứng**

Đôi khi, có thể có nhóm đối chứng ngẫu nhiên hay thử nghiệm nhưng chúng lại không khả thi về mặt chính trị. Trong trường hợp này, có thể tiến hành lựa chọn nhóm đối chứng một cách ngẫu nhiên bằng cách sử dụng bất kỳ kế hoạch nào nhằm thí điểm dự án tại các khu vực hạn chế nhất định. Các vùng thí điểm dự án lúc đầu có thể được lựa chọn ngẫu nhiên, và với các vùng dự án tiềm năng trong tương lai đóng vai trò nhóm đối chứng. Theo thời gian, các cộng đồng bổ sung có thể được đưa vào trong dự án một cách ngẫu nhiên. Có ba ví dụ minh họa cách thức giải quyết tình huống mà trong đó sự lựa chọn ngẫu nhiên không khả thi về mặt chính trị hay vì một lý do nào đó. Ở Việt Nam, một dự án giao thông nông thôn sẽ được đánh giá với các thông tin hạn chế và không có lựa chọn ngẫu nhiên. Quỹ đầu tư xã hội Ôndurat là một ví dụ về cách thức xây dựng một nhóm đối chứng trong các dự án do nhu cầu thúc đẩy, sử dụng phương pháp so sánh tương hợp căn cứ vào một tập hợp dữ liệu chéo duy nhất. Đánh giá các dự án do nhu cầu thúc đẩy có thể đặc biệt khó khăn do không thể biết trước được những dự án hay cộng đồng nào sẽ tham dự vào dự án. Đánh giá thứ ba là đánh giá về Quỹ đầu tư xã hội Bôlivia tại khu vực Chacon. Đây là một ví dụ tốt về cách thức sử dụng phương pháp lựa chọn ngẫu nhiên trong các dự án do nhu cầu thúc đẩy theo một cách thức cho phép định hướng đối tượng của dự án.

**Dự án đường bộ nông thôn ở Việt Nam.** Dự án này nhằm làm giảm tình trạng đói nghèo ở các vùng nông thôn bằng cách tăng cường khả năng tiếp cận các cộng đồng nông thôn và liên kết chúng với mạng lưới đường sá cấp huyện và cấp tỉnh. Thiết kế đánh giá tác động tập trung vào các dữ liệu gốc và điều tra tiếp theo thu thập cho một mẫu tham gia dự án và các cộng đồng thuộc nhóm đối chứng được xác định bằng phương pháp so sánh tương hợp. Các thông tin ban đầu và sau khi có chương

trình về các chỉ số, chẳng hạn như năng suất nông nghiệp của xã, đa dạng hoá nguồn thu nhập, cơ hội việc làm, sự sẵn có của hàng hoá, dịch vụ và cơ sở vật chất, giá trị và sự phân phối tài sản, sẽ được thu thập từ một mẫu ngẫu nhiên bao gồm các xã tham gia dự án (nhóm tham dự) và các xã không tham gia (nhóm đối chứng). Các dữ liệu này sẽ được sử dụng để so sánh việc thay đổi trong các kết quả trước và sau khi có chương trình giữa các xã có dự án và các xã không có dự án bằng phương pháp "khác biệt kép."

Lý tưởng nhất là các xã thuộc hai nhóm này cần phải tương đương nhau trong tất cả các đặc điểm quan sát và không quan sát được của chúng, và sự khác biệt duy nhất giữa chúng là việc các xã trong nhóm tham dự là đối tượng thụ hưởng lợi ích từ chương trình trong khi các xã trong nhóm đối chứng thì không phải. Do không có sự lựa chọn ngẫu nhiên đối với nhóm tham dự và nhóm đối chứng, và vì vào thời điểm lựa chọn mẫu, không có các dữ liệu cần thiết cho việc lựa chọn các xã so sánh phù hợp, nên người ta chọn ra các mẫu ngẫu nhiên của các xã tham dự và các xã không tham dự. Cụ thể, các xã tham dự được lựa chọn từ một danh sách bao gồm tất cả các xã có các dự án dự kiến trong từng tỉnh. Sau đó các xã so sánh được lựa chọn từ một danh sách tất cả các xã còn lại không có dự án dự kiến nhưng ở cùng huyện với xã tham dự chương trình. Sử dụng các thông tin thu thập được để đánh giá, các kỹ thuật tương hợp mức độ xu hướng sẽ được sử dụng để đảm bảo là các xã không tham gia dự án là nhóm đối chứng phù hợp. Bất kỳ các đơn vị so sánh nào có các thuộc tính bất thường so với các xã tham dự đều sẽ bị loại bỏ ra khỏi mẫu (phụ lục 1.15).

**Quỹ đầu tư xã hội Ondurát.** Quỹ đầu tư xã hội Ondurát (FHIS) (xem Ngân hàng Thế giới 1998b) nhằm vào việc nâng cao các điều kiện sống của những nhóm xã hội đói nghèo bằng cách tài trợ cho các dự án nhỏ về cơ sở hạ tầng xã hội và kinh tế. FHIS là một thể chế phục vụ theo nhu cầu. Quỹ này đã đáp ứng sáng kiến của chính quyền các địa phương, các bộ thuộc

chính phủ, các tổ chức phi chính phủ, và các nhóm xã hội bằng cách tài trợ cho việc đầu tư vào cơ sở hạ tầng, trang thiết bị và đào tạo. Đánh giá tác động của FHIS sử dụng các kỹ thuật so sánh tương hợp. Trong đánh giá này, người ta rút ra mẫu của nhóm tham dự một cách ngẫu nhiên từ danh sách các cộng đồng tại đó có các dự án FHIS hoạt động được ít nhất một năm. Mặt khác, nhóm đối chứng lại được lựa chọn từ một danh sách các dự án "đang được chuẩn bị"- các dự án đã yêu cầu và chấp thuận nhưng vẫn chưa nhận được đầu tư từ FHIS. Trên lý thuyết, các cộng đồng trong nhóm đối chứng tự động phù hợp với các cộng đồng trong nhóm tham gia dự án căn cứ theo tiến trình tự lựa chọn và vào các tiêu chí chấp thuận dự án của FHIS. Một đợt điều tra hộ gia đình được tiến hành sau đó ở cả các cộng đồng thuộc nhóm tham dự và nhóm đối chứng. Điều tra này được bổ sung bằng sự đánh giá định tính (dưới hình thức các nhóm tập trung (focus group) và phỏng vấn những người cung cấp thông tin quan trọng) được tiến hành trong một tập hợp nhỏ bao gồm các cộng đồng tham gia chương trình. Đánh giá ban đầu này là bước đầu tiên để tiến tới thiết lập một hệ thống giám sát và đánh giá hiện nay đang hoạt động, hệ thống này sau đó sẽ có thể được kết hợp vào trong hoạt động của FHIS. Cụ thể, dữ liệu thu thập được từ các cộng đồng sẽ có dự án của FHIS trở thành nguồn dữ liệu ban đầu quan trọng, được sử dụng làm căn cứ để theo dõi các thay đổi trong tương lai của các chỉ số tác động, sau khi các cộng đồng này đã thực sự nhận được đầu tư của FHIS.

**Đầu tư cho giáo dục tại khu vực Chaco ở Bolivia.** Các dự án giáo dục được Quỹ đầu tư xã hội Bolivia (SIF) tài trợ nhằm nâng cấp cơ sở vật chất và đào tạo giáo viên tại các trường công ở nông thôn. Sự trì hoãn thực hiện dự án tại Khu vực Chaco và nguồn vốn hạn chế cho việc nâng cấp trường học là dịp để sử dụng các đánh giá thử nghiệm, trong khi vẫn đảm bảo những trường học có nhu cầu bức xúc nhất sẽ được hưởng lợi từ dự án. Các trường học trong Khu vực Chaco được xếp hạng theo

chỉ số chất lượng trường học căn cứ vào tổng cộng năm chỉ số về cơ sở hạ tầng và trang thiết bị của trường học là: số bóng điện, hệ thống nước thải, nguồn nước, số bàn trên mỗi học sinh, và diện tích mét vuông trên mỗi học sinh. Chỉ những trường không đạt được giá trị ngưỡng mới được tiếp nhận chương trình của SIF. Trong số các trường đủ điều kiện, các trường nghèo nhất được tự động lựa chọn để nhận đầu tư của SIF. Nhóm ưu tiên thứ hai bao gồm 120 trường, nhưng ngân sách của Quỹ chỉ đủ để nâng cấp chưa đầy một nửa số trường này. Do vậy, các trường đủ điều kiện trong nhóm ưu tiên thứ hai đã được lựa chọn một cách ngẫu nhiên vào nhóm tham dự và nhóm đối chứng, nhờ đó đáp ứng được các điều kiện cho một thiết kế đánh giá thử nghiệm (Phụ lục 1.4).

### Kết hợp các phương pháp

Đối với hầu hết các đánh giá, cần sử dụng một số kỹ thuật để có được các kết quả vững chắc nhằm trả lời một số câu hỏi đánh giá. Mỗi câu hỏi có thể đòi hỏi các kỹ thuật khác nhau, ngay cả trong cùng một thiết kế dự án. Ba ví dụ sau minh họa cách thức kết hợp một số kỹ thuật trong cùng một đánh giá, đó là, ví dụ về Quỹ xã hội Bôlivia, đánh giá về TRABAJAR ở Áchentina, và Cải cách trường học ở Nicaragua.

**Quỹ xã hội Bôlivia.** Quỹ xã hội thường bao gồm một số loại dự án phụ khác nhau, và do vậy việc thiết kế đánh giá có thể đòi hỏi một số phương pháp. Trong quỹ xã hội Bôlivia, cách thức thực hiện dự án, quyết định các phương pháp đánh giá được sử dụng. Trong các dự án giáo dục, các trường học sẽ nhận được chương trình đã được xác định từ trước, do đó không thể sử dụng phương pháp lựa chọn ngẫu nhiên mà cần đến các phương pháp tương hợp. Trong các dự án y tế, các phương pháp phản thân được sử dụng do chương trình sẽ được thực hiện tại tất cả các trung tâm y tế trong khu vực (xem Phụ lục 1.4).

**Sử dụng kết hợp nhiều thành phần đánh giá- TRABAJAR**

**II ở Áchentina.** Đánh giá TRABAJAR bao gồm một loạt các thành phần được thiết kế để đánh giá mức độ thành công của chương trình trong việc đạt các mục tiêu chính sách. Thành phần thứ nhất, căn cứ vào số liệu điều tra hộ gia đình để đánh giá các lợi ích về thu nhập đối với những người tham dự TRABAJAR. Thành phần thứ hai, giám sát việc phân bổ vốn của chương trình (tính trọng điểm), theo dõi các thay đổi theo thời gian do kết quả của cải cách. Thành phần này tạo ra sự phản hồi định kỳ nửa năm một lần để điều chỉnh trọng tâm của chương trình. Các thành phần đánh giá bổ sung bao gồm một phân tích chi phí- lợi ích về các dự án cơ sở hạ tầng, và đánh giá xã hội nhằm thu thập ý kiến phản hồi của cộng đồng về việc thực thi dự án. Mỗi thành phần này đều được tiến hành hai lần. Ba thành phần tương lai khác cũng được lên kế hoạch. Kỹ thuật nghiên cứu so sánh tương hợp sẽ được áp dụng lại để đánh giá tác động của sự tham dự chương trình TRABAJAR tới hoạt động của thị trường lao động. Chất lượng dự án cơ sở hạ tầng sẽ được đánh giá lại, lần này đối với các dự án đã hoàn thành trong ít nhất một năm để đánh giá sự bền vững, tình trạng bảo dưỡng và tỷ lệ sử dụng. Cuối cùng, một thành phần nghiên cứu định tính sẽ kiểm tra các hoạt động và thủ tục của chương trình bằng cách phỏng vấn các cán bộ làm việc tại các cơ quan bảo trợ dự án, cũng như những người thụ hưởng chương trình.

Các kết quả đánh giá cung cấp định hướng rõ ràng cho sự cải cách chính sách. Thành phần đánh giá thứ nhất cho thấy chương trình TRABAJAR đã rất thành công trong việc định hướng đối tượng là người nghèo- việc tự lựa chọn đối tượng tham dự của chương trình do đặt ra mức lương thấp đã có hiệu quả ở Áchentina, và những người tham dự đã gia tăng được thu nhập nhờ tham dự chương trình. Thành phần đánh giá thứ hai là sự cải thiện theo thời gian trong phân bổ nguồn vốn của chương trình theo khu vực địa lý- chương trình hiện nay đã thành công hơn trong việc định hướng các nguồn vốn tới các khu vực đói nghèo. Tuy nhiên, quá trình đánh giá hiện tại cho thấy có những kết quả khác nhau và các kết quả thường xuyên

yếu kém ở một số tỉnh- những địa phương này hiện nay đang nhận được sự quan tâm nhiều hơn về mặt chính sách. Các kết quả đánh giá đầy thất vọng về chất lượng dự án cơ sở hạ tầng đã thúc đẩy các nỗ lực to lớn của nhóm dự án nhằm cải thiện hoạt động trong lĩnh vực này thông qua việc cải cách chính sách. Cụ thể là, đòi hỏi phải có nhiều lần tham quan thực địa để đánh giá và giám sát hơn, trừng phạt các cơ quan có kết quả yếu kém vào thời điểm kết thúc dự án, và tăng cường các tài liệu đánh giá. Cuối cùng, các đánh giá xã hội đã cho thấy, cần có sự trợ giúp kỹ thuật tốt hơn đối với các NGO và các chính quyền nông thôn trong giai đoạn soạn thảo và thực thi dự án, cũng như cần có sự công khai và minh bạch hơn trong các thông tin về chương trình TRABAJAR (Phụ lục 1.1).

**Cải cách tự chủ trường học ở Nicaragua.** Vào năm 1993, Chính phủ Nicaragua đã tiến hành các bước đi quyết định nhằm thực hiện một sáng kiến phân quyền trong lĩnh vực giáo dục, trao quyền tự chủ trong quản lý và ngân sách cho một số trường tiểu học và trung học. Mục đích của cuộc cải cách này là nâng cao kết quả học tập của học sinh: khi sự quản lý trường trở nên dân chủ và thu hút nhiều sự tham dự hơn, thì có thể định hướng việc quản lý trường học ở địa phương và cách thức chi tiêu nhằm trực tiếp tăng cường tính sư phạm và nâng cao kết quả học tập của học sinh. Ảnh hưởng của cuộc cải cách này hiện đang được đánh giá bằng cách sử dụng kết hợp các kỹ thuật định tính và định lượng nhằm đánh giá các kết quả, cũng như tiến trình phân quyền. Mục tiêu của thành phần định tính là làm rõ liệu có thực sự có các cải cách về quản lý và ngân sách dự kiến tại các trường học không, và tại sao có, hay tại sao không có. Thành phần đánh giá định lượng bổ sung cho kết quả này bằng cách trả lời câu hỏi: "Liệu các thay đổi trong quản lý và ngân sách trường học có thực sự tạo ra kết quả học tập tốt hơn đối với trẻ em hay không?". Kết quả định tính cho thấy, việc thực hiện thành công cải cách này phụ thuộc lớn vào bối cảnh và môi trường của trường học (tức là phụ thuộc vào mức

độ đói nghèo của cộng đồng), trong khi các kết quả định lượng cho thấy, tăng cường quyền ra quyết định của các trường học trên thực tế cũng kèm theo sự cải thiện đáng kể kết quả học tập của học sinh.

Những phương pháp và nguồn dữ liệu khác nhau nhưng bổ sung cho nhau đã được sử dụng để kết hợp cả hai cách tiếp cận này. Một mặt, đánh giá định lượng được tiến hành sau một thiết kế bán thử nghiệm, trong đó điểm thi của một mẫu bao gồm học sinh ở các trường tự chủ (nhóm tham dự) được so sánh với kết quả điểm thi từ một nhóm tương tự bao gồm các trường công và trường tư không tự chủ (nhóm đối chứng). Dữ liệu cho thành phần đánh giá này được thu thập từ một nhóm mẫu trong hai đợt điều tra trường học và hộ gia đình phù hợp và từ kết quả điểm thi của học sinh. Mặc khác, thiết kế đánh giá định tính bao gồm một loạt các cuộc phỏng vấn những người cung cấp thông tin quan trọng và các cuộc thảo luận nhóm tập trung với các cán bộ trường học và các cha mẹ học sinh khác nhau trong một mẫu nhỏ bao gồm các trường tự chủ và các trường truyền thống nằm trong nghiên cứu định lượng.

Việc sử dụng cả hai kỹ thuật nghiên cứu định tính và định lượng đã tạo ra sự kết hợp có giá trị của các kết quả hữu ích và hợp lý về mặt chính sách. Nghiên cứu định lượng cho phép có sự khái quát chung, hợp lệ về mặt thống kê về điều kiện và kết quả của các trường học; nghiên cứu định tính nâng cao các kết quả này nhằm xác định xem tại sao đạt thành công trong một số kết quả dự kiến của chương trình nhưng lại thất bại với các kết quả khác, và do đó, góp phần chỉ dẫn cho sự điều chỉnh chính sách. Hơn nữa, vì có tính trực quan nên nghiên cứu định tính dễ tiếp cận hơn, và vì vậy, lý thú hơn đối với cán bộ của bộ, do đó khuyến khích xây dựng năng lực nhanh chóng và tạo ra uy tín cho tiến trình đánh giá ở trong bộ (Phụ lục 1.11).

### Khai thác các nguồn dữ liệu hiện có

Các nguồn dữ liệu hiện có, như điều tra hộ gia đình toàn quốc,

điều tra dân số, sổ sách hành chính của chương trình, hay dữ liệu của chính quyền địa phương, có thể cung cấp các đầu vào có giá trị cho nỗ lực đánh giá. Việc sử dụng các nguồn dữ liệu hiện có làm giảm nhu cầu thu thập dữ liệu tốn kém chỉ để phục vụ đánh giá, như được minh họa trong tình huống đánh giá về Đường nông thôn ở Việt Nam. Hơn nữa, mặc dù các dữ liệu hiện có có thể không chứa đựng tất cả thông tin mà chúng ta cần phải thu thập cho đánh giá, nhưng các kỹ thuật đánh giá sáng tạo có thể bù đắp cho những dữ liệu thiếu, như được chỉ ra trong Dự án Khuyến nông toàn quốc ở Kenya.

**Dự án đường bộ nông thôn ở Việt Nam.** Dữ liệu sử dụng cho đánh giá này căn cứ vào sự kết hợp giữa những nguồn dữ liệu hiện có ở cấp quốc gia và địa phương với các kỳ điều tra được tiến hành chỉ để phục vụ cho đánh giá. Điều tra hộ gia đình được thiết kế một cách có hiệu quả, lặp lại một số câu hỏi trong Điều tra mức sống dân cư, để từ thông tin chung của hai cuộc điều tra này, có thể sử dụng các kỹ thuật hồi quy nhằm ước lượng vị trí của hộ gia đình trong phân phối phúc lợi trên toàn quốc.

Đánh giá này căn cứ chủ yếu vào dữ liệu cấp xã được các xã thu thập hàng năm, trong đó đề cập tới các đặc điểm về nhân khẩu, tình hình sử dụng đất, và hoạt động sản xuất. Nguồn dữ liệu này được bổ sung bằng một đợt điều tra cấp xã được tiến hành nhằm phục vụ riêng cho đánh giá. Hai cơ sở dữ liệu bổ sung được thiết lập trên cơ sở sử dụng thông tin sẵn có. Một cơ sở dữ liệu phong phú cấp tỉnh đã được thiết lập để giúp hiểu được lý do lựa chọn các tỉnh trong dự án. Cơ sở dữ liệu này bao gồm tất cả các tỉnh của Việt Nam và bao gồm dữ liệu về nhiều biến kinh tế xã hội. Cuối cùng, một cơ sở dữ liệu cấp độ dự án cho mỗi lĩnh vực dự án được điều tra cũng được xây dựng nhằm kiểm chứng quy mô của dự án và phương pháp thực hiện trong đánh giá tác động của dự án (Phụ lục 1.15).

**Dự án khuyến nông quốc gia của Kenya (NEP).** Hoạt động của Dự án khuyến nông quốc gia (NEP) gây ra nhiều tranh cãi

và là một phần trong cuộc tranh luận lớn hơn về tính hiệu quả chi phí của phương pháp đào tạo và tham quan (T&V) trong dịch vụ khuyến nông. Trong bối cảnh của Kênia, cuộc tranh luận này còn trở nên gay gắt hơn, một phần do có một nghiên cứu (Bindlish, Evenson và Gbetibouo 1993, 1997) ước tính tỷ lệ doanh lợi của phương pháp T&V rất cao, trong khi trên thực tế thiếu các kết quả có tính thuyết phục, kể cả hiện trạng tồi tệ của ngành nông nghiệp ở Kênia trong những năm gần đây.

Sự không nhất trí về hoạt động của NEP đã làm trì hoãn kết quả của đánh giá này. Đánh giá sử dụng một phương pháp thực nghiệm mạnh mẽ để đánh giá sự phát triển thể chế của chương trình và tác động tới kết quả nông nghiệp. Đánh giá sử dụng kết hợp các phương pháp định tính và định lượng nhằm đặt ra những câu hỏi liên quan tới chính sách, và cho thấy những yếu kém nghiêm trọng trong chương trình. Đó là: (a) Sự phát triển thể chế của NEP là hạn chế, và sau 15 năm, vẫn có rất ít thành công trong hiệu quả của dịch vụ khuyến nông; (b) chất lượng và số lượng cung cấp dịch vụ khuyến nông yếu kém; (c) các dịch vụ khuyến nông chỉ có tác động tích cực nhỏ tới hiệu suất nông nghiệp và không có tác động gì tới năng suất nông nghiệp.

Đánh giá có thể rút ra một số kết luận chính sách cụ thể từ các kết quả này, trong đó có nhiều kết luận liên quan tới thiết kế của các dự án khuyến nông trong tương lai. Thứ nhất, đánh giá cho thấy, cần tăng cường tính định hướng đối tượng của phương pháp T&V, tập trung vào những lĩnh vực và nhóm dân cư mà phương pháp này có thể tạo ra tác động lớn nhất. Hơn nữa, cần điều chỉnh các nhu cầu tư vấn một cách cẩn thận để đáp ứng nhu cầu của nông dân, có tính đến sự khác biệt về điều kiện công nghệ và kinh tế của địa phương. Để đạt được mức độ định hướng dịch vụ như thế, đòi hỏi phải có các luồng thông tin phù hợp, đúng lúc và đáng tin cậy - do đó, cần có một hệ thống M&E của chương trình để tạo ra một luồng cố định bao gồm các thông tin phản hồi về nội dung dịch vụ từ những người thụ hưởng. Nhằm nâng cao hiệu năng của chương

trình, hoạt động huấn luyện và thăm quan cần mỏng hơn và ít căng thẳng hơn, nhưng lại phải bao phủ một phạm vi rộng hơn, thì sẽ nâng cao được hiệu quả chi phí. Cách tiếp cận sử dụng một phương pháp duy nhất của chương trình (thăm quan các nông trang) để chuyển tải các thông điệp tiêu chuẩn, cũng làm hạn chế hiệu quả của chương trình. Sự cải cách thể chế có thể nâng cao hiệu quả cung cấp dịch vụ. Phân quyền trong thiết kế chương trình, kể cả tạo ra các cơ chế tham dự cho phép nông dân có tiếng nói (ví dụ như việc chia sẻ chi phí và các tổ chức nông dân) cũng cần trở thành một thành phần trong cơ chế cung cấp dịch vụ. Cuối cùng, sự hoàn lại chi phí, dù chỉ là một phần, sẽ tạo ra động cơ hợp lý, giúp giải quyết được các vấn đề về trách nhiệm giải trình và kiểm chứng chất lượng, làm cho các dịch vụ có tính định hướng theo nhu cầu và đáp ứng nhu cầu tốt hơn, đồng thời cải thiện về ngân sách (Phụ lục 1.8).

### Các chi phí và tài trợ

Rõ ràng có nhiều chi phí liên quan tới việc tiến hành đánh giá tác động. Đây là lý do tại sao một số nước lại ngần ngại khi tài trợ cho các nghiên cứu này. Các chi phí bao gồm chi phí thu thập dữ liệu và giá trị thời gian của tất cả thành viên trong nhóm đánh giá. Tài trợ cho đánh giá tác động có thể có từ bên trong dự án, từ những nguồn khác của chính phủ, từ tài trợ nghiên cứu, hay từ một nhà tài trợ bên ngoài. Thông tin cho một mẫu, gồm các đánh giá của Ngân hàng Thế giới cho thấy, mặc dù nhiều nước đảm nhận phần lớn chi phí đánh giá nhưng việc thực hiện thành công đánh giá tác động đòi hỏi có các nguồn lực đáng kể từ bên ngoài, nhiều hơn những gì đã được cung cấp trong khoản vay hay tín dụng của một dự án. Số tiền này xuất phát từ sự kết hợp các nguồn sau: (a) khoản cho vay của Ngân hàng Thế giới hay tài trợ tín dụng cho việc thu thập và xử lý số liệu; (b) chính phủ, thông qua tiền lương trả cho các cán bộ địa phương được giao nhiệm vụ thực hiện đánh giá (như giải thích trong bảng 4.1, các chi phí cán bộ này không được đưa

vào trong tính toán chi phí đánh giá ở trong các nghiên cứu tình huống trong sách do hạn chế về số liệu); (c) các khoản tài trợ nghiên cứu của Ngân hàng Thế giới và các khoản viện trợ song phương nhằm tài trợ việc hỗ trợ kỹ thuật của các nhà tư vấn có kinh nghiệm đặc thù cho đánh giá; và (d) ngân sách hành chính của Ngân hàng Thế giới được tính bằng thời gian của các cán bộ Ngân hàng Thế giới trong việc hướng dẫn đánh giá tác động và thường tham gia tích cực vào công việc phân tích.

Mặc dù rất ít đánh giá tác động cho biết về chi phí thực hiện công việc này, nhưng bảng 4.2 đã đưa ra ước tính chi phí cho một mẫu bao gồm các đánh giá tác động có sự tham gia của Ngân hàng Thế giới. Tuy vậy, ước tính chi phí này không bao gồm giá trị thời gian mà cán bộ Ngân hàng Thế giới đóng góp cho nước chủ nhà (giá trị này có thể khá lớn) vì không có thông tin. Trong tám trường hợp nêu ra ở đây, việc có đến năm cán bộ tham dự vào công việc đánh giá trong vài năm không phải là chuyện hiếm. Mức độ tham dự như vậy đã làm tăng đáng kể chi phí đánh giá trong nhiều trường hợp.

Chi phí trung bình ước tính của đánh giá dự án là 433.000 đôla Mỹ. Số liệu này phản ánh sự thay đổi từ 263.000 đôla Mỹ cho đánh giá một chương trình đào tạo kỹ năng nghề nghiệp cho các thanh niên thất nghiệp ở Totoridát và Tobagó, tới 878.000 đôla Mỹ cho đánh giá Quỹ Đầu tư xã hội Bolivia. Chi tiêu cho đánh giá tác động của các dự án này tính trung bình bằng 0,6 phần trăm tổng chi phí của dự án (chi phí này đôi khi bao gồm cả tài trợ từ một số nhà tài trợ), hay 1,3 phần trăm chi phí các khoản vay của Ngân hàng Quốc tế cho Tái thiết và Phát triển (IBRD) hay tín dụng của Hiệp hội Phát triển Quốc tế (IDA). Chi phí tối kém nhất trong các đánh giá nêu ra ở đây là chi phí thu thập dữ liệu và chi phí tư vấn, kể cả tư vấn trong nước và nước ngoài. Trong nhiều trường hợp, chi phí đi lại bao gồm cả các chuyến đi lại của cán bộ trong nước để gặp gỡ cán bộ của Ngân hàng Thế giới và các nhà nghiên cứu tại các cuộc họp chiến lược và các khóa huấn luyện, bởi vì việc xây dựng năng lực cho cán bộ của nước chủ nhà cũng là một mục tiêu

**Bảng 4.2. Tóm tắt các chi phí ước tính từ các đánh giá tác động của Ngân hàng Thế giới**

Ước tính chi phí đánh giá (\$) <sup>a</sup>	Tỷ lệ phần trăm chi phí tổng chi phí dự án*	Tỷ lệ phần trăm trên giá trị khoản vay IBRD hay tín dụng IDA*	Tỷ lệ phần trăm chi phí trên giá trị khoản vay IBRD hay tín dụng IDA*	Chia chi phí đánh giá theo (%)			
				Di lại	Cán bộ của Ngân hàng Thế giới	Nhà tư vấn	Thu thập dữ liệu
Quản lý trường học ở Nicaragoa	495.000	1,26	1,5	8,1	18,1	39,0	34,8
Quản lý trường học ở En Xanvado	443.000	0,60	1,3	7,7	7,4	25,8	59,2
Chương trình phiếu (coupon) ở Côlômbia	266.000	0,20	0,3	9,4	9,8	21,8	59,0
Quỹ xã hội ở Ôndurat	263.000	0,23	0,9	3,0	11,5	53,2	32,3
Quỹ xã hội ở Nicaragoa	449.000	0,30	0,8	4,9	33,0	7,8	55,7
Quỹ xã hội ở Bôlivia	878.000	0,50	1,4	3,4	14,6	12,9	69,1
Đào tạo thanh niên ở Trinidát và Tôbagô	238.000	0,80	1,2	7,6	11,5	17,9	63,1
Trung bình	433.000	0,56	1,0	6,3	15,1	25,5	53,3

a. Chi phí này không bao gồm chi phí của các nhóm đối tác trong nước không được tài trợ bằng vốn vay hay tín dụng. Số liệu đề cập tới khoảng thời gian lựa chọn dự án trong mẫu đánh giá, chứ không phải tổng tài trợ được Ngân hàng thế giới hay các nhà tài trợ khác cung cấp cho những thể chế này.

b. Chi phí này được tính như là tỷ lệ phần trăm của khoản vay hay tín dụng hay của dự án được trình bày chỉ để tham khảo. Trong nhiều trường hợp, tài trợ thực sự cho đánh giá có được từ các nguồn khác với tài trợ dự án.

c. Các ước tính về chi phí di lại bao gồm các chuyến công vụ của cán bộ Ngân hàng Thế giới và các nhà tư vấn quốc tế tới nước chủ nhà, cũng như các chuyến đi lại của các đối tác trong nước, đặc biệt là khi để tham dự các kỳ họp chiến lược và các hội thảo phân tích cùng với các nhà tư vấn quốc tế và các cán bộ của Ngân hàng Thế giới.

Nguồn: Các tài liệu lưu trữ dự án của Ngân hàng Thế giới.

quan trọng. Hai ví dụ dưới đây về đánh giá tác động của các dự án ở Totoridát và Tôbagô, và Bôlivia minh họa một số thay đổi có thể xuất hiện trong chi phí của chương trình.

Đánh giá chương trình đào tạo kỹ năng nghề nghiệp ở Totoridát và Tôbagô tận dụng một cuộc điều tra về thu nhập và việc làm quốc gia để lấy mẫu nhiều hơn đối với những người tốt nghiệp chương trình và tạo một nhóm đối chứng từ một tập hợp con trong mẫu toàn quốc. Hơn nữa, nhóm đánh giá đã giúp đỡ thiết kế và sử dụng các dữ liệu hành chính sẵn có từ những ghi chép về những người đăng ký tham gia chương trình, nhờ đó có được dữ liệu trước khi có chương trình và không cần phải liệt kê xác định những người tốt nghiệp từ chương trình. Quy mô mẫu tổng cộng đối với ba nghiên cứu theo dõi là gần bằng 2500 thanh niên, tính cả nhóm tham dự và nhóm đối chứng. Chỉ có một bảng câu hỏi ngắn được sử dụng trong điều tra, và bảng câu hỏi này chỉ được áp dụng với những người tốt nghiệp chương trình. Totoridát và Tôbagô là một nước nhỏ, có thể dễ tiếp cận các cộng đồng bằng đường ô tô, và tiếng Anh được sử dụng phổ biến trên toàn quốc và trong số những người tốt nghiệp chương trình.

Đánh giá Quỹ Xã hội Bôlivia (SIF) sử dụng các điều tra cơ sở và tiếp theo về các nhóm tham dự và nhóm đối chứng nhằm đánh giá hoạt động của chương trình trong các lĩnh vực y tế, giáo dục, nước, và vệ sinh. Không có sẵn các điều tra trên toàn quốc để tiến hành phân tích hay thực hiện lấy mẫu nhiều hơn, và điều này làm cho việc thu thập dữ liệu trong đánh giá trở nên rất nặng nề. Mẫu bao gồm nhóm tham dự và nhóm đối chứng và có gần 7.000 hộ gia đình và 300 cơ quan được phỏng vấn trong hai lần điều tra: điều tra cơ sở năm 1993 và điều tra tiếp theo năm 1998.

Tại Bôlivia, các công cụ thu thập dữ liệu để đánh giá tác động bao gồm phỏng thí nghiệm lưu động để kiểm định chất lượng nước, các bài kiểm tra kết quả học tập, và tám bảng câu hỏi cho những người cung cấp thông tin từ các hộ gia đình và các cơ quan. Tám bảng câu hỏi này bao gồm hai bảng câu hỏi

cho các hộ gia đình (một cho những người cung cấp thông tin chính và một cho phụ nữ ở độ tuổi sinh đẻ), một bảng câu hỏi cho cộng đồng, bốn bảng câu hỏi khác nhau cho các trung tâm y tế, căn cứ vào các loại hình trung tâm y tế khác nhau (từ các trạm xá nhỏ trong cộng đồng cho tới bệnh viện), và một bảng câu hỏi trường học dành cho hiệu trưởng. Để đánh giá tính định hướng đối tượng, đánh giá còn có một thước đo nghèo khổ dựa theo tiêu dùng. Thước đo này đòi hỏi phải thu thập dữ liệu chi tiết về tiêu dùng từ các hộ gia đình và dữ liệu giá cả trong khu vực từ các cộng đồng. Công việc thực địa được tiến hành ở vùng nông thôn, nơi tập trung đa số các dự án của SIF và bao gồm một mẫu ngẫu nhiên có các hộ gia đình nông thôn mà chỉ có thể tiếp cận bằng đi bộ hay đi ngựa. Cuối cùng, các bảng câu hỏi phải được xây dựng và thực hiện bằng các thứ tiếng Tây Ban Nha, Quechua, và Aymara.

### Các vấn đề kinh tế chính trị

Có một số vấn đề kinh tế chính trị không chỉ tác động tới việc đánh giá có được tiến hành hay không, mà còn tác động tới cách thức tiến hành đánh giá. Quyết định tiến hành một đánh giá tuỳ thuộc nhiều vào sự ủng hộ chính trị mạnh mẽ. Nhiều chính phủ không nhận thấy giá trị của đánh giá dự án, và do đó không muốn đầu tư nguồn lực vào việc này. Hơn nữa, chính phủ có ngần ngại khi cho phép đánh giá độc lập mà sự đánh giá đó có thể thu được kết quả trái ngược với chính sách của chính phủ, đặc biệt là trong các chế độ độc đoán và khép kín. Tuy nhiên, các chính phủ cởi mở hơn có thể xem đánh giá và sự phổ biến kết quả đánh giá như là một phần quan trọng trong tiến trình dân chủ.

Đánh giá cũng nhạy cảm với các thay đổi chính trị. Ba trong số tám nghiên cứu tác động được trình bày trong bảng 4.2 đã phải huỷ bỏ vì các vấn đề kinh tế chính trị. Sự thay đổi chế độ hay thay đổi các vị trí quan trọng tại một cơ quan chính phủ đổi tác và sự biến chuyển trong các chiến lược, chính sách, có

thể tác động không chỉ tới nỗ lực đánh giá, mà cơ bản hơn, tới cả việc thực thi chương trình được đánh giá. Một ví dụ về rủi ro này là từ kinh nghiệm của một nhóm thiết kế và đánh giá tác động của một chương trình thí điểm quản lý trường học ở Péru như là một phần trong dự án về giáo dục tiểu học do Ngân hàng Thế giới tài trợ. Nhóm này bao gồm các quan chức của Bộ Giáo dục, cán bộ Ngân hàng Thế giới, và các nhà tư vấn quốc tế và trong nước. Nhóm đã làm việc hơn một năm để phát triển các mô hình quản lý tự chủ trường học thí điểm, xác lập thiết kế thử nghiệm, thiết kế các công cụ điều tra và kiểm tra kết quả, và thu thập dữ liệu cơ sở về đặc điểm của trường học và kết quả của học sinh. Trước khi áp dụng thí điểm chương trình ở những trường học được lựa chọn một cách ngẫu nhiên, các quan chức chính phủ cấp cao đã huỷ bỏ thử nghiệm quản lý trường học này do thử nghiệm này có lẽ đã bị thất bại về mặt chính trị. Một cải cách tương tự đã được áp dụng vài năm sau đó, nhưng không có được lợi ích từ sự kiểm tra thí điểm hay đánh giá.

Tại Vénéxuêla, một đánh giá về sức khoẻ sinh sản và sức khoẻ trẻ em và về chương trình dinh dưỡng đã phải thiết kế lại ba lần với ba đối tác khác nhau, do chính phủ thay đổi trách nhiệm phụ trách đánh giá từ cơ quan này sang cơ quan khác. Mỗi lần thay đổi lại kèm theo việc đàm phán lại hợp đồng với hằng tư nhân được thuê thu thập dữ liệu và tiến hành phần lớn phân tích cho đánh giá. Khi bắt đầu xảy ra vấn đề liên quan tới tính hợp thức của đối tác chính phủ thứ ba, hằng này bèn thực hiện vô hiệu hoá hợp đồng, và việc đánh giá đã phải bị huỷ bỏ. Các sự kiện này xảy ra trong một thời kỳ náo động chính trị được đặc trưng bằng vô số các cuộc cải tổ nội các, mà kết cục là sự sụp đổ của chính phủ dân cử đắm nhận vai trò đối tác trong dự án, vì vậy, đánh giá khó mà tránh khỏi tác động của sự bất ổn chính trị. Tuy vậy, trong cả hai trường hợp ở Péru và Vénéxuêla, một điều an ủi là mức độ nguồn lực bị lãng phí không quá lớn. Một ví dụ đáng buồn hơn là tác động của thay đổi chính sách tới các chiến lược đánh giá ở En-

Xanvađo, nơi mà thành công được công nhận của cải cách tại các trường học nông thôn đã khuyến khích chính phủ đưa ra áp dụng ngay lập tức chương trình cải cách giáo dục tương tự tại tất cả các trường học ở thành phố, trong khi lẽ ra nên phân kỳ ngẫu nhiên để thực hiện như dự kiến ban đầu. Quyết định này đã loại trừ khả năng sử dụng thiết kế thử nghiệm và làm cho phương pháp so sánh phản thân - một phương pháp ít tin cậy - trở thành phương án thiết kế đánh giá duy nhất có thể trong trường hợp này.

---

---

## Tài liệu tham khảo

Từ “đã xử lý” mô tả các công trình nghiên cứu đã được tái bản một cách không chính thức có thể không sẵn có trong các thư viện

- Abadie, A., J. Angrist, and G. Imbens. 1998. “Instrumental Variables Estimation of Quartile Treatment Effects.” National Bureau of Economic Research Working Paper Series, No. 229:1–28.
- Atkinson, Anthony. 1987. “On the Measurement of Poverty.” *Econometrica* 55: 749–64.
- Bamberger, Michael. 2000. *Integrating Quantitative and Qualitative Methods in Development Research*. Washington, D.C.: World Bank.
- Barnes, Carolyn. 1966. “Assets and the Impact of Microenterprise Finance Programs.” USAID AIMS Project Brief 6. Washington, D.C.: USAID.
- Barnes, Carolyn, and Erica Keogh. 1999. “An Assessment of the Impact of Zambuko’s Microenterprise Program in Zimbabwe: Baseline Findings.” USAID AIMS Project Brief 23. Washington, D.C.: USAID.
- Benus, Jacob, Neelima Grover, and Recep Varcin. 1998. *Turkey: Impact of Active Labor Programs*. Bethesda, Md.: Abt Associates.
- Benus, Jacob, Neelima Grover, Jiri Berkovsky, and Jan Rehak. 1998. *Czech Republic: Impact of Active Labor Programs*. Bethesda, Md.: Abt Associates.
- Besharov, Douglas J., Peter Germanis, and Peter H. Rossi. 1997.

- Evaluating Welfare Reform: A Guide for Scholars and Practitioners.* College Park, Md.: University of Maryland.
- Bindlish, V., R. Evenson, and Mathurin Gbetibouo. 1993. *Evaluation of T & V-Based Extension in Burkina Faso.* Washington, D.C.: World Bank.
- Bloom, Howard S., Larry L. Orr, George Cave, Stephen H. Bell, Fred Doolittle, and Winston Lin. 1994. *The National JTPA Study: Overview, Impacts, Benefits, and Costs of Title II-A.* Bethesda, Md.: Abt Associates.
- Blundell, Richard W., and R. J. Smith. 1993. "Simultaneous Microeconometric Models with Censoring or Qualitative Dependent Variables." In G. S. Maddala, C. R. Rao, and H. D. Vinod, eds., *Handbook of Statistics.* Vol. 11. Amsterdam: North Holland.
- Bourgignon, Francois, Jaime de Melo, and Akiko Suwa. 1991. "Distributional Effects of Adjustment Policies: Simulations for Archetype Economies in Africa and Latin America." *World Bank Economic Review* 5 (2): 339-66.
- Burtless, Gary. 1995. "The Case for Randomized Field Trials in Economic and Policy Research." *Journal of Economic Perspectives* 9 (Spring): 63-84.
- Card, David, and Philip K. Robins. 1996. "Do Financial Incentives Encourage Welfare Recipients to Work? Evidence from a Randomized Evaluation of the Self-Sufficiency Project." National Bureau of Economic Research Paper 5701, August. Cambridge, Mass.: NBER.
- Carvalho, Soniya, and Howard White. 1994. "Indicators for Monitoring Poverty Reduction." *World Bank Discussion Papers* 254. Washington, D.C.: World Bank.
- Chen, Martha Alter, and Donald Snodgrass. 1999. "An Assessment of the Impact of Sewa Bank in India: Baseline Findings." USAID AIMS Project, Processed. Washington, D.C.: USAID.

- Cohen, Monique, and Gary Gaile. 1997. "Highlights and Recommendations of the Virtual Meeting of the CGAP Working Group on Impact Assessment Methodologies." USAID AIMS Project. Washington, D.C.: Management Systems International.
- . 1998. "Highlights and Recommendations of the Second Virtual Meeting of the CGAP Working Group on Impact Assessment Methodologies." USAID AIMS Project. Washington, D.C.: Management Systems International.
- Dar, Amit, and Indermit S. Gill. 1998. "Evaluating Retraining Programs in OECD Countries: Lessons Learned." *The World Bank Research Observer* 13 (February): 79–101.
- Dar, Amit, and Zafiris Tzannatos. 1999. "Active Labor Market Programs: A Review of Evidence from Evaluations." World Bank Social Protection Discussion Paper 9901. Washington, D.C.: World Bank.
- Dehejia, Rajeev H., and Sadek Wahba. 1998. "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs." NBER Working Paper Series 6586. Cambridge, Mass.: NBER  
<http://www.nber.org/papers/w6586>.
- Dennis, Michael L., and Robert F. Boruch. 1989. "Randomized Experiments for Planning and Testing Projects in Developing Countries: Threshold Conditions." *Evaluation Review* 13 (June): 292–309.
- Diagne, Aliou, and Manfred Zeller. 1998. "Determinants of Access to Credit and Its Impact on Income and Food Security in Malawi." Manuscript submitted to the International Food Policy Research Institute (IFPRI) Publication Review Committee for consideration as an IFPRI Research Report.
- Diop, F., A. Yazbeck, and R. Bitran. 1995. "The Impact of Alternative Cost Recovery Schemes on Access and Equity in Niger." *Health Policy and Planning* 10 (3): 223–40.
- Donecker, Jane, and Michael Green. 1998. *Impact Assessment in*

- Multilateral Development Institutions.* London: Department for International Development.
- Dunn, Elizabeth. 1999. "Microfinance Clients in Lima, Peru: Baseline Report for AIMS Core Impact Assessment." USAID AIMS Project, Processed. Washington, D.C.: USAID.
- Edgecomb, Elaine L., and Carter Garber. 1998. "Practitioner-Led Impact Assessment: A Test in Honduras." USAID AIMS. Washington, D.C.: USAID.
- Ezemenari, Kene, Anders Ruqvist, and K. Subbarao. 1999. "Impact Evaluation: A Note on Concepts and Methods." World Bank Poverty Reduction and Economic Management Network, Processed. Washington, D.C.: World Bank.
- Foster, James, J. Greer, and Erik Thorbecke. 1984. "A Class of Decomposable Poverty Measures." *Econometrica* 52: 761-65.
- Friedlander, Daniel, and Gayle Hamilton. 1993. *The Saturation Work Initiative Model in San Diego: A Five-Year Follow-Up Study.* New York: Manpower Demonstration Research Corporation.
- Friedlander, Daniel, and Philip K. Robins. 1995. "Evaluating Program Evaluations: New Evidence on Commonly Used Nonexperimental Methods." *American Economic Review* 85 (September): 923-37.
- Friedlander, Daniel, David H. Greenberg, and Philip K. Robins. 1997. "Evaluating Government Training Programs for the Economically Disadvantaged." *Journal of Economic Literature* 35 (December): 1809-55.
- Fuller, Bruce, and Magdalena Rivarola. 1998. *Nicaragua's Experiment to Decentralize Schools: Views of Parents, Teachers and Directors.* Working Paper Series on Impact Evaluation of Education Reforms, paper no. 5. Washington, D.C.: World Bank.
- Gaile, Gary L., and Jenifer Foster. 1996. "Review of Methodological Approaches to the Study of the Impact of

- MicroEnterprise Credit Programs." USAID AIMS Brief 2. Washington, D.C.: USAID
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 1998. "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya." Development Research Group, World Bank, October 28.
- Goodman, Margaret, Samuel Morley, Gabriel Siri, and Elaine Zuckerman. 1997. *Social Investment Funds in Latin America: Past Performance and Future Role*. Washington D.C.: Inter American Development Bank. March.
- Government of Denmark. 1995. *Methods for Evaluation of Poverty Oriented Aid Interventions*. Copenhagen: Ministry of Foreign Affairs.
- Greene, W. H. 1997. *Econometric Analysis*. Hemel Hempstead, New Jersey: Prentice Hall Press.
- Greenberg, David, and Mark Shroder. 1997. *The Digest of Social Experiments*, 2nd ed. Washington, D.C.: The Urban Institute Press.
- Grosh, Margaret E., and Juan Muñoz. 1996. "A Manual for Planning and Implementing the Living Standards Measurement Study Survey." LSMS Working Paper #126. World Bank, Washington, D.C.
- Grossman, Jean Baldwin. 1994. "Evaluating Social Policies: Principles and U.S. Experience." *The World Bank Research Observer* 9 (July): 159–80.
- Habicht, J. P., C. G. Victoria, and J. P. Vaughan. 1999. "Evaluation Designs for Adequacy, Plausibility and Probability of Public Health Programme Performance and Impact." *International Journal of Epidemiology* 28: 10–18.
- Harrell, Adele, Martha Burt, Harry Hatry, Shelli Rossman, Jeffrey Roth, and William Sabol. 1996. *Evaluation Strategies for Human Service Programs: A Guide for Policymakers and Providers*. Washington, D.C.: Urban Institute.

- Heckman, James. 1997. "Instrumental Variables. A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *Journal of Human Resources* 32 (3): 441-61.
- Heckman, James, R. Lalonde, and J. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, vol. III. New York: Elsevier Science Publishing Co.
- Heckman, James, and Richard Robb. 1985. "Alternative Methods of Evaluating the Impact of Interventions: An Overview." *Journal of Econometrics* 30: 239-67.
- Heckman, James J., and Jeffrey A. Smith. 1995. "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9 (2): 85-110.
- Heckman, J., H. Ichimura, J. Smith, and P. Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66: 1017-99.
- Hicks, Norman. 1998. *Measuring the Poverty Impact of Projects in LAC*. World Bank Latin America and the Caribbean Region, Processed. Washington, D.C.: World Bank.
- Holder, Harold D., Andrew J. Trelo, Robert F. Saltz, and Joel W. Grube. 1997. "Summing Up: Recommendations and Experiences for Evaluation of Community-Level Prevention Programs." *Evaluation Review* 21 (April): 268-78.
- Hollister, Robinson G., and Jennifer Hill. 1995. *Problems in the Evaluation of Community-Wide Initiatives*. New York: Russell Sage Foundation.
- Hulme, David. 1997. "Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice." Manchester, U.K.: Institute for Development Policy and Management, University of Manchester.
- International Food Policy Research Institute. 1998. *Programa Nacional de Educaciün, Salud, y Alimentaciün (PROGRESA)*: A

- Proposal for Evaluation* (with technical appendix). Washington, D.C.: IFPRI.
- Jalan, Jyotsna, and Martin Ravallion. 1998a. "Income Gains from Workfare: Estimates for Argentina's TRABAJO Program Using Matching Methods." Washington D.C.: Development Research Group, World Bank.
- . 1998b. "Transfer Benefits from Workfare: A Matching Estimate for Argentina." Washington, D.C.: Development Research Group, World Bank.
- Jamison, Dean T., Barbara Serle, Klaus Galda, and Stephen P. Heyneman. 1981. "Improving Elementary Mathematics Education in Nicaragua: An Experimental Study of the Impact of Textbooks and Radio on Achievement." *Journal of Educational Psychology* 73 (4): 556-67.
- Karoly, Lynn A., Peter W. Greenwood, Susan S. Everingham, Jill Houš, M. Rebecca Kilburn, C. Peter Rydell, Matthew Sanders, and James Chiesa. 1998. *Investing in Our Children: What We Know and Don't Know about the Costs and Benefits of Early Childhood Interventions*. Santa Monica, Calif.: Rand.
- Kemple, James J., Fred Doolittle, and John W. Wallace. 1993. *The National JTPA Study: Site Characteristics and Participation Patterns*. New York: Manpower Demonstration Research Corporation.
- Khandker, Shahidur R. 1998. *Fighting Poverty with Microcredit: Experience in Bangladesh*. New York: Oxford University Press for the World Bank.
- Killick, Tony. 1995. *IMF Programmes in Developing Countries, Design and Impact*. London: England.
- King, Elizabeth, and Berk Ozler. 1998. *What's Decentralization Got to Do with Learning? The Case of Nicaragua's School Autonomy Reform*. Working Paper Series on Impact Evaluation of Education Reforms, paper no. 9. Washington, D.C.: World Bank.

- King, Elizabeth, Berk Ozler, and Laura Rawlings. 1999. *Nicaragua's School Autonomy Reform: Fact or Fiction?* Washington, D.C.: World Bank.
- Kish, Leslie. 1965. *Survey Sampling*. New York: John Wiley and Sons.
- Levinson, James F., Beatrice Lorge Rogers, Kristin M. Hicks, Thomas Schaetzel, Lisa Troy, and Collette Young. 1998. *Monitoring and Evaluation: A Guidebook for Nutrition Project Managers in Developing Countries*. Medford, Mass.: International Food and Nutrition Center, Tufts University School of Nutrition Science and Policy, March.
- Manski, Charles, and Irwin Garfinkel, eds. 1992. *Evaluating Welfare and Training Programs*. Cambridge, Mass.: Harvard University Press.
- Manski, Charles, and Steven Lerman. 1977. "The Estimation of Choice Probabilities from Choice-Based Samples." *Econometrica* 45: 1977-88.
- McKay, H., A. McKay, L. Siniestra, H. Gomez, and P. Lloreda. 1978. "Improving Cognitive Ability in Chronically Deprived Children" *Science* 2000(21): 270-78.
- Meyer, Bruce D. 1995. "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistics*, 13 (April): 151-161.
- Miles, Matthew B., and A. Michael Huberman. 1994. *Qualitative Data Analysis*. London: Sage Publications.
- MkNelly, Barbara, and Karen Lippold. 1998. "Practitioner-Led Impact Assessment: A Test in Mali." USAID AIMS Brief 21. Washington, D.C.: USAID
- MkNelly, Barbara, and Christopher Dunford (in collaboration with the Program in International Nutrition, University of California, Davis). 1998. "Impact of Credit with Education on Mothers' and Their Young Children's Nutrition: Lower Pra Rural Bank Credit with Education Program in Ghana."

- Freedom from Hunger Research Paper No. 4, March.
- Moffitt, Robert. 1991. "Program Evaluation with Nonexperimental Data." *Evaluation Review* 15 (3): 291–314.
- Mohr, Lawrence B. 1995. *Impact Analysis for Program Evaluation*, 2nd ed. Thousand Oaks, Calif.: Sage Publications.
- \_\_\_\_\_. 1999. "The Qualitative Method of Impact Analysis." Processed.
- Morduch, Jonathan. 1998a. "The Microfinance Schism." Harvard Institute for International Development, Development Discussion Paper 626, February.
- \_\_\_\_\_. 1998b. "Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh." Harvard University Institute for International Development, Processed, June.
- \_\_\_\_\_. 1999. "The Microfinance Promise." *Journal of Economic Literature* 37(4): 1569–1614.
- Newman, John, Menno Pradhan, Laura Rawlings, Ramiro Coa, and Jose Luis Evia. 2000. "An Impact Evaluation of Education, Health, and Water Supply Investments of the Bolivia Social Investment Fund." Processed.
- Newman, John, Laura Rawlings, and Paul Gertler. 1994. "Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries." *The World Bank Research Observer* 9 (July): 181–202.
- Operations Evaluation Department, World Bank. 1994. "An Overview of Monitoring and Evaluation in the World Bank." June 30.
- Poppele, J. S. Summarto, and L. Pritchett. 1999. "Social Impacts of the Indonesia Crisis: New Data and Policy Implications." Social Monitoring Early Response Unit, World Bank, Washington, D.C. Processed.
- Pradhan, Menno, Laura Rawlings, and Geert Ridder. 1998. "The

- Bolivian Social Investment Fund: An Analysis of Baseline Data for Impact Evaluation." *World Bank Economic Review* 12 (3): 457–82.
- Ravallion, Martin. 1994. *Poverty Comparisons, Fundamentals in Pure and Applied Economics*, vol. 56. Harwood Academic Publishers.
- \_\_\_\_\_. 1999. *Monitoring Targeting Performance When Decentralized Allocations to the Poor Are Unobserved*. World Bank, Washington, D.C. Processed.
- Ravallion, Martin, and Quentin Wodon. 1998. *Evaluating a Targeted Social Program When Placement Is Decentralized*. World Bank Policy Research Working Paper 1945. Washington, D.C.: World Bank.
- Ravallion, Martin, Dominique van de Walle, and Madhur Gautam. 1995. "Testing a Social Safety Net." *Journal of Public Economics* 57: 175–99.
- Rawlings, Laura. 2000. "Assessing Educational Management and Quality in Nicaragua." In Bamberger: *Integrating Quantitative and Qualitative Methods in Development Research*. Washington, D.C.: World Bank.
- Rebien, Claus C. 1997. "Development Assistance Evaluation and the Foundations of Program Evaluation." *Evaluation Review* 21 (August): 438–60.
- Revenga, Ana, Michelle Riboud, and Hong Tan. 1994. "The Impact of Mexico's Retraining Program on Employment and Wages." *World Bank Economic Review* 8 (2): 247–77.
- Rivers, Douglas, and Quang H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39: 347–66.
- Robinson, Sherman. 1991. "Macroeconomics, Financial Variables, and Computable General Equilibrium Models." *World Development*, 19(11): 1509–25.

- Rosenbaum, P., and D. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70: 41–55.
- . 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score." *American Statistician* 39: 35–39.
- Rossman, Gretchen B., and Bruce L. Wilson. 1985. "Numbers and Words: Combining Quantitative and Qualitative Methods in a Single Large-Scale Evaluation Study. *Evaluation Review* 9 (3): 627–43.

---

---

## Phụ lục 1

# Các nghiên cứu tình huống

### **Phụ lục 1.1: Đánh giá lợi ích tới người nghèo từ các chương trình công việc: Chương trình TRABAJAR của Áchentina**

#### **I. Giới thiệu**

**Mô tả dự án.** Chương trình TRABAJAR của Áchentina có mục tiêu giảm đói nghèo bằng cách đồng thời tạo ra cơ hội việc làm cho người nghèo và cải thiện cơ sở hạ tầng xã hội tại các cộng đồng nghèo. TRABAJAR I là một chương trình thí điểm được triển khai vào năm 1996 nhằm đối phó với tình trạng khủng hoảng kinh tế lan tràn và tỷ lệ thất nghiệp vượt quá 17 phần trăm. TRABAJAR II được khởi đầu vào năm 1997 như là một chương trình mở rộng và cải cách từ chương trình thí điểm, và TRABAJAR III đã bắt đầu phê chuẩn các dự án vào năm 1998. Chương trình TRABAJAR đưa ra mức lương thấp nhằm hấp dẫn (“tự lựa chọn”) chỉ những người nghèo, những người lao động thất nghiệp tham gia chương trình. Chính quyền địa phương và các tổ chức phi chính phủ (NGO) đề xuất các dự án cơ sở hạ tầng thuê mướn người tham dự và phải trang trải những khoản chi phí không phải tiền lương trong dự án. Các dự án được chấp thuận ở cấp độ khu vực, căn cứ vào hướng dẫn của chính quyền trung ương.

Chương trình này đã có những thay đổi về thủ tục thiết kế và hoạt động nhờ quá trình đánh giá. TRABAJAR II bao gồm một số cải cách nhằm cải thiện định hướng mục tiêu của dự án. Hệ thống phân bổ ngân sách của chính quyền trung ương hiện

nay chịu tác động mạnh mẽ của mức độ đói nghèo và chỉ số thất nghiệp của các tỉnh, và áp dụng trọng số cao hơn cho các đề xuất dự án ở những vùng nghèo theo hướng dẫn chấp thuận dự án. Ở cấp độ địa phương, đã có những nỗ lực trong hai chương trình TRABAJAR II và III để nâng cao năng lực của các cơ quan cấp tỉnh nhằm giúp đỡ các khu vực nghèo tiến hành dự án và nâng cao các tiêu chuẩn của chất lượng cơ sở hạ tầng.

**Đánh giá tác động.** Nỗ lực đánh giá được bắt đầu trong quá trình chuẩn bị dự án cho TRABAJAR II và hiện vẫn đang tiếp tục. Mục tiêu của đánh giá là xác định xem chương trình có đạt được mục tiêu chính sách hay không, và chỉ ra những lĩnh vực cần có cải cách trong chương trình nhằm tối đa hóa hiệu quả của nó. Đánh giá bao gồm một số nghiên cứu riêng biệt nhằm xem xét (a) lợi ích thu nhập rộng của những người tham dự chương trình, (b) phân bổ nguồn lực của chương trình giữa các khu vực (tính định hướng), (c) chất lượng của các dự án cơ sở hạ tầng được tài trợ, và (d) vai trò của cộng đồng và các tổ chức phi chính phủ trong kết quả của dự án.

Có hai thành phần đánh giá nổi bật về những kỹ thuật thực nghiệm tốt nhất được sử dụng. Thứ nhất, nghiên cứu về lợi ích thu nhập rộng, minh họa các kỹ thuật tốt nhất trong so sánh tương hợp cũng như khả năng sử dụng phong phú đối với các dữ liệu điều tra hộ gia đình toàn quốc để thực hiện so sánh tương hợp. Thứ hai, nghiên cứu về các kết quả định hướng, đưa ra một kỹ thuật mới nhằm đánh giá sự định hướng khi không thể quan sát được phạm vi tác động của chi tiêu công cộng tại cấp địa phương. Toàn bộ thiết kế đánh giá cũng cho thấy sự kết hợp tốt nhất giữa các thành phần và các kỹ thuật nghiên cứu - từ phân tích định lượng cho đến các cuộc tham quan công trình tại chỗ và đánh giá xã hội - nhờ đó cung cấp kết quả đúng lúc.

## II. Thiết kế việc đánh giá

Đánh giá TRABAJAR bao gồm nhiều thành phần được thiết kế

để đánh giá mức độ thành công trong việc đạt được các mục tiêu chính sách của chương trình. Thành phần đầu tiên được rút ra từ dữ liệu điều tra hộ gia đình để đánh giá lợi ích thu nhập đối với những người tham dự TRABAJAR. Nghiên cứu này cải tiến so với sự đánh giá truyền thống về các chương trình công việc, theo đó lợi ích thu nhập của những người tham dự chương trình được xác định bằng tổng số thu nhập nhận được. Trái lại, nghiên cứu này đo lường lợi ích thu nhập bằng cách ước tính lợi ích thu nhập ròng. Sử dụng các thành tựu mới đây trong kỹ thuật so sánh tương hợp, nghiên cứu này tính đến thu nhập bỏ lỡ (thu nhập phải từ bỏ do các thành viên tham dự chương trình TRABAJAR), do đó có kết quả chính xác, và lợi ích thu nhập ròng của người tham dự thấp hơn so với cách tính truyền thống. Thành phần đánh giá thứ hai theo dõi sự phân bổ nguồn vốn của chương trình (tính định hướng), bám sát những thay đổi theo thời gian do kết quả của cải cách. Thông qua việc sử dụng khôn khéo các dữ liệu phổ biến sẵn có (sự phân bổ tài trợ của chương trình giữa các tỉnh và điều tra dân số toàn quốc), thiết kế của thành phần đánh giá này thể hiện một phương pháp mới trong đánh giá tính định hướng tới đối nghèo khi không có dữ liệu thực sự về diện đối nghèo của chương trình. Phân tích này được bắt đầu bằng công việc giám sát (vào tháng 11-1997) và được cập nhật mỗi năm hai lần kể từ thời điểm này.

Các thành phần đánh giá bổ sung bao gồm một phân tích về chi phí - lợi ích được tiến hành trên một mẫu nhỏ gồm những dự án cơ sở hạ tầng, cùng với các đánh giá xã hội để cung cấp phản hồi về sự thực hiện dự án. Mỗi hoạt động trong số này đều được tiến hành hai lần, cho cả TRABAJAR II và TRABAJAR III. Có ba hoạt động đánh giá tương lai được lên kế hoạch. Kỹ thuật nghiên cứu so sánh tương hợp sẽ được áp dụng lại để đánh giá tác động của sự tham dự chương trình TRABAJAR tới hoạt động của thị trường lao động. Chất lượng các dự án cơ sở hạ tầng sẽ được đánh giá lại, lần này đối với các dự án đã hoàn thành trong ít nhất một năm để đánh giá mức độ bền vững, tình

hình bảo trì và tỷ lệ sử dụng. Cuối cùng, một thành phần nghiên cứu định tính sẽ điều tra về hoạt động và thủ tục của chương trình bằng cách phỏng vấn cán bộ của những cơ quan bảo trợ dự án, cũng như những người thụ hưởng dự án.

### III. Thu thập dữ liệu và kỹ thuật phân tích

Đánh giá về lợi ích thu nhập ròng tới những người tham dự chương trình được căn cứ vào hai nguồn dữ liệu, điều tra mức sống dân cư toàn quốc (Encuesta de Desarrollo Social- EDS) và điều tra về những người tham dự TRABAJAR được tiến hành nhằm phục vụ mục đích đánh giá này. (Điều tra EDS được tài trợ trong một dự án khác của Ngân hàng Thế giới. Nó được thiết kế nhằm cải thiện chất lượng thông tin về phúc lợi của hộ gia đình ở Argentina, đặc biệt là trong lĩnh vực tiếp cận các dịch vụ hạ tầng và các chương trình xã hội của chính phủ.) Các đợt điều tra được thực hiện trong tháng Tám (EDS) và tháng Chín (điều tra người tham dự TRABAJAR) năm 1997 do cơ quan thống kê toàn quốc, sử dụng chung bảng câu hỏi và nhóm phỏng vấn. Mẫu cho điều tra EDS bao gồm 85 phần trăm dân số toàn quốc, trừ một số khu vực nông thôn và các cộng đồng rất nhỏ. Mẫu cho điều tra người tham dự TRABAJAR được rút ra từ một mẫu ngẫu nhiên các dự án TRABAJAR II nằm trong khung mẫu EDS và tạo dữ liệu cho 2.802 thành viên tham gia chương trình hiện tại (tổng số người tham gia TRABAJAR II từ giữa tháng Năm 1997 và tháng Một năm 1998 là 65.321 người). Tính tin cậy của kỹ thuật tương hợp được tăng cường nhờ khả năng sử dụng cùng một bảng câu hỏi cho cả nhóm tham dự và nhóm đối chứng, vào cùng thời gian, và đảm bảo cả hai nhóm này có chung môi trường kinh tế.

Để tạo ra nhóm đối chứng phù hợp từ điều tra EDS, nghiên cứu này sử dụng một kỹ thuật có tên là mức độ xu hướng. (Việc bảng câu hỏi EDS rất phức tạp, thu thập dữ liệu chi tiết về các đặc điểm của hộ gia đình giúp dự đoán sự tham dự chương trình, đã tạo điều kiện thuận lợi cho việc sử dụng kỹ thuật mức

độ xu hướng). Một cặp tương hợp lý tưởng sẽ gồm hai cá nhân, một người thuộc nhóm mẫu tham dự và người kia thuộc nhóm đối chứng, hai người này có tất cả các biến số ( $x$ ) dự đoán sự tham dự chương trình như nhau. Vấn đề thường gặp phải trong việc tạo ra một cặp như vậy là tính thiếu thực tế do có quá nhiều biến số trong  $x$ . Tuy nhiên, vẫn có thể tạo ra các cặp tương hợp trên cơ sở tính toán mức độ xu hướng của mỗi cá nhân, hiểu đơn giản là xác suất tham dự phụ thuộc vào ( $x$ ). (Mức độ xu hướng được tính cho mỗi quan sát trong nhóm mẫu tham dự và nhóm đối chứng bằng cách sử dụng mô hình logit chuẩn tắc). Dữ liệu về thu nhập của nhóm đối chứng phù hợp gồm những người không tham dự- cho phép ước lượng được phần thu nhập mất đi của những người đã tham dự chương trình TRABAJAR II. Thu nhập ròng có được từ sự tham gia chương trình được tính bằng tổng tiền lương của chương trình trừ đi phần thu nhập mất đi.

Phân tích định hướng đáng lưu ý ở chỗ, không cần có sự thu thập dữ liệu đặc biệt nào. Công trình thực nghiệm căn cứ vào dữ liệu từ văn phòng dự án của bộ về phân bổ nguồn tài chính theo các khu vực cho TRABAJAR I (từ 3/1996 đến 4/1997) và trong sáu tháng đầu của TRABAJAR II (từ tháng 5 đến tháng 10/1997). Nghiên cứu này cũng căn cứ vào chỉ số đói nghèo của mỗi khu hành chính (có 510 khu), được tính từ điều tra dân số năm 1991 như là tỷ lệ của các hộ gia đình có “nhu cầu cơ bản chưa được đáp ứng.” Đây là một chỉ số tổng hợp đại diện cho mật độ cư trú, trang bị vệ sinh, chất lượng nhà ở, trình độ học vấn của người lớn, tỷ lệ nhập học của trẻ em, việc làm và sự phụ thuộc (tỷ lệ thành viên gia đình làm việc chia cho những người không làm việc). Chỉ số này tương đối cũ, mặc dù nó có ưu điểm là biến ngoại sinh đối với (không chịu tác động bởi) chương trình TRABAJAR. Để phân tích phạm vi tác động dự định, dữ liệu về chi tiêu công cộng theo khu vực địa lý- trong trường hợp này, là các khu hành chính- được hồi quy theo tỷ lệ đói nghèo của khu vực địa lý tương ứng. Hệ số kết quả ước tính một cách nhất quán “chênh lệch trong định hướng” được tính

bằng chênh lệch giữa phân bổ tài chính trung bình của chương trình cho người nghèo và cho người không nghèo. Chênh lệch trong định hướng trên quy mô toàn quốc có thể được tách ra để đánh giá phần đóng góp của cơ chế định hướng của chính phủ (phân bổ tài chính cho các khu hành chính) so với sự định hướng ở cấp tỉnh do chính quyền địa phương.

Phân tích chi phí - lợi ích do một kỹ sư xây dựng tiến hành, người này thực hiện một nghiên cứu hai giai đoạn về các dự án cơ sở hạ tầng của TRABAJAR. Trong giai đoạn thứ nhất, bà ta đến thăm một mẫu gồm 50 dự án TRABAJAR I đã hoàn thành và đánh giá chúng căn cứ vào sáu loại chỉ số: kỹ thuật, thể chế, môi trường, kinh tế xã hội, giám sát, và vận hành và bảo trì. Các dự án này sau đó được xếp hạng chất lượng chung theo một hệ thống điểm, và các phân tích chi phí - lợi ích sẽ được tiến hành khi phù hợp (không áp dụng đối với các trường học hay trung tâm y tế). Một nghiên cứu tiếp theo về 120 dự án TRABAJAR II được thực hiện một năm sau đó, nhằm theo dõi tác động của cải cách tới chất lượng cơ sở hạ tầng.

Các đánh giá về mặt xã hội được thực hiện trong giai đoạn chuẩn bị dự án đối với cả TRABAJAR I và TRABAJAR II. Các đánh giá này cung cấp sự phản hồi về các vấn đề thực hiện dự án, như vai trò của các tổ chức phi chính phủ, sự sẵn có của hỗ trợ kỹ thuật trong thực hiện dự án và xây dựng, và sự lựa chọn đối tượng thụ hưởng. Cả hai đánh giá xã hội đều do các nhà xã hội học tiến hành dưới hình thức các nhóm tập trung và các cuộc phỏng vấn.

## VI. Kết quả

Việc tính đến thu nhập bô lõ rất quan trọng nếu muốn có được một bức tranh chính xác về lợi ích của chương trình phúc lợi việc làm. Các số liệu thống kê mô tả những người tham dự TRABAJAR II cho thấy, nếu không có khả năng tiếp cận chương trình (thu nhập trên đầu người của gia đình trừ đi tiền lương của chương trình), thì 85 phần trăm những người tham dự

chương trình sẽ rơi vào nhóm 20 phần trăm số dân có thu nhập thấp nhất - và do đó, sẽ nằm trong số những người nghèo ở Argentina. Tuy nhiên, kết quả ước lượng về thu nhập bỏ lỡ theo phương pháp tương hợp khá lớn, do vậy, sự gia tăng thu nhập rộng trung bình nhờ tham dự chương trình chỉ bằng khoảng một nửa tiền lương của TRABAJAR. Khi không có chương trình, những người tham dự chương trình sẽ không có đủ tiền để có thể tiếp tục sống trong tình trạng thất nghiệp; do đó việc tham dự chương trình đã làm mất đi một số thu nhập nhất định. Thu nhập bị mất đi này được ước lượng bằng cách quan sát thu nhập của những người không tham dự chương trình có những đặc điểm "giống với" những người tham dự. Tuy nhiên, sau khi đã tính đến thu nhập bị mất, thì sự phân bổ lợi ích của chương trình vẫn hoàn toàn theo hướng có lợi cho người nghèo, với việc 80 phần trăm những người tham dự chương trình nằm trong nhóm 20 phần trăm dân số có thu nhập thấp nhất. Sự tham dự của phụ nữ vào chương trình ở mức thấp (chiếm 15 phần trăm) nhưng các lợi ích thu nhập rộng hầu như giống nhau đối với cả những người tham dự là nữ và nam, trong khi những người tham dự trẻ tuổi hơn có lợi ích thu nhập thấp hơn một cách đáng kể.

Tính định hướng đối tượng được cải thiện đáng kể nhờ các cải cách trong TRABAJAR II. Sự phân bổ nguồn lực ngầm cho các hộ gia đình nghèo đã tăng lên bảy lần giữa TRABAJAR I và TRABAJAR II. Một phần ba sự gia tăng này là do sự định hướng tốt hơn ở cấp trung ương, và hai phần ba là do định hướng tốt hơn ở cấp tỉnh. Tuy nhiên, cũng có những khác biệt đáng kể giữa các kết quả định hướng của các tỉnh. Một khu hành chính có 40 phần trăm số dân được xếp vào nhóm người nghèo có thể kỳ vọng nhận được số tiền phân bổ dao động từ không đến năm lần số tiền phân bổ trung bình, tùy thuộc vào việc nó thuộc tỉnh nào. Hơn nữa, kết quả định hướng có xu hướng kém hơn tại các tỉnh nghèo nhất.

Chất lượng của các dự án cơ sở hạ tầng được ghi nhận là đạt

yêu cầu, nhưng đáng thất vọng là các cải cách của TRABAJAR đã không mang lại những cải thiện đáng kể về chất lượng của các dự án này. Lý do một phần là do sự mở rộng mạnh mẽ của chương trình, điều này làm cho chương trình khó đảm bảo đáp ứng được một số tiêu chuẩn vận hành đã được xác định từ trước. Tuy nhiên, các dự án đã đáp ứng những nhu cầu ưu tiên của cộng đồng tốt hơn. Sự đánh giá xã hội cho thấy, cần có sự hỗ trợ kỹ thuật tốt hơn đối với các tổ chức phi chính phủ và các vùng nông thôn, cũng như cần có sự công khai và minh bạch thông tin hơn nữa về chương trình TRABAJAR.

## V. Ứng dụng chính sách

Kết quả đánh giá đã cung cấp bằng chứng rõ ràng: những người tham dự TRABAJAR chủ yếu là những người nghèo. Việc tự lựa chọn người tham dự bằng cách đưa ra mức tiền lương thấp là một chiến lược đã có hiệu quả ở Argentina, và những người tham dự đã tăng được thu nhập nhờ tham dự chương trình (mặc dù lợi ích ròng nhỏ hơn tổng tiền lương, do còn có phần thu nhập bị mất). Chương trình này có lẽ không có sự phân biệt đối xử đối với phụ nữ. Những cải cách trong TRABAJAR II đã nâng cao hơn nữa thành công của các kết quả định hướng theo khu vực địa lý - chương trình giờ đây đã thành công hơn trong việc định hướng các luồng vốn tới các vùng nghèo. Tuy nhiên, các kết quả của chương trình vẫn rất khác nhau và vẫn tiếp tục có những kết quả yếu kém ở một số tỉnh. Những tỉnh này cần nhận được sự quan tâm chính sách hơn nữa. Cuối cùng, các kết quả đáng thất vọng về chất lượng dự án cơ sở hạ tầng đã tạo ra những nỗ lực to lớn của nhóm công tác dự án để cải thiện kết quả trong lĩnh vực này bằng cách tăng cường các thủ tục vận hành như, yêu cầu phải đến thực địa nhiều hơn trong đánh giá và giám sát, trừngh phạt các cơ quan có kết quả yếu kém khi hoàn thành dự án, và tăng cường các tài liệu hướng dẫn đánh giá.

## VI. Chi phí và tổ chức quản lý công tác đánh giá

**Chi phí.** Chi phí thực hiện điều tra TRABAJAR (cho nghiên cứu về lợi ích thu nhập ròng) và chi phí xử lý dữ liệu xấp xỉ bằng 350.000 đô la. Hai đánh giá về chất lượng các dự án phụ (phân tích chi phí- lợi ích) có chi phí bằng khoảng 10.000 đôla cho mỗi đánh giá. Chi phí cho đánh giá xã hội cũng tương tự, do đó, tổng chi phí đánh giá ước tính bằng khoảng 390.000 đôla.

**Tổ chức quản lý.** Đánh giá được cán bộ Ngân hàng Thế giới Martin Ravallion thiết kế và do Ngân hàng Thế giới kết hợp với nhóm dự án của Áchentina thực hiện. Trong các giai đoạn khác nhau, đánh giá cũng đòi hỏi sự hợp tác của một số cơ quan chính quyền địa phương, bao gồm Cơ quan Thống kê, Bộ Lao động (kể cả các văn phòng thực địa), và Vụ Phân tích chính sách trong Bộ Phát triển xã hội.

## VII. Các bài học

Tầm quan trọng của việc tính đến thu nhập bỏ lỡ khi đánh giá lợi ích của chương trình phúc lợi việc làm. Thu nhập bỏ lỡ chiếm tỷ trọng khá lớn (bằng khoảng một nửa) trong tổng tiền lương mà những người tham dự chương trình phúc lợi việc làm ở Áchentina nhận được. Kết quả này cho thấy, các phương pháp đánh giá truyền thống (chỉ sử dụng tổng tiền lương) đã đánh giá cao hơn thực tế một cách đáng kể lợi ích về thu nhập của chương trình, và do đó, đã phỏng đại về tình cảnh những người tham dự nghèo sê ra sao khi không có chương trình này.

**Phương pháp tương hợp mức độ xu hướng.** Khi sử dụng kỹ thuật đánh giá so sánh tương hợp, mức độ xu hướng cho phép có những cặp so sánh có tính tin cậy cao giữa mẫu bao gồm đối tượng tham dự và mẫu bao gồm đối tượng không tham dự (nhóm đối chứng).

**Sử dụng khôn ngoan các nguồn dữ liệu quốc gia hiện có.**

Thông thường, các nguồn dữ liệu hiện có, như điều tra toàn quốc hay điều tra hộ gia đình, có thể cung cấp các đầu vào có giá trị cho công việc đánh giá. Sử dụng những nguồn dữ liệu hiện có làm giảm nhu cầu thu thập dữ liệu tốn kém chỉ để phục vụ mục đích đánh giá. Có thể sử dụng các kỹ thuật đánh giá sáng tạo để bù đắp cho những dữ liệu bị thiếu, như đánh giá về các kết quả định hướng theo khu vực địa lý của TRABAJAR cho thấy.

**Nhiều thành phần đánh giá khác nhau.** Thiết kế đánh giá TRABAJAR minh họa cách kết hợp có hiệu quả những công cụ và kỹ thuật đánh giá khác nhau. Phân tích dữ liệu điều tra, đi thực địa và đánh giá xã hội, đều được sử dụng để tạo ra những kết quả phong phú, cung cấp đầu vào có giá trị để tăng hiệu quả của dự án và chỉ ra những lĩnh vực cần cải cách.

**Sự đúng lúc của kết quả.** Nhiều thành phần đánh giá được thiết kế rõ ràng, căn cứ vào chu kỳ dự án, chọn thời điểm để tạo ra các kết quả trong giai đoạn chuẩn bị dự án, nhờ vậy có thể sử dụng một cách có hiệu quả các kết quả của đánh giá để cung cấp thông tin cho chính sách. Một số thành phần đánh giá hiện nay đang tạo ra dữ liệu có tính định kỳ trong tiến trình theo dõi liên tục dự án.

### VIII. Nguồn

Jalan, Jyotsna, and Martin Ravallion. 1999. "Income Gains from Workfare and Their Distribution." World Bank, Washington, D.C. Processed.

Ravallion, Martin. 1999. Monitoring Targeting Performance When Decentralized Allocations to the Poor Are Unobserved." World Bank, Washington, D.C. Processed.

**Phụ lục 1.2: Liệu tài chính vi mô có thực sự  
giúp đỡ người nghèo không? Các bằng chứng  
mới từ những chương trình tiên phong  
ở Băngladét**

### I. Giới thiệu

**Mô tả dự án.** Các chương trình tài chính vi mô của Ngân hàng Grameen, Uỷ ban tiến bộ nông thôn Băngladét và Ban Phát triển nông thôn Băngladét là những chương trình tiên phong cho các chương trình tương tự ở nhiều nước khác. Các chương trình này cung cấp những khoản cho vay nhỏ đối với các hộ gia đình nghèo sở hữu dưới nửa hécta đất. Các khoản vay này được kèm theo các hợp đồng và lịch trình vay có tính sáng tạo. Các chương trình này đã phục vụ cho hơn 4 triệu người nghèo ở Băngladét, và có lẽ đã khá thành công. Chẳng hạn, nhóm phần tư vay tiền nhiều nhất trong số những người vay tiền từ Ngân hàng Grameen tiêu dùng nhiều hơn 15 phần trăm, có tỷ lệ trẻ em trai đi học nhiều gấp đôi và tỷ lệ trẻ em gái đi học tăng lên đáng kể so với nhóm phần tư vay tiền ít nhất.

**Các đặc điểm nổi bật của đánh giá.** Đánh giá này xem xét tác động của chương trình tới 1.800 hộ gia đình ở Băngladét và so sánh với một nhóm các hộ gia đình đối chứng sống tại những khu vực không nhận được tín dụng vi mô. Đóng góp chủ yếu của nghiên cứu này cho thấy rằng, các ước lượng đơn giản về tác động của chương trình có thể đã bị phóng đại đáng kể và việc sửa chữa sai lệch do lựa chọn mẫu đã làm cho những lợi ích tưởng là rất đáng kể trở về con số không. Đánh giá này chỉ ra rằng, hầu hết những lợi ích nhận thấy ban đầu là do các đặc điểm khác biệt của những người nhận được khoản vay: những người này có xu hướng khá giả hơn và làm việc nhiều hơn so với nhóm đối chứng. Một khi các kỹ thuật phù hợp được sử

dụng, thì việc vay tiền không tạo ra tác động gì tới tiêu dùng, và trẻ em ở các khu vực có chương trình tín dụng vi mô thực tế còn có kết quả kém hơn trẻ em ở các khu vực so sánh. Nhân tố quyết định chính tới hiện tượng này là việc cho vay của chương trình đã không tuân thủ các chỉ dẫn về đối tượng tham dự- trên thực tế, nhiều người vay tiền có diện tích đất đai sở hữu lớn hơn mức tối đa quy định là một nửa hécta.

Đánh giá này sử dụng một kỹ thuật điều tra và các kỹ thuật kinh tế lượng một cách sáng tạo. Một khía cạnh lý thú khác là đánh giá này xem xét cả tác động của tác động chương trình tới phương sai cũng như giá trị trung bình của kết quả, và phát hiện ra rằng, lợi ích chủ yếu của chương trình là làm giảm rủi ro chứ không phải là làm tăng các kết quả trung bình.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Các nhà nghiên cứu quan tâm tới việc xác định tác động của các chương trình tài chính vi mô tới log tiêu dùng trên đầu người, phương sai của log tiêu dùng, log mức lao động trung bình cho một người trưởng thành trong tháng trước, phương sai của log mức lao động cho một người trưởng thành, số giờ lao động của nam giới trong tháng trước, số giờ lao động của nữ giới trong tháng trước, tỷ lệ phần trăm nhập học của nam giới (tuổi từ 5 đến 17) và tỷ lệ phần trăm nhập học của nữ giới (tuổi từ 5 đến 17).

Đánh giá này căn cứ vào điều tra và bao gồm 87 làng được điều tra trong ba đợt, từ năm 1991 đến 1992. Các làng được lựa chọn một cách ngẫu nhiên từ một cuộc điều tra dân số và từ các danh sách hành chính, tại năm huyện ly đóng vai trò là nhóm đối chứng và 24 huyện ly là nơi thực hiện chương trình. Mỗi làng có 20 hộ gia đình được điều tra.

Các nhà nghiên cứu chia các hộ gia đình thành năm loại khác nhau, tuỳ thuộc vào tiêu chí hợp lệ là sở hữu một nửa hécta đất. Việc tái tạo lại lược đồ của chương trình rất có ích, nó minh

hoặc cách thức tạo ra các biến giả đặc trưng cho hệ thống và cho biết về sai lệch xảy ra do lựa chọn.

Làng 1: Có chương trình		Làng 2: Đối chiếu
A Không hợp lệ [b=1;e=0;c=0]	Hộ gia đình có nhiều hơn 1/2 hécta đất	B Sẽ không hợp lệ [b=0;e=0;c=0]
C Hợp lệ nhưng không tham dự [b=1;e=1; c=0]	D Người tham dự [b=1;e=1; c=1]	E Hộ gia đình có từ 1/2 hécta đất trở xuống

Việc so sánh kết quả của nhóm D với nhóm C gặp rất nhiều khó khăn về lựa chọn: bằng chứng cho thấy các hộ gia đình thuộc nhóm C không tham dự chương trình vì họ lo ngại rằng sẽ không thể hoàn trả được tiền vay. Tuy nhiên, nếu việc sở hữu đất đai là biến ngoại sinh, thì nhóm C và D có thể được so sánh với nhóm E, bởi vì sự khác biệt trong kết quả phụ thuộc vào sự tham dự chương trình chứ không phải do sự tự lựa chọn. Tất nhiên, điều này sẽ không còn đúng nếu như có nhiều khác biệt giữa các làng với nhau. Nếu có sự khác nhau (có thể do sự lựa chọn không ngẫu nhiên) thì phương pháp khác biệt trong khác biệt sẽ là phương pháp tốt hơn. Vì vậy, người đánh giá có thể tính toán các kết quả trung bình của C và D, các kết quả trung bình của A, và sau đó tính toán chênh lệch kết quả. Cũng tương tự, có thể tính chênh lệch giữa kết quả trung bình của E và của B, và sau đó có thể so sánh sự khác nhau trong phạm vi một làng.

### III. Dữ liệu

Nhà nghiên cứu thu thập dữ liệu về 1.798 hộ gia đình, trong đó có 1.538 hộ gia đình đủ tư cách tham dự và 905 hộ thực sự tham

dự chương trình. Các đợt điều tra thu thập dữ liệu được tiến hành trong năm 1991 và 1992, sau ba vụ thu hoạch lúa chính. Các biến quan trọng cần quan tâm là tiêu dùng trên đầu người trong tuần trước đó, mức tín dụng nhận được, diện tích đất sở hữu, cung lao động trong tháng trước và các đặc điểm nhân khẩu học. Một nguồn dữ liệu thứ cấp về các giao dịch đất đai cũng được sử dụng để kiểm tra hoạt động thị trường về đất đai.

#### **IV. Các kỹ thuật kinh tế lượng**

Có ba thành phần lý thú trong những kỹ thuật kinh tế lượng được sử dụng. Thành phần thứ nhất là việc sử dụng dữ liệu hành chính để kiểm tra các giả định quan trọng cần thiết cho sử dụng mô hình hồi quy không liên tục; tức là kiểm tra tính ngoại sinh của sở hữu đất đai. Thành phần thứ hai là việc sử dụng các kỹ thuật vẽ đồ thị không có tham số để mô tả xác suất một hộ gia đình đủ tư cách tham dự chương trình và xác suất nhận được khoản vay dưới dạng là một hàm số của biến về sở hữu đất đai. Công việc này đi kèm với thảo luận về việc liệu sử dụng mô hình hồi quy không liên tục có phù hợp không, do phân tích đồ thị cho thấy không có điểm chặn rõ ràng ở mức 0,5 hécta. Cuối cùng, nghiên cứu chủ yếu sử dụng các kỹ thuật khác biệt và khác biệt trong khác biệt.

#### **V. Ai là người tiến hành**

Dữ liệu do Viện Nghiên cứu Phát triển Băngladét thay mặt cho Ngân hàng Thế giới thu thập. Việc phân tích do nhà nghiên cứu Jonathan Morduch thực hiện.

#### **VI. Kết quả**

Các kết quả cho thấy hầu như toàn bộ các lợi ích hình dung của chương trình là kết quả của sai lệch do lựa chọn, có nguyên

nhân từ sự định hướng sai lầm các khoản vay. Cụ thể, các tác giả nhận thấy có khoảng 20 đến 30 phần trăm những người đi vay sở hữu nhiều hơn yêu cầu tối đa là nửa hécta đất, điều này cho thấy những người thực hiện chương trình có lẽ đã bẻ cong các quy định theo những cách thức không thể quan sát được. Khi sự so sánh được giới hạn chỉ trong những người vay tiền thoả mãn điều kiện hạn chế về đất đai, các tác giả nhận thấy tiêu dùng trung bình trong các làng tiếp cận được chương trình tài chính vi mô, thấp hơn tại các làng trong nhóm đối chứng khi áp dụng phương pháp khác biệt hay khác biệt kép. Điều này cho thấy có sự sai đích đáng kể khi cho vay, và vì vậy, không thể sử dụng các phương pháp hồi quy không liên tục để phân tích tác động của chương trình.

Đánh giá này cũng có ích trong việc so sánh các kết quả khi áp dụng những kỹ thuật kinh tế lượng khác nhau: các kết quả khác nhau đáng kể khi sử dụng các phương pháp tác động cố định, khác biệt trong khác biệt hay là khác biệt đơn giản. Đánh giá này cho thấy phương pháp đầu tiên không hợp lý khi sự khác biệt không thể quan sát của nhóm mục tiêu được sử dụng làm căn cứ để đưa ra quyết định về địa điểm. Tuy nhiên, có các kết quả trái ngược trong hai phương pháp còn lại về việc liệu chương trình có làm giảm sự biến động về tiêu dùng và thu nhập không, điều này cho thấy, cần phải có các dữ liệu theo chiều dài thời gian. Tác động tới giáo dục trên thực tế đã đảo ngược sau khi khắc phục sai lệch lựa chọn.

Cũng cần chú ý là, mặc dù phân tích này cho thấy có rất ít tác động của nhóm tham dự so với nhóm đối chứng, nhưng trên thực tế, nhóm đối chứng có thể không thiếu khả năng tiếp cận nguồn vốn bởi vì nguồn vốn này do các tổ chức phi chính phủ cung cấp. Tuy nhiên, việc chi tiêu hàng triệu đôla để tài trợ cho các chương trình tín dụng vi mô là một vấn đề cần phải cân nhắc.

## VII. Các bài học

Có một số bài học rất quan trọng từ nghiên cứu này. Thứ nhất

là, tầm quan trọng của việc kiểm tra xem chương trình có vận hàng đúng như mô tả không. Thứ hai là, xem xét về sự phù hợp của phương pháp hồi quy rời rạc so với phương pháp khác biệt trong khác biệt hay khác biệt đơn giản. Thứ ba là xem xét tác động của sự tham dự chương trình tới mô - men thứ nhất và thứ hai của phân phối, bởi vì bản thân giảm rủi ro cũng là một kết quả có ích. Còn có một bài học cơ bản tuy không được nêu ra một cách trực tiếp nhưng cũng có được từ nghiên cứu này, đó là bài học về kinh tế chính trị: nếu có một động cơ mạnh mẽ để bẻ cong các quy tắc, thì chắc chắn những quy tắc này sẽ bị bẻ cong.

### VIII. Nguồn

Morduch, Jonathan. 1998. "Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh." Processed, June 17.

Cũng tham khảo:

Khandker, Shahidur R. 1998. *Fighting Poverty with Microcredit: Experience in Bangladesh*. New York: Oxford University Press for the World Bank.

## **Phụ lục 1.3. Chương trình đổi thực phẩm lấy giáo dục ở Băngladέt: Đánh giá một chương trình xã hội có định hướng khi có sự phân quyền trong tổ chức chương trình**

### **I. Giới thiệu**

**Mô tả dự án.** Chương trình Đổi thực phẩm lấy giáo dục (FFE) ở Băngladέt được thiết kế nhằm tăng tỷ lệ nhập học tiểu học bằng cách cung cấp lúa gạo hay lúa mỳ cho các hộ gia đình được lựa chọn để tạo ra động cơ cho các cha mẹ học sinh. Chương trình này bắt đầu như là một chương trình thí điểm, nhưng đã gia tăng quy mô và tầm quan trọng: tỷ lệ ngân sách của Ban giáo dục tiểu học và đại chúng đã tăng từ 11 phần trăm năm 1993-94 lên 26 phần trăm năm 1995-96, và đạt con số 2,2 triệu trẻ em, hay bằng 13 phần trăm tổng số học sinh. Thiết kế của chương trình khá thú vị: chương trình được định hướng theo thứ bậc trong đó FFE được triển khai tại tất cả các trường học trong các khu vực địa lý lạc hậu về kinh tế, có tỷ lệ nhập học thấp. Sau đó sẽ lựa chọn các hộ gia đình được nhận thực phẩm từ các nhóm cộng đồng trong khu vực địa lý, dựa trên những tiêu chí có phần nào tuỳ ý (các hộ gia đình không có đất đai, do phụ nữ làm chủ hộ, có thu nhập thấp). Trẻ em trong các hộ gia đình này phải tham dự ít nhất 85 phần trăm các lớp học trong mỗi tháng.

**Các đặc điểm nổi bật của đánh giá.** Đánh giá này đặc biệt có ích bởi vì nó nhấn mạnh những gì có thể làm được khi thiết kế chương trình không tuân thủ các kỹ thuật đánh giá chuẩn tắc và khi cần phải tiến hành đánh giá trên cơ sở những nguồn dữ liệu sẵn có. Trên thực tế, cách tiếp cận trong FFE gần như là một cực đối ngược với cách lựa chọn đối tượng hoàn toàn ngẫu nhiên. Chương trình này không chỉ lựa chọn các khu vực địa lý

theo những đặc điểm nhất định, mà còn lựa chọn các cá nhân sinh sống trong các khu vực địa lý này tham dự chương trình, bởi vì những người này cần đến sự giúp đỡ. Như vậy, do chương trình nhằm mục tiêu là những người nghèo nhất trong số những người nghèo, nên việc phân tích đơn giản sẽ không cho thấy được mức độ tác động thực sự.

Thiết kế của chương trình tạo ra một khó khăn lớn trong việc tạo ra tình huống phản chứng (counterfactual) bởi vì sự lựa chọn tham gia chương trình do nhu cầu tham dự chương trình của hộ gia đình quyết định. Đánh giá này cung cấp một phương pháp sáng tạo – và có thể khái quát hoá – để khắc phục sai lệch do sự lựa chọn bằng cách dựa vào sự phân quyền trong quá trình ra quyết định. Tóm lại, bởi vì chính quyền trung ương phân bổ chi tiêu cho các khu vực địa lý, nhưng các cơ quan địa phương lại thực hiện phân bổ trong khu vực nên đánh giá sử dụng các kỹ thuật biến công cụ dựa trên tiêu chí địa lý để làm giảm sự sai lệch xảy ra trong thủ tục lựa chọn nội sinh. Việc áp dụng phương pháp này dẫn đến có thể ước lượng được các tác động của FFE cao hơn nhiều so với khi sử dụng phương pháp bình quân nhỏ nhất thông thường.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Câu hỏi nghiên cứu là định lượng hoá tác động của FFE tới việc đi học, được đo lường bằng tỷ lệ nhập học của mỗi hộ gia đình. Thiết kế đánh giá rất đơn giản: đánh giá được tiến hành dựa trên các dữ liệu có từ trước- cụ thể, sử dụng cả một kỳ điều tra chi tiêu hộ gia đình có tính đại diện trên quy mô toàn quốc và một kỳ điều tra cộng đồng chi tiết. Đánh giá hồi tưởng (retrospective) được thiết kế trên thực tế nhằm làm giảm nhu cầu điều tra gốc; đánh giá này chỉ cần đến các điều tra bao gồm đặc điểm của hộ gia đình và đặc điểm địa lý cụ thể của khu vực mà hộ gia đình sống. Phần tiếp theo cung cấp chi tiết về cách thức cấu trúc công việc điều tra để phản ánh chính xác tác động của chương trình.

### III. Dữ liệu

Dữ liệu được rút ra từ Điều tra chi tiêu hộ gia đình năm 1995-96 (HES), một kỳ điều tra có tính đại diện trên toàn quốc do Cơ quan Thống kê Bangladesh tiến hành, trong đó có các câu hỏi về việc tham dự FFE và có một thành phần điều tra cấp địa phương. Các tác giả sử dụng những câu trả lời về đặc điểm nhân khẩu học của hộ gia đình, sở hữu đất đai, trường học, và các biến của chương trình từ 3.625 hộ gia đình ở nông thôn nhằm xác định tác động tới việc đi học. Việc đi học của mỗi trẻ em thực tế đã được đo lường trực tiếp trong HES, tính cả những ngày chúng nghỉ học và những ngày trường đóng cửa. Biến phụ thuộc là tỷ lệ số ngày đi học trung bình của hộ gia đình chia cho số ngày đi học có thể thực hiện được. Cả hai thành phần này của điều tra đều đóng vai trò quan trọng. Một mặt, thông tin về hộ gia đình giúp vào việc nắm bắt được ảnh hưởng của các đặc điểm nhân khẩu học tới việc đi học. Mặt khác, thông tin về các đặc điểm của khu vực địa lý giúp cho việc mô hình hoá chiến lược ra quyết định của chính quyền tập trung và làm giảm sai lệch do lựa chọn được nêu ra ở trên.

### IV. Các kỹ thuật kinh tế lượng

Đánh giá trả lời hai vấn đề rất quan trọng mà các nhà nghiên cứu thực địa phải đối mặt. Vấn đề thứ nhất là, sự tham dự chương trình được phân quyền, và do đó, quyết định phân bổ nguồn lực dựa trên điều kiện các biến số mà nhà kinh tế lượng không thể quan sát được nhưng những người ra quyết định lại có thể quan sát được. Điều này có nghĩa là, đánh giá đòi hỏi có một thước đo quyết định sự tham dự chương trình ở cấp cá nhân nhưng không tương quan với hạng tử sai số (và do vậy, với kết quả chương trình). Vấn đề thứ hai là, chỉ có một điều tra chéo duy nhất, và không có điều tra gốc về những người tham dự chương trình, do vậy, khó có thể ước lượng được tác động thuần tuý của chương trình.

Dánh giá này đặc biệt sáng tạo do nó sử dụng một quy trình phân bổ gồm hai bước. Đặc điểm quan trọng cần thiết để thực hiện việc này là dữ liệu chéo, bao gồm cả đặc điểm của hộ gia đình và đặc điểm địa lý. Trong trường hợp đặc biệt này, mô hình là như sau:

$$W_i = \alpha IP_i + \beta X_i + \eta Z_i + \mu i \quad (1)$$

Trong đó  $W$  là kết quả phúc lợi của mỗi cá nhân,  $X$  và  $Z$  bao gồm các đặc điểm của gia đình và các đặc điểm về địa lý, và  $IP$  chính là sự tham dự của cá nhân vào trong chương trình, có tương quan với hạng tử sai số. Rõ ràng, và điều này rất quan trọng trong các tài liệu về đánh giá, ước lượng bình phương nhỏ nhất của  $\forall$  sẽ bị sai lệch.

Dánh giá sử dụng sự khác nhau về địa lý như là biến công cụ đối với sự tham dự của mỗi cá nhân, bởi vì nó không tương quan với hạng tử sai số, cũng như với đặc điểm của hộ gia đình. Có thể mô tả mối quan hệ này như sau

$$IP_i = \gamma GP_i + \pi X_i + v_i \quad (2)$$

Điều chú ý ở đây là  $Z$  cần bao gồm tất cả các thông tin được sử dụng trong việc đưa ra quyết định lựa chọn về địa lý. Trong trường hợp này, chúng ta sử dụng hai tập hợp biến số địa lý. Tập hợp biến số địa lý thứ nhất tương đối chuẩn tắc và tác động trực tiếp tới quyết định đi học, chẳng hạn như khoảng cách tới trường học, loại trường học, và các biến số về chất lượng trường học. Tập hợp thứ hai liên quan tới quyết định lựa chọn đối tượng, và mặc dù tập hợp này dài nhưng cũng cần lưu ý ở đây để minh họa. Các biến số bao gồm phân phối đất đai; mật độ tưới tiêu; chất lượng đường xá; mức độ điện khí hoá; khoảng cách và thời gian tới các trụ sở hành chính địa phương và tới thủ đô; khoảng cách tới các cơ sở y tế và tài chính; mức độ xảy ra thiên tai; trình độ học vấn trung bình của người chủ hộ và vợ/chồng; tôn giáo chính ở trong làng; và quy mô dân số của làng. Các biến số này được tính toán ở cấp độ làng và có thể dự đoán cho sự lựa chọn tương đối tốt: một phương trình

hồi quy dạng probit trên tổng số 166 làng đã cho các kết quả tương đối phù hợp (có  $R^2$  giả bằng 0,55). Kết quả này cho thấy các biến số trên thực tế đã phản ánh được sự lựa chọn thành phần tham dự nói chung.

Sau đó có thể mô hình hoá cho tập hợp các phương trình này bằng phương pháp bình phương nhỏ nhất ba giai đoạn và so sánh với các kết quả có được trong phương pháp hồi quy nhỏ nhất thông thường.

## V. Ai là người tiến hành

Đánh giá này do Martin Ravallion và Quentin Wodon của Ngân hàng Thế giới tiến hành như là một phần trong công việc hợp tác lâu dài giữa Cơ quan Thống kê Băngladét và Đơn vị Giảm đói nghèo và Quản lý Kinh tế của Khu vực Nam Á, Ngân hàng Thế giới.

## VI. Kết quả

Có sự khác biệt rõ ràng trong hai phương pháp: tác động ước tính của FFE sử dụng phương pháp bình phương nhỏ nhất ba giai đoạn cao hơn 66 phần trăm so với phương pháp bình phương nhỏ nhất thông thường khi không có sự kiểm soát về địa lý và cao hơn 49 phần trăm khi có sự kiểm soát này. Nói cách khác, ước lượng đơn giản chỉ kiểm soát sự khác nhau giữa các hộ gia đình (phương pháp bình phương nhỏ nhất thông thường có kiểm soát địa lý) sẽ làm giảm đi đáng kể hiệu quả của chương trình. Ngay cả khi đã đưa các công cụ kiểm soát địa lý cũng không làm mất đi sai lệch từ việc lựa chọn người tham dự. Lượng ngũ cốc trung bình trong chương trình làm tăng tỷ lệ tham dự lên 24 phần trăm, khi phương pháp nêu ở phần trên được sử dụng.

Đáng chú ý là, nhân tố quan trọng làm cho phương pháp này trở nên hợp lý là việc có đầy đủ biến số để mô hình hoá quyết định về đối tượng của chương trình, và những biến số này cũng

gần gũi với những biến số được những người quản lý chương trình sử dụng. Nếu vẫn còn các biến số bị bỏ sót, thì các kết quả vẫn tiếp tục bị sai lệch.

## VII. Các bài học

Nhiều đánh giá không bao gồm thiết kế chiến lược thu thập dữ liệu từ ban đầu, hoặc vì đánh giá này không phải là một phần không thể tách rời trong dự án ngay từ đầu, hoặc chỉ đơn giản vì lý do chi phí. Tình huống nghiên cứu này là một đánh giá cần nghiên cứu vì hai lý do. Thứ nhất, nó cho biết mức độ sai lệch xảy ra nếu như sử dụng phương pháp kinh tế lượng không đúng. Thứ hai, nó mô tả phương pháp kinh tế lượng hợp lý để ước lượng tác động của chương trình mà không có độ trễ về chi phí và thời gian liên quan tới một đánh giá trong tương lai.

## VIII. Nguồn

Ravallion and Wodon. 1998. *Evaluating a Targeted Social Program When Placement Is Decentralized*. Policy Research Working Paper 1945, World Bank, Washington, D.C.

## Phụ lục 1.4: Đánh giá Quỹ đầu tư xã hội của Bôlivia

### I. Giới thiệu

**Mô tả dự án.** Quỹ đầu tư xã hội Bôlivia (SIF) được thành lập năm 1991 như là một thể chế tài chính khuyến khích đầu tư bền vững vào các lĩnh vực xã hội, đáng chú ý là vào sức khoẻ, giáo dục và vệ sinh. Mục tiêu chính sách là định hướng đầu tư tới những khu vực không được các mạng lưới dịch vụ công cộng phục vụ, nhất là các cộng đồng dân cư nghèo. Nguồn vốn của SIF do đó được phân bổ tùy theo chỉ số đói nghèo của vùng, nhưng trong phạm vi của vùng thì chương trình này lại định hướng theo nhu cầu, nhằm đáp ứng nhu cầu của cộng đồng cho các dự án ở cấp khu vực. Hoạt động của SIF được phân quyền thêm vào năm 1994, nâng cao vai trò của các bộ chuyên trách và chính quyền khu vực trong thiết kế và chấp thuận dự án. SIF của Bôlivia là thể chế đầu tiên thuộc loại này trên thế giới và đã đóng vai trò nguyên mẫu cho nhiều quỹ tương tự bắt đầu hoạt động sau đây ở châu Mỹ Latinh, châu Phi và châu Á.

**Đánh giá tác động.** Mặc dù các quỹ xã hội được áp dụng rộng rãi trong thập kỷ 1990, nhưng vẫn có rất ít các nỗ lực mạnh mẽ để đánh giá tác động của các quỹ này tới thành tích giảm đói nghèo. Đánh giá SIF của Bôlivia, do Ngân hàng Thế giới và SIF đồng tiến hành, được bắt đầu từ năm 1991, và hiện vẫn đang tiếp tục. Nghiên cứu bao gồm các dữ liệu điều tra gốc (1993) và tiếp theo (1997), nhờ đó cho phép có đánh giá tác động trước và sau khi có chương trình. Đánh giá này bao gồm các đánh giá riêng biệt về các dự án giáo dục, y tế, và nước. Đánh giá này đặc biệt vì nó áp dụng nhiều phương pháp đánh giá khác nhau, và xem xét lợi ích cũng như nhược điểm của các phương pháp này. Các kết quả đánh giá ban đầu đã được hoàn tất và hiện

dang được trình bày cho các nhà tài trợ và các cơ quan chính phủ để nhận được phản hồi. Kết quả cuối cùng và các vấn đề phương pháp luận sẽ được khai thác kỹ lưỡng hơn khi kết hợp với Báo cáo Quỹ Đầu tư Xã hội năm 2000 và cùng với một phân tích về hiệu quả chi phí.

## II. Thiết kế đánh giá

Quy trình đánh giá SIF của Bôlivia được bắt đầu từ năm 1991 và hiện vẫn đang diễn ra. Thiết kế đánh giá bao gồm những đánh giá riêng biệt về các dự án giáo dục, sức khoẻ, và nước, trong đó đánh giá tính định hướng của chương trình tới người nghèo, cũng như tác động của đầu tư dịch vụ xã hội tới các kết quả đáng mong muốn cho cộng đồng như tăng tỷ lệ nhập học, điều kiện y tế, và mức độ sẵn có của nước. Đánh giá này minh họa cho các kỹ thuật đánh giá tốt nhất, sử dụng dữ liệu cơ sở trong phân tích tác động. Đánh giá cũng có tính sáng tạo do nó áp dụng hai phương pháp đánh giá khác nhau - lựa chọn ngẫu nhiên và so sánh tương hợp - trong phân tích các dự án giáo dục và sau đó so sánh các kết quả nhận được của mỗi phương pháp với nhau. Đây là một đóng góp quan trọng bởi vì lựa chọn ngẫu nhiên (lựa chọn ngẫu nhiên người thụ hưởng chương trình trong một nhóm đủ điều kiện) được nhìn nhận rộng rãi như là một phương pháp vững về mặt thống kê, nhưng so sánh tương hợp (sử dụng một quy trình không ngẫu nhiên nhằm lựa chọn một nhóm kiểm soát "phù hợp" nhất với các đặc điểm của đối tượng thụ hưởng chương trình) lại được sử dụng nhiều nhất trên thực tế.

## III. Các kỹ thuật thu thập và phân tích dữ liệu

Các nỗ lực thu thập dữ liệu cho đánh giá SIF của Bôlivia khá lớn và bao gồm cả một đợt điều tra ("cơ sở") về đầu tư trước

khi tiến hành SIF II năm 1993 và một đợt điều tra tiếp theo vào năm 1997. Các điều tra này được áp dụng cho cả các thể chế nhận được tài trợ từ SIF và các hộ gia đình cũng như các cộng đồng có lợi từ khoản đầu tư này. Các dữ liệu tương tự cũng được thu thập từ những thể chế và hộ gia đình so sánh (nhóm đối chứng). Điều tra hộ gia đình thu thập dữ liệu về nhiều đặc điểm, gồm có tiêu dùng, sự tiếp cận các dịch vụ cơ bản, tình trạng sức khoẻ và giáo dục của mỗi thành viên trong gia đình. Đánh giá bao gồm những mẫu riêng biệt cho các dự án sức khoẻ (gồm 4.155 hộ gia đình, 190 trung tâm y tế), các dự án giáo dục (1.894 hộ gia đình, 156 trường học), các dự án nước (1.071 hộ gia đình, 18 dự án nước), và các dự án nhà vệ sinh (231 hộ gia đình, 15 dự án).

Điều tra hộ gia đình bao gồm ba mẫu nhỏ: (a) một mẫu ngẫu nhiên, gồm tất cả các hộ gia đình sống ở nông thôn Bôlivia và khu vực Chaco (một tỉnh); (b) một mẫu gồm các hộ gia đình sống gần trường học trong nhóm tham dự hay nhóm đối chứng cho các dự án giáo dục; và (c) một mẫu gồm các hộ gia đình sẽ được hưởng lợi từ các dự án nước hay nhà vệ sinh.

Để phân tích liệu các khoản đầu tư của SIF có thực sự định hướng tới người nghèo hay không, nghiên cứu này sử dụng các dữ liệu gốc (đầu tư trước SIF) và thông tin về địa điểm mà sau đó thực sự tiến hành đầu tư SIF để tính toán xác suất các cá nhân này sẽ trở thành người thụ hưởng từ SIF với điều kiện về thu nhập của họ. Nghiên cứu này sau đó kết hợp các dữ liệu của điều tra gốc và điều tra tiếp theo để ước lượng tác động trung bình của SIF tại các cộng đồng nhận được đầu tư SIF, sử dụng các kỹ thuật hồi quy. Ngoài tác động trung bình, nghiên cứu còn tìm hiểu liệu có thể xác định được những đặc điểm của cộng đồng, trường học, hay các trung tâm y tế đi kèm với những tác động lớn hơn mức trung bình đáng kể hay không.

Trong lĩnh vực giáo dục, lĩnh vực mà đầu tư của SIF được phân bổ ngẫu nhiên cho những cộng đồng có tư cách tham dự tương tự nhau, nghiên cứu này áp dụng thiết kế thử nghiệm lứa

chọn ngẫu nhiên “lý tưởng” (trong đó có thể trực tiếp quan sát tình huống phản chứng). Trong các dự án sức khoẻ và vệ sinh, nơi các dự án không được phân bổ một cách ngẫu nhiên, nghiên cứu này sử dụng phương pháp “biến số công cụ” để bù đắp cho việc thiếu tình huống phản chứng trực tiếp. Các biến số công cụ có tương quan với chương trình, nhưng không có tương quan trực tiếp với kết quả của chương trình.

## VI. Kết quả

Đầu tư của SIF II trong lĩnh vực giáo dục và sức khoẻ không cải thiện một cách rõ ràng cơ sở hạ tầng và trang thiết bị. Các dự án giáo dục tác động rất ít tới tỷ lệ bỏ học, nhưng các kết quả thi của học sinh lớp sáu tại các nhà trường tham gia SIF lại cao hơn đáng kể. Trong lĩnh vực sức khoẻ, đầu tư của SIF làm tăng tỷ lệ sử dụng dịch vụ y tế và giảm tỷ lệ tử vong. Các dự án nước của SIF đi kèm với sự cải thiện không đáng kể chất lượng nước, nhưng cải thiện khả năng tiếp cận và khối lượng nước và cũng làm giảm tỷ lệ tử vong.

So sánh các kết quả sử dụng phương pháp lựa chọn ngẫu nhiên và phương pháp so sánh tương hợp trong giáo dục, chúng ta thấy phương pháp so sánh tương hợp tạo ra các nhóm tham dự và nhóm đối chứng ít tương đương nhau hơn, và vì vậy, dẫn đến các kết quả ít chắc chắn hơn khi đánh giá tác động của chương trình. Để minh họa kết quả này, có bằng chứng cho thấy, cơ sở hạ tầng trường học được cải thiện (chúng ta dự kiến điều này sẽ xảy ra tại các trường học tham gia SIF) trong thiết kế đánh giá lựa chọn ngẫu nhiên nhưng không được cải thiện trong thiết kế đánh giá so sánh tương hợp.

Cuối cùng, kết quả cho thấy, đầu tư của SIF II thường không định hướng tốt tới người nghèo. Các dự án sức khoẻ và vệ sinh có lợi cho các hộ gia đình tương đối khá giả tính theo thu nhập trên đầu người, và không có mối liên hệ nào giữa thu nhập trên đầu người và lợi ích giáo dục của SIF.

## V. Ứng dụng chính sách

Các kết quả về sự định hướng đối tượng của chương trình cho thấy sự xung đột giữa định hướng tới người nghèo và bản chất được dẫn dắt bởi nhu cầu của SIF. Cùng với việc triển khai luật tham dự đại chúng vào năm 1994, các dự án nhỏ phải được trình lên chính quyền khu vực. Các kết quả định hướng cho thấy, ngay cả trong một hệ thống phân quyền cao độ, thì cũng cần phải giám sát tiến trình định hướng. Trong trường hợp của Bolivia, có lẽ các cộng đồng khá giả và có tổ chức có nhiều khả năng nhận được đầu tư từ SIF hơn là các cộng đồng nghèo nhất. Đặc biệt, trong các dự án vệ sinh của SIF, khó mà khắc phục được sự thiên vị theo hướng bất lợi cho các cộng đồng nghèo. Đầu tư vào các trang bị vệ sinh cơ bản là phương thức đầu tư hiệu quả nhất ở các khu vực đông dân hiện đã có thể tiếp cận được hệ thống nước, nhờ đó, dự án có thể tận dụng được tính kinh tế theo quy mô.

Hiện trạng đầu tư của SIF không tạo ra tác động đáng kể tới tỷ lệ nhập học, đã dẫn đến sự tái cơ cấu hoạt động của SIF trong lĩnh vực này. Thay vì tập trung hoàn toàn vào việc cung ứng cơ sở hạ tầng, các dự án sẽ cung cấp kết hợp các đầu vào nhằm nâng cao chất lượng trường học. Tương tự, các kết quả đáng thất vọng về chất lượng nước (không hề được cải thiện nhờ các dự án của SIF so với trước kia) đã thu hút nhiều sự chú ý, và thiết kế dự án trong lĩnh vực này cũng đang được xem xét lại.

## VI. Các bài học

**Hiệu quả của kỹ thuật lựa chọn ngẫu nhiên.** Thiết kế nghiên cứu ngẫu nhiên, trong đó một nhóm đối chứng được lựa chọn một cách ngẫu nhiên từ những người có khả năng trở thành đối tượng thụ hưởng chương trình, có hiệu quả hơn nhiều trong việc phát hiện ra tác động của chương trình so với phương pháp so

sánh tương hợp để tạo ra một nhóm đối chứng. Việc lựa chọn ngẫu nhiên cần được đưa vào trong thiết kế chương trình ngay từ giai đoạn đầu nhằm xác định tiến trình lựa chọn người thụ hưởng chương trình. Sự lựa chọn ngẫu nhiên không phải luôn luôn có tính khả thi. Tuy nhiên, khi nguồn vốn của chương trình không đủ để trang trải cho tất cả các đối tượng thụ hưởng, thì có thể cần lựa chọn ngẫu nhiên từ một tập hợp lớn hơn các đối tượng đủ điều kiện tham gia chương trình.

**Tầm quan trọng của việc thể chế hóa tiến trình đánh giá.** Đánh giá có thể đặc biệt phức tạp và tốn thời gian. Đánh giá của Bôlivia được tiến hành trong thời kỳ bảy năm nhằm đánh giá đầy đủ tác động của dự án, và đã đạt được các kết quả tốt, xét trên phương diện này. Tuy nhiên, khó có thể quản lý được đánh giá trong một khoảng thời gian dài như vậy và khi có sự tham gia của nhiều tác nhân khác nhau (các cơ quan chính phủ và các thể chế tài trợ). Việc quản lý và thực hiện một nỗ lực đánh giá có thể được đơn giản hóa bằng cách kết hợp các tiến trình này trong nội dung hoạt động của các bộ tại địa phương ngay từ giai đoạn ban đầu. Hơn nữa, có thể chỉ cần đến những nỗ lực đánh giá phức tạp trong một số ít chương trình- chẳng hạn, chỉ trong các chương trình lớn có các kết quả thiếu chắc chắn. Tại các chương trình này, lợi ích của nỗ lực đánh giá có thể cũng là lớn nhất.

## VII. Các chi phí và tổ chức quản lý đánh giá

**Chi phí.** Tổng chi phí ước tính của đánh giá SIF ở Bôlivia tính đến thời điểm hiện tại là 878.000 đôla, tương đương với 0,5 phần trăm tổng chi phí dự án. Thu thập dữ liệu chiếm tỷ trọng chi phí tương đối cao trong chi phí này (69 phần trăm), phần còn lại là chi phí đi lại, phí tổn thời gian của cán bộ Ngân hàng Thế giới, và chi phí tư vấn.

**Tổ chức quản lý.** Đánh giá do cán bộ của Ngân hàng Thế giới

thiết kế và được đồng tài trợ của Ngân hàng Thế giới, KfW, và Chính phủ các nước Hà Lan, Thụy Điển, và Đan Mạch. Công việc điều tra do Viện Thống kê quốc gia Bôlivia tiến hành và do các đối tác của SIF quản lý trong giai đoạn một, và sau đó, trong giai đoạn hai do Bộ Tài chính.

## VII. Nguồn

Pradhan, Menno, Laura Rawlings, and Geert Ridder. 1998. "The Bolivian Social Investment Fund: An Analysis of Baseline Data for Impact Evaluation." *World Bank Economic Review* 12 (3): 457- 82.

thiết kế và được đồng tài trợ của Ngân hàng Thế giới, KfW, và Chính phủ các nước Hà Lan, Thụy Điển, và Đan Mạch. Công việc điều tra do Viện Thống kê quốc gia Bôlivia tiến hành và do các đối tác của SIF quản lý trong giai đoạn một, và sau đó, trong giai đoạn hai do Bộ Tài chính.

## VII. Nguồn

Pradhan, Menno, Laura Rawlings, and Geert Ridder. 1998. "The Bolivian Social Investment Fund: An Analysis of Baseline Data for Impact Evaluation." *World Bank Economic Review* 12 (3): 457– 82.

## Phụ lục 1.5: Tác động của các chương trình lao động tích cực: Cộng hoà Séc

### I. Giới thiệu

**Mô tả dự án.** Nhiều nước đang phát triển hiện phải đổi mới với khó khăn trong việc đào tạo lại người lao động khi giảm quy mô các doanh nghiệp nhà nước. Công việc này đặc biệt phức tạp ở các nền kinh tế đang chuyển đổi được đặc trưng bằng tỷ lệ thất nghiệp cao, và tình trạng trì trệ hay tiền lương giảm sút. Tuy nhiên, tất cả các chương trình đào tạo lại đều không tương đương nhau. Một số chương trình chỉ đơn giản là thanh toán chấm dứt hợp đồng trả hình cho những công nhân bị mất việc; các chương trình khác lại là các chương trình thất nghiệp trả hình. Điều này làm cho việc đánh giá các chương trình trở nên đặc biệt hấp dẫn.

Tuy vậy, việc đánh giá các chương trình đào tạo đặc biệt khó khăn, và trường hợp đánh giá ở Séc không phải là ngoại lệ. Về cơ bản, có một số chương trình khác nhau được thành lập để phục vụ những nhóm đối tượng khác nhau. Cũng có nhiều cách thức đo lường kết quả, kể cả số liệu về việc làm, công việc tự làm chủ, thu nhập hàng tháng, và thu nhập theo giờ. Mức độ tác động ở đây phụ thuộc nhiều vào thời gian hơn các loại đánh giá khác nhau và có thể đạt được các kết quả rất khác, tuỳ thuộc vào việc đánh giá xảy ra sau một tháng, sáu tháng, một năm, hay năm năm, kể từ khi có chương trình.

**Các đặc điểm nổi bật của đánh giá.** Đánh giá này lượng hoá tác động của bốn chương trình thị trường lao động tích cực (ALP) tại Cộng hoà Séc, sử dụng phương pháp thiết kế gần như thử nghiệm bằng cách lựa chọn một nhóm những người không tham dự tương tự, phù hợp nhất với những người tham dự ALP. Dữ liệu từ điều tra hành chính và dữ liệu điều tra tiếp đều

được sử dụng để đánh giá các kết quả khác nhau của chương trình, như thời gian thất nghiệp, khả năng có việc làm, công việc tự làm chủ, và thu nhập. Phân tích hồi quy được sử dụng để ước lượng tác động của mỗi chương trình trong số năm chương trình tới các kết quả này, sau khi đã kiểm soát đối với các đặc điểm nhân khẩu học cơ bản.

Có một số bài học quan trọng được rút ra từ đánh giá này. Một số bài học rất thực tế đó là: làm thế nào để thiết kế một đánh giá phức tạp, làm thế nào để sử dụng các dữ liệu hành chính, làm thế nào để giải quyết các khó khăn liên quan tới việc tiến hành điều tra, và cơ chế tạo ra mẫu tương hợp. Bài học thứ hai là, làm thế nào để cấu trúc một phân tích nhằm cung cấp những thông tin liên quan tới chính sách - có thể thông qua đánh giá chi tiết tác động theo nhóm nhỏ. Điều này dẫn đến khuyến nghị chính sách nhằm định hướng các chương trình ALP theo những loại khách hàng đặc biệt và đưa đến kết luận: có một dạng ALP không có hiệu quả trong thay đổi việc hay thu nhập.

## **II. Các câu hỏi nghiên cứu và thiết kế đánh giá**

Đánh giá này là một phần trong kế hoạch đánh giá có quy mô rộng hơn được thực hiện ở bốn nước: Cộng hoà Séc, Ba Lan, Hunggari, và Thổ Nhĩ Kỳ. Bởi cảnh chung là mỗi nước này đều có tỷ lệ thất nghiệp cao, một phần do giảm quy mô của các doanh nghiệp nhà nước. Để đối phó, các nước này đã triển khai các chương trình hỗ trợ thu nhập thụ động như trợ cấp thất nghiệp và trợ giúp xã hội. Bởi cảnh này kết hợp với các chương trình ALP là đối tượng của chương trình. Có năm chương trình ALP là: các chương trình Việc làm có ích cho xã hội (tạo việc làm mới), Việc làm có ích cho công cộng (việc làm công cộng ngắn hạn), Chương trình cho những người tốt nghiệp phổ thông (trợ cấp cho việc tuyển dụng những người mới tốt nghiệp), Đào tạo lại (đào tạo nghề nghiệp kéo dài từ vài tuần đến vài tháng), và Chương trình dành cho những người tàn tật và thiêt thời.

Chương trình cuối cùng tương đối nhỏ và không được đưa vào trong đánh giá.

Có hai câu hỏi nghiên cứu. Câu hỏi thứ nhất là kiểm định liệu các thành viên tham dự các ALP khác nhau có thành công hơn khi tái tham gia thị trường lao động so với những người không tham dự ALP không, và liệu kết quả này có khác nhau giữa các nhóm nhỏ và với các điều kiện thị trường lao động khác nhau không. Câu hỏi thứ hai là xác định hiệu quả chi phí của mỗi ALP và đưa ra đề xuất nhằm cải thiện hiệu quả.

Thiết kế đánh giá là thiết kế bán thử nghiệm, được đưa ra sau khi có chương trình- thực chất, chính là việc sử dụng một nhóm phù hợp. Nhóm tham dự được so sánh với một nhóm không tham dự tương tự (với thông tin được lấy ra từ sổ sách hành chính) gồm những người có đăng ký với cơ quan dịch vụ việc làm của nhà nước nhưng không được lựa chọn tham gia ALP. Điểm cơ bản ở đây là một cá nhân được lựa chọn ngẫu nhiên từ nhóm tham dự ALP. Kết quả của cá nhân này sau đó được so sánh với các cá nhân khác ở nhóm không tham dự (căn cứ vào tuổi tác, giới tính, trình độ giáo dục, số tháng thất nghiệp, quy mô thị trấn, tình trạng hôn nhân, và dạng công việc gần nhất). Đánh giá này đặc biệt đáng chú ý khi phân tích chi tiết nhóm đối chứng với nhóm tham dự.

Có một số vấn đề không tránh khỏi trong cách tiếp cận này, và chúng đã được nêu ra trong nhiều tài liệu (Burtless 1995, và Heckman và Smith 1995). Một mối lo ngại hiển nhiên mà các phương pháp không phải lựa chọn ngẫu nhiên thường gặp phải là việc những người tham dự đã bị “gạn lọc” bởi các chương trình đào tạo theo những đặc điểm mà nhà nghiên cứu không thể quan sát hay đo lường được. Mối lo ngại thứ hai là những người không tham dự có thể đã thay thế những hình thức đào tạo khác bằng việc tham gia chương trình đào tạo lại lao động của nhà nước. Mối lo ngại thứ ba là trợ cấp cho thuê mướn lao động có thể sẽ chỉ dẫn đến việc thay thế một số công nhân bằng một số người khác.

### III. Dữ liệu

Một thành phần rất lý thú của đánh giá này là việc sử dụng dữ liệu hành chính của chính phủ để tạo ra một khuôn khổ chọn mẫu cho điều tra. Nhóm điều tra đã tham quan Bộ Lao động và Các vấn đề xã hội (MOLSA) tại Praha và ba văn phòng thí trường lao động ở các địa phương nhằm tăng thêm sự hiểu biết về công việc quản lý và thực hiện ALP, và để có được các dữ liệu hành chính về những người tham dự ALP. Từ đó, người ta lựa chọn ra 20 huyện để tiến hành điều tra, dựa trên tiêu chí phân bố địa lý và các đặc điểm công nghiệp khác nhau. Giữa các huyện cũng có tỷ lệ thất nghiệp khác nhau. Điều tra bao gồm các câu hỏi định lượng về những kết quả quan trọng của chương trình và các câu hỏi định tính về sự đánh giá của người tham dự đối với chương trình.

Một thành phần có giá trị khác là việc thực hiện một đợt điều tra thí điểm ở bốn huyện. Phương pháp này - một phương pháp luôn đóng vai trò quan trọng- xác định được không chỉ những vấn đề về mặt kỹ thuật, mà còn cả một vấn đề về mặt pháp lý thường phát sinh khi sử dụng dữ liệu hành chính. Vấn đề ở đây chính là cách thức diễn giải luật về tính riêng tư: trong trường hợp này, MOLSA đã không cho phép người điều tra gửi thư trực tiếp và đòi hỏi phải có sự đồng ý của người được hỏi tiềm năng thì cơ quan lao động mới được cho những người điều tra biết địa chỉ của những người này. Yêu cầu này làm trì hoãn quá trình đánh giá, tăng chi phí, và hạ thấp tỷ lệ trả lời rất nhiều.

Điều tra được tiến hành vào đầu năm 1997 trên một mẫu ngẫu nhiên gồm 24.973 người đăng ký tại các văn phòng lao động. Trong số này, 9.477 người đã tham dự ALP trong giai đoạn 1994-95. Tỷ lệ trả lời từ phía những người không tham dự là 14 phần trăm; của người tham dự là 24,7 phần trăm, với tổng số người trả lời là 4.537 người. Tỷ lệ trả lời thấp liên quan trực tiếp tới phán quyết về mặt pháp luật: hầu hết mọi người đều không trả lời khi được yêu cầu lần đầu, nhưng trong số những

người đồng ý để người điều tra biết được địa chỉ thì tỷ lệ trả lời ở mức cao. Một điều tồi tệ hơn là chúng ta không biết gì về mức độ sai lệch do hiện tượng này gây ra.

#### IV. Kỹ thuật kinh tế lượng

Khó khăn trong đo lường bản chất thời vụ và sự phức tạp của các kết quả trên thị trường lao động được minh họa bằng việc sử dụng bảy thước đo kết quả khác nhau: tỷ lệ phần trăm những người hiện đang làm việc, tỷ lệ phần trăm những người hiện đang làm việc tự trả lương, tỷ lệ phần trăm những người từng có việc làm, độ dài thời gian thất nghiệp, độ dài thời gian nhận được trợ cấp thất nghiệp, tổng số tiền trợ cấp thất nghiệp, và thu nhập hàng tháng hiện tại.

Tuy nhiên, phương pháp đánh giá tương đối rõ ràng. Đánh giá sử dụng đồng thời phương pháp khác biệt đơn giản giữa các nhóm và phương pháp bình phương nhỏ nhất thường có các biến số giả cho từng nhóm để đánh giá tác động của chương trình. Người ta tính toán tác động chung của chương trình, sau đấy là tác động ước tính của mỗi nhóm phụ (theo tuổi tác, giới tính, trình độ giáo dục và quy mô doanh nghiệp đối với các kết quả thu nhập). Phân tích cuối cùng này đặc biệt có ích, bởi vì nó xác định ra các nhóm phụ gồm có các cá nhân mà tác động thực tế của chương trình tới những người này có thể khác nhau, do đó có những gợi ý chính sách hoàn toàn khác nhau. Trên thực tế, đa số các khuyến nghị của đánh giá là cần định hướng mục tiêu chặt chẽ hơn trong các ALP.

#### V. Ai là người tiến hành

Đánh giá này là một phần trong đánh giá tại bốn nước về các chương trình lao động tích cực, với mục đích hiểu tác động của ALP trong những điều kiện kinh tế khác nhau. Đánh giá do một

uỷ ban điều hành dự án giám sát, với các đại diện từ Ngân hàng Thế giới, bốn nước tham dự, các cơ quan tài trợ bên ngoài, và nhà thầu hỗ trợ kỹ thuật (Abt Associates và Viện Upjohn).

Nhóm nghiên cứu đã ký hợp đồng với một doanh nghiệp điều tra tư nhân để tiến hành điều tra- vì những lý do về chất lượng dữ liệu, cũng như để giảm khả năng gây ra sự lo ngại nếu như chính các cơ quan lao động địa phương tiến hành điều tra. Cần chú ý là, uy tín của nghiên cứu có thể bị tác động xấu nếu như cơ quan lao động chịu trách nhiệm tiến hành điều tra. Trên thực tế, khó khăn “rủi ro đạo đức” thường trở thành một vấn đề nghiêm trọng nếu như cơ quan chịu trách nhiệm đào tạo cũng đồng thời chịu trách nhiệm thu thập thông tin về các kết quả của chương trình đào tạo.

## VI. Kết quả

Các kết quả có tính điển hình cho các chương trình đào tạo. Một số dự án có thể có một số tác động (mặc dù tương đối yếu) tới một số dạng công nhân trong những tình huống nhất định. Một điểm quan trọng trong đánh giá là nó phát hiện ra một chương trình có lẽ đã lãng phí tiền bạc, chương trình này không tạo ra tác động gì tới toàn bộ người lao động hay tới một nhóm nhỏ nào. Một điểm mạnh khác của đánh giá nằm ở phần trình bày đánh giá, điều này rất quan trọng nếu như đánh giá sẽ được các nhà hoạch định chính sách nghiên cứu. Trong đánh giá, mỗi chương trình đều có các bảng tóm tắt những lợi ích kết hợp về tiền lương và việc làm, tính cho toàn bộ người tham dự và cho từng nhóm nhỏ.

Một điểm rất tiêu cực trong đánh giá này là đã không có sự phân tích chi phí - lợi ích nào được thực hiện, mặc dù có sự hứa hẹn ban đầu. Việc có thông tin tóm tắt về lợi ích so sánh với các chi phí hiện và ẩn của chương trình, có lẽ sẽ rất có ích. Chẳng hạn, mặc dù người đánh giá có thể nhận thấy một chương trình làm tăng xác suất có việc làm, nhưng cần chú ý

là, lợi ích này đã đi kèm với phí tổn của một chương trình đào tạo kéo dài chín tháng. Việc tính toán đầy đủ tỷ lệ lợi tức của đầu tư sẽ kết hợp được với chi phí hiện của chương trình và chi phí cơ hội về thời gian của người tham dự, và so sánh chi phí này với sự gia tăng thu nhập và việc làm.

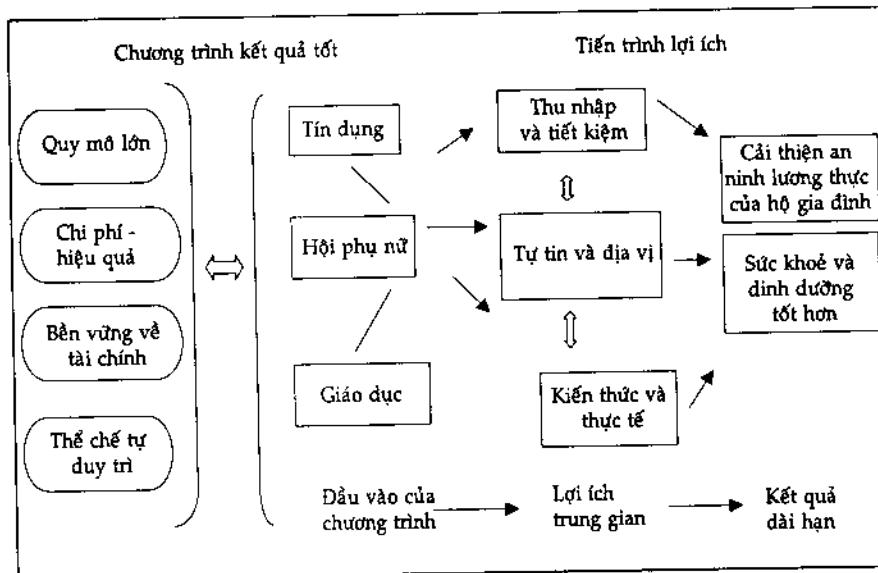
## VII. Các bài học

Có một số bài học quan trọng từ nghiên cứu này. Thứ nhất, phần giới thiệu đã trình bày về những thành phần thực tế của đánh giá, đặc biệt là tầm quan trọng của việc xem xét môi trường chính trị khi thiết kế một chương trình đánh giá. Sự thất bại trong việc thuyết phục cơ quan dịch vụ việc làm về tầm quan trọng của tiến trình đánh giá cũng có nghĩa là các công cụ đánh giá đã bị hạn chế rất nhiều. Thứ hai, nghiên cứu này minh chứng cho lợi ích của việc tạo ra một mẫu có những đặc điểm tương hợp với người tham dự. Cuối cùng, đánh giá này minh họa tầm quan trọng của việc tiến hành phân tích không chỉ với dữ liệu tổng quát, mà còn với từng nhóm nhỏ, nhờ đó, làm tăng khả năng định hướng mục tiêu thành công của chương trình.

## VIII. Nguồn

- Benus, Jacob, Grover Neelima, Jiri Berkovsky, and Jan Rehak. 1998. *Czech Republic: Impact of Active Labor Market Programs*. Cambridge, Mass., and Bethesda, Md.: Abt Associates, May.
- Burtless, Gary. 1995. "The Case for Randomized Field Trials in Economic and Policy Research." *Journal of Economic Perspectives* 9 (2): 63– 84.
- Heckman, James J., and Jeffrey A. Smith. 1995. "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9 (2) : 85– 110.

## Lược đồ sử dụng trong thiết kế đánh giá các chương trình lao động tích cực của Séc



**Phụ lục 1.6: Tác động của tín dụng cho giáo dục tới dinh dưỡng của bà mẹ và trẻ em:  
Chương trình Ngân hàng nông thôn  
vùng Hạ Pra ở Gana**

### I. Giới thiệu

**Mô tả dự án.** Chương trình Tín dụng cho Giáo dục kết hợp các thành phần của chương trình Ngân hàng Grameen và giáo dục trên cơ sở sức khoẻ, dinh dưỡng, thời điểm và khoảng cách sinh đẻ, và kỹ năng kinh doanh nhỏ. Mục tiêu là cải thiện tình trạng dinh dưỡng và an ninh lương thực của các hộ gia đình nghèo ở Gana. Tổ chức Không còn Đói kém, cùng với Chương trình Dinh dưỡng Quốc tế tại Đại học California Davis, cung cấp các dịch vụ Tín dụng cùng với Giáo dục cho phụ nữ nông thôn nghèo ở huyện Shama Ahanta East tại Miền Tây Gana. Một liên minh gồm năm ngân hàng nông thôn được thành lập để cung cấp các dịch vụ này- có hơn 9.000 khoản cho vay, với giá trị tổng cộng 600.000 đôla đã được thực hiện cho tới tháng Ba năm 1997, và tỷ lệ trả nợ chưa bao giờ thấp hơn 92 phần trăm.

**Các đặc điểm nổi bật của đánh giá.** Đánh giá này lý thú vì ba lý do. Thứ nhất, thiết kế mẫu hoàn toàn phù hợp: chương trình được thực hiện ở 19 cộng đồng và các dữ liệu được thu thập trong ba nhóm mẫu phụ nữ khác nhau. Ba nhóm này bao gồm những người đã tham dự chương trình ít nhất là một năm, những người không tham dự nhưng ở trong cộng đồng nhận được chương trình, và những người ở trong các cộng đồng đóng vai trò nhóm đối chứng. Thứ hai, nghiên cứu này mô tả rõ ràng phương pháp của mình: nó xác định và đánh giá các kết quả trung hạn và lâu dài. Cuối cùng, đánh giá này bao gồm cả các kết quả định tính và định lượng, và thường nhấn mạnh kết quả định lượng với một ví dụ nhỏ minh họa.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Các câu hỏi nghiên cứu tập trung vào tác động của chương trình tới (a) tình trạng dinh dưỡng của trẻ em; (b) năng lực kinh tế của phụ nữ (thu nhập, tiết kiệm, thời gian) để đầu tư vào thực phẩm và chăm sóc sức khoẻ; (c) kiến thức của phụ nữ và việc áp dụng cách thức nuôi con bằng sữa mẹ, cai sữa, và phòng chống tiêu chảy; và (d) khả năng của phụ nữ cung cấp đủ dinh dưỡng cho con cái của họ.

Để làm được điều này, đánh giá tách rời các mục tiêu cuối cùng nhằm cải thiện an toàn lương thực và tình trạng dinh dưỡng của hộ gia đình khỏi những lợi ích trung hạn có được nhờ sự thay đổi hành vi, giảm đói nghèo và tăng cường quyền lực của phụ nữ.

Một thiết kế bán thử nghiệm đã được sử dụng trên thực địa trong hai cuộc điều tra (vào năm 1993 và 1996) nhằm đánh giá tác động của chiến lược tới tình trạng dinh dưỡng của trẻ em; năng lực kinh tế của các bà mẹ, quyền lực của phụ nữ và việc áp dụng các biện pháp chăm sóc sức khoẻ và tăng cường dinh dưỡng trẻ em của các bà mẹ. Có tổng cộng 299 cặp mẹ con được điều tra trong thời kỳ đầu và 290 cặp khác trong thời kỳ thứ hai. Cả thông tin định tính và thông tin định lượng đều được thu thập.

Thiết kế đánh giá tương đối phức tạp. Ngân hàng Nông nghiệp vùng Hạ Pra đã xác định được 19 cộng đồng vẫn chưa nhận được dịch vụ Tín dụng cùng với Giáo dục, và các nhà tu vấn đã phân chia các cộng đồng thành những cộng đồng lớn và nhỏ (khoảng 800 cộng đồng), và sau đó lại phân chia căn cứ vào khoảng cách của các cộng đồng này tới những con đường chính. Trong mỗi nhóm, người ta phân công 13 trong số 19 cộng đồng vào trong nhóm tham dự hay nhóm đối chứng. Ba cộng đồng được tham dự vì những lý do chính trị, và có ba cộng đồng khác được lựa chọn như là các nhóm đối chứng, tương tự ba cộng đồng trên, căn cứ vào sự gần gũi về khoảng cách, trình độ phát triển thương mại, quy mô và khả năng tiếp cận những con

đường chính. Có hai cộng đồng đã bị bỏ ra khỏi mẫu do thiếu sự quan tâm và do số cộng đồng theo cách phân loại này ít. Do vậy, trong lần nghiên cứu tiếp theo, chỉ có 17 cộng đồng được điều tra.

Trong các cộng đồng nhỏ, có mười cặp mẹ con, trong đó các trẻ em từ 12 đến 23 tháng tuổi được lựa chọn tham gia điều tra gốc. Trong các cộng đồng lớn, có 30 cặp mẹ con được lựa chọn. Ở đây phát sinh hai vấn đề quan trọng. Vấn đề thứ nhất là cách lựa chọn như vậy làm cho điều tra không theo dõi được những phụ nữ được phỏng vấn lần trước bởi vì rất ít phụ nữ từng tham dự điều tra gốc cũng có con nhỏ trong kỳ điều tra vào năm 1996. Vấn đề thứ hai là sự hạn chế về lứa tuổi làm mẫu thứ hai giảm đi rất nhiều, khiến cho nó phải được mở rộng tới những phụ nữ có con dưới ba tuổi vào năm 1996. Một ưu điểm quan trọng của thiết kế đánh giá phức tạp này là việc nó cho phép phân loại phụ nữ trong mẫu cơ sở thành những đối tượng tham dự và không tham dự trong tương lai.

Có ba dạng phụ nữ được điều tra: những người tham dự, những người không tham dự nhưng sinh sống trong các cộng đồng nhận được chương trình, và những người sống ở các cộng đồng đóng vai trò nhóm đối chứng. Tất cả các thành viên tham dự chương trình đều được điều tra; hai dạng sau được lựa chọn ngẫu nhiên từ những phụ nữ có con dưới ba tuổi. Cần lưu ý là, quy mô mẫu tổng cộng (bằng 360) được tính toán dựa theo độ lệch chuẩn tìm thấy trong các nghiên cứu trước đây. Có một yêu cầu là mẫu phải có thể phát hiện được chênh lệch bằng 0,4 trong giá trị tính theo z của nhóm đối chứng và nhóm mục tiêu tại mức độ ý nghĩa bằng 0,05 và độ vững mạnh (power) bằng 0,8.

### III. Dữ liệu

Người ta thu thập cả các dữ liệu định lượng và định tính về các hộ gia đình, bà mẹ và trẻ em, tập trung vào các thước đo trung

hạn và dài hạn - và đặc biệt vào tính đa chiều của các kết quả.

Đối với các kết quả trung hạn, sẽ có một tập hợp các câu hỏi nhằm đo lường năng lực kinh tế của phụ nữ (thu nhập, lợi nhuận, đóng góp thu nhập của hộ gia đình, tiết kiệm, kỹ năng kinh doanh, và chi tiêu cho lương thực và hộ gia đình). Cũng tương tự, có một tập hợp các thước đo khác nhằm vào kiến thức về sức khoẻ và dinh dưỡng của phụ nữ (nuôi con bằng sữa mẹ, nuôi trẻ, phòng và chống bệnh tiêu chảy và tiêm chủng). Còn có một tập hợp thước đo khác đo lường quyền lực của phụ nữ (sự tự tin và hy vọng vào tương lai, địa vị và quyền ra quyết định trong gia đình, địa vị và mạng lưới xã hội trong cộng đồng). Đối với các kết quả cuối cùng, như tình trạng dinh dưỡng và an ninh lương thực, người ta sử dụng các thước đo trực tiếp hơn (các thước đo nhân trắc học cho tình trạng dinh dưỡng và các câu hỏi về tình trạng đói trong vấn đề về an ninh lương thực).

Mặc dù đã dự kiến quy mô mẫu tổng cộng gồm 360 cặp mẹ và con, nhưng chỉ có 299 cặp được phỏng vấn trong đợt điều tra đầu (chủ yếu vì có hai cộng đồng đã bị bỏ qua) và 290 cặp trong đợt điều tra thứ hai. Các đặc điểm của những bà mẹ và hộ gia đình được so sánh trong mỗi nhóm ở cả ba nhóm này và người ta không nhận thấy có sự khác biệt đáng kể nào cả.

#### IV. Các kỹ thuật kinh tế lượng

Các kỹ thuật kinh tế lượng được sử dụng tương đối đơn giản và khai thác điểm mạnh của thiết kế điều tra. Người ta tính giá trị trung bình của nhóm cho mỗi thước đo kết quả và sau đó tiến hành kiểm định  $t$  để đánh giá liệu có sự khác biệt đáng kể giữa nhóm đối chứng và nhóm tham dự không. Đây thực chất là phương pháp khác biệt đơn giản. Các kết quả được bổ sung bằng nhiều hình vẽ.

Tuy nhiên, vẫn có một số câu hỏi quan trọng chưa được trả lời. Thứ nhất, thiết kế mẫu bị phân mảnh, và bởi vì, chủ yếu là

do thiết kế, kết quả của mỗi cặp mẹ con sẽ có tương quan với các kết quả khác trong cộng đồng, nên sai số chuẩn sẽ bị sai lệch thấp hơn và thống kê sẽ bị sai lệch cao hơn thực tế. Trong trường hợp thái cực, tất cả các kết quả đều tương quan hoàn hảo với nhau và quy mô mẫu thực tế chỉ còn là 17 thay vì 300. Hiện tượng này làm cho các kết quả trở nên có ý nghĩa trong khi thực ra chúng không có ý nghĩa. Thứ hai, mặc dù thiết kế được phân lớp một cách rõ ràng, nhưng tác động của sự phân lớp này vẫn chưa được xác định. Đó là câu hỏi liệu các cộng đồng có quy mô lớn hay nhỏ sẽ hưởng lợi nhiều hơn từ dự án, hoặc liệu các cộng đồng sống gần đường có trở nên giàu có hơn những cộng đồng ở cách xa đường hay không. Điều này thật đáng ngạc nhiên, bởi vì lý do có thể cho một thiết kế mẫu như vậy là căn cứ từ các ngụ ý chính sách. Thứ ba, mặc dù các vấn đề về sai lệch do lựa chọn được đưa ra thảo luận nhưng không có một sự phân tích chính thức hay biện pháp để khắc phục vấn đề quan trọng này. Cuối cùng, mặc dù có sự khác biệt đáng kể về tỷ lệ không trả lời ở từng khoản mục - điều này cho thấy còn có khả năng xảy ra sai lệch lựa chọn ngay cả ở trong điều tra này, nhưng vấn đề này vẫn không được giải quyết hay nêu ra trong đánh giá.

## V. Ai là người tiến hành

Một viện nghiên cứu quốc tế hoạt động phi lợi nhuận có tên là Không còn Đói kém, đã xây dựng chương trình Tín dụng cho Giáo dục và phối hợp với Chương trình Dinh dưỡng quốc tế tại Đại học California Davis để đánh giá chương trình này. Viện này kết hợp với Ngân hàng Nông thôn vùng Hạ Pra (một ngân hàng độc lập, do Ngân hàng Gana quản lý), và sau này là với bốn ngân hàng nông thôn khác ở Gana, để thực hiện chương trình. Ngân hàng Nông thôn vùng Hạ Pra đóng vai trò quan trọng trong việc xác định và lựa chọn các cộng đồng sẽ được điều tra.

## VI. Kết quả

Mục tiêu trung hạn nói chung đã đạt được: mặc dù thu nhập và chi tiêu của phụ nữ không tăng, nhưng kỹ năng kinh doanh và tỷ lệ tiết kiệm của phụ nữ đã tăng lên đáng kể. Hiểu biết của phụ nữ về sức khoẻ và dinh dưỡng nói chung cũng được cải thiện. Phụ nữ có lẽ cũng nhận thấy sự gia tăng quyền lực của mình. Đối với mục tiêu cuối cùng, đánh giá này cho thấy chương trình đã cải thiện được tình trạng an ninh lương thực và dinh dưỡng trẻ em cho các hộ gia đình, nhưng vẫn chưa cải thiện được tình trạng dinh dưỡng của bà mẹ.

## VII. Các bài học

Một đóng góp quan trọng của đánh giá này là có thiết kế mẫu rất thú vị: việc phân lớp và lựa chọn các nhóm tham dự và không tham dự liên quan tới sự tham dự trong tương lai của các đối tượng này là một phương pháp rất có ích. Một bài học khác là việc sử dụng có hiệu quả các chỉ tiêu do lưỡng kết quả - đôi khi bao gồm cả các nhân tố phi lượng hoá, như quyền lực của phụ nữ. Một bài học quan trọng khác là giá trị của dữ liệu phi lượng hoá nhằm minh họa cho các kết quả định lượng.

## VIII. Nguồn

MkNelly, Barbara, and Christopher Dunford (in collaboration with the Program in International Nutrition, University of California Davis). 1998. "Impact of Credit with Education on Mothers' and their Young Children's Nutrition: Lower Pra Rural Bank Credit with Education Program in Ghana." Freedom from Hunger Research Paper No. 4, March.

## Phụ lục 1.7: Sách giáo khoa và điểm thi: Bằng chứng từ một đánh giá tiềm năng ở Kenya

### I. Giới thiệu

**Mô tả dự án.** Đánh giá tác động của các hình thức chi tiêu cho giáo dục khác nhau tới kết quả học tập của học sinh có vai trò đặc biệt quan trọng ở các nước đang phát triển. Các nghiên cứu trước đây cho thấy, cung cấp sách giáo khoa là một phương cách hiệu quả làm tăng kết quả điểm thi. Kenya, với những nguồn lực đặc biệt khan hiếm đối với các nhà giáo dục, là một tình huống nghiên cứu tốt. Những người đánh giá nhận thấy, chỉ có một trên sáu trẻ em ở nước này hiện đang học các lớp 3, 4 và 5 là có sách giáo khoa; con số này tăng lên một bốn ở các lớp cao hơn. Hơn nữa, trang thiết bị vật chất đặc biệt nghèo nàn và có nhiều trẻ em phải ngồi trên sàn để học.

Đánh giá này đánh giá tác động tới kết quả học tập của một chương trình được thực hiện vào năm 1996. Trong chương trình này, người ta lựa chọn ngẫu nhiên và cung cấp sách giáo khoa cho tất cả các lớp ở 25 trong số 100 trường tiểu học ở nông thôn. Sách giáo khoa tiếng Anh được trao cho các lớp từ lớp 3 đến lớp 7, với tỷ lệ sáu quyển sách cho 10 trẻ em; sách giáo khoa môn toán được cung cấp cho các lớp 3, 5 và 7 với tỷ lệ 50 phần trăm; và sách giáo khoa môn khoa học được dành cho lớp 8, với tỷ lệ 60 phần trăm.Thêm vào đó, mỗi lớp học nhận được một cuốn sách hướng dẫn giáo viên . Người ta tiến hành kiểm tra kết quả học tập của học sinh trước khi phân phát sách giáo khoa và kiểm tra lại vào 10 tháng sau đó. Các bài thi tương tự cũng được áp dụng ở các trường đóng vai trò đơn vị đối chiếu. Phương pháp này là sự kết hợp giữa thiết kế lựa chọn ngẫu nhiên với so sánh phản thân.

**Các đặc điểm nổi bật của đánh giá.** Đánh giá này là một minh

hoạ tuyệt vời về việc phát triển và thực thi một thiết kế điều tra tốt, và sau đó áp dụng các kỹ thuật kinh tế lượng phù hợp. Đánh giá đặc biệt đáng tin cậy khi chỉ ra cách thức rút ra kết luận từ những kết quả ở các cấp với dữ liệu có quy mô lớn, việc sử dụng các ước lượng khác biệt trong khác biệt, cách thức khắc phục sai lệch do lựa chọn và triệt tiêu lẫn nhau (selection and attrition bias), cũng như sai số thống kê và những vấn đề làm thoái giảm (crowding-out issues). Một cấu thành rất đáng chú ý khác của đánh giá là sự tập trung vào tác động của chương trình tới học sinh tại tất cả các đoạn phân phối. Cuối cùng, việc thừa nhận và phân tích về các tác động thứ cấp tiềm tàng là một ví dụ rất tốt cho việc xem xét tất cả các khía cạnh của một chương trình.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Nghiên cứu này tập trung chủ yếu vào đánh giá tác động của sách giáo khoa tới kết quả học tập. Do đây là một khái niệm phức tạp nên các kết quả được đo lường bằng sự khác biệt giữa những trường nhận được sách giáo khoa và những trường so sánh trên một số khía cạnh: điểm kỳ kiểm tra sau khi có chương trình, mức gia tăng điểm thi, chênh lệch giữa tổ hợp môn học- lớp học nhận và không nhận sách giáo khoa, cũng như hoạt động của trẻ em và của giáo viên. Đánh giá này cũng xem xét những tác động thứ cấp khác (mà thường bị bỏ qua), cụ thể là khả năng hình thức trợ cấp này sẽ làm giảm đi sự đóng góp của cha mẹ, nhất là bằng cách lấn át các nguồn vốn huy động khác.

Thiết kế đánh giá này tương đối phức tạp. Bộ Giáo dục lựa chọn ra 100 trường có nhu cầu tham dự chương trình vào năm 1995. Các trường này được chia thành bốn nhóm- trước tiên căn cứ vào vị trí địa lý, và sau đó căn cứ vào danh sách theo bảng chữ cái các trường nằm trong khu vực địa lý này. Sau đó người ta lựa chọn các trường trong bốn nhóm trên cơ sở danh sách theo bảng chữ cái. Nhóm đầu tiên nhận được sự hỗ trợ về sách

giáo khoa vào năm 1996, nhóm thứ hai vào năm 1997, và cứ thế tiếp tục. Sách giáo khoa các môn Toán, Tiếng Anh và Khoa học được cung cấp cho các lớp khác nhau - chủ yếu là các lớp từ 3 đến 7.

### III. Dữ liệu

Các kỳ thi Toán, Tiếng Anh và Khoa học được tổ chức cho học sinh tại tất cả các lớp học trong 100 trường học trước khi phân phát sách giáo khoa. Tuy nhiên, đánh giá cũng đã tận dụng các kỳ kiểm tra trước được thực hiện ở các lớp từ lớp 3 đến lớp 7 vào tháng Mười năm 1996 và các kỳ kiểm tra sau vào tháng Mười năm 1997. Nhờ vậy, những người đánh giá có được dữ liệu của hơn 8.800 học sinh (ở tất cả các lớp) cho mỗi môn học ở 100 trường và có tổng cộng 26.000 quan sát. Do có 25 trường nhận được sách giáo khoa trong thời kỳ này, nên học sinh ở những trường này trở thành nhóm "có sách giáo khoa"; 75 trường còn lại là nhóm đối chứng. Ngoài điểm thi, người ta còn tiến hành thu thập dữ liệu về tình hình tài chính và phương pháp sư phạm của các trường.

Thông tin về việc sử dụng sách giáo khoa tại lớp học được thu thập bởi những người quan sát được huấn luyện, những người này sẽ tham quan từng trường học và ghi chép liên tục về tám hoạt động có thể ở lớp học (từ hoạt động của giáo viên và học sinh cho tới việc sử dụng sách giáo khoa của giáo viên và học sinh). Các ghi chép này ghi nhận các hoạt động trong vòng 15 phút, và sau đó được sử dụng để xác định tỷ lệ phần trăm thời gian mà giáo viên và học sinh dành cho mỗi hoạt động khác nhau trong tổng cộng 551 thời kỳ tại lớp học. Cán bộ thực địa phỏng vấn từ bốn đến năm học sinh ở mỗi lớp và điền vào bảng câu hỏi căn cứ vào câu trả lời của học sinh.

Cuối cùng, dữ liệu về tình hình tài chính được thu thập từ bảng câu hỏi dành cho các trường học và uỷ ban trường học vào năm 1997. Bảng câu hỏi này nêu ra các câu hỏi về tình hình huy động tài trợ của các nhà trường.

#### IV. Các kỹ thuật kinh tế lượng

Cần lưu ý tới các vấn đề thú vị được kỹ thuật chọn mẫu này tạo ra. Các điểm thi ở trong cùng một trường, cũng như các điểm thi ở trong cùng một lớp có thể có tương quan với nhau. Tương tự, điểm thi các môn học khác nhau của cùng một học sinh cũng sẽ tương quan với nhau. Cũng có thể đánh giá chương trình này căn cứ vào tác động của chương trình tới trình độ học tập của học sinh, hoặc tới lợi ích học tập của học sinh. Nhìn chung, tác động của chương trình cần phải có tính vững vàng (robust) khi sử dụng các kỹ thuật kinh tế lượng và các phương pháp xem xét dữ liệu khác nhau. Đánh giá ở đây đã đạt được yêu cầu này.

Đánh giá được tiến hành trước hết bằng cách cung cấp các ước lượng từ một phương trình hồi quy đơn giản có biến giả. Biến giả này được tạo ra cho mỗi kết hợp điểm thi - môn học phản ánh tác động ngẫu nhiên của trường học, lớp học, và môn học (biến phụ thuộc là thay đổi điểm thi giữa kỳ thi trước và kỳ thi sau). Một điểm hấp dẫn của phương pháp này là có thể kết hợp các biến giả theo những cách thức rất có ích:

- Kết hợp một số lớp với nhau để ước lượng tác động của sách giáo khoa tới một môn học
- Kết hợp tất cả các điểm thi để ước lượng tác động trung bình của sách giáo khoa tới một lớp học; và
- Kết hợp tất cả các lớp và các môn để ước lượng tác động bình quân gia quyền của sách giáo khoa tới tất cả các lớp và các môn.

Rõ ràng, cơ cấu của tác động ngẫu nhiên khác nhau trong mỗi cách tiếp cận nêu trên, và đánh giá đã nêu rất rõ ràng nội dung này.

Đánh giá sau đó sử dụng phương pháp khác biệt trong khác biệt. Phương pháp này tương đối đơn giản do chỉ so sánh điểm thi trong các kỳ thi sau và kỳ thi trước giữa các trường trong

nhóm tham dự và nhóm đối chứng với nhau.

Phương pháp thứ ba cần được thảo luận nhiều hơn. Phương pháp này tương đối phức tạp hơn do nó khai thác sự khác biệt ở trong mỗi trường. Phương trình hồi quy được sử dụng ở đây liên quan tới việc hồi quy điểm thi theo các biến giả phản ánh việc học sinh có (a) ở trong một trường nhận được sách giáo khoa và (b) ở trong một kết hợp môn học- lớp học nhận được sách giáo khoa hay không. Điều này làm giảm các vấn đề xảy ra do hiện tượng phương sai không đồng nhất (heterogeneity) của các trường, cũng như các vấn đề trong lựa chọn mẫu. Sở dĩ phương pháp này làm giảm các vấn đề trong lựa chọn mẫu vì nó cho thấy tác động tới điểm thi của cùng một học sinh phụ thuộc vào việc học sinh này có nhận được sách giáo khoa hay không. Tuy vậy, phương pháp này giả định là điểm thi ở các kết hợp lớp học- môn học khác nhau có thể được cộng và trừ cho nhau, và giả định rất rõ ràng này có thể là lý do dẫn đến những kết quả rất khác nhau trong phương pháp này.

Một chủ đề thường gặp phải trong đánh giá là ước muốn không chỉ nắm bắt được tác động trung bình của chương trình, mà còn cả tác động của chương trình tới những nhóm phụ trong số những người tham dự chương trình. Đánh giá này là một minh họa rất có ích cho việc sử dụng các hạng tử tương tác (interaction terms) và hồi quy theo khoảng phân nhóm. Phương pháp thứ nhất liên quan đến sự tương tác giữa điểm thi ban đầu và biến giả sách giáo khoa để xác định tác động của sách giáo khoa tới những học sinh khá giả so với học sinh nghèo, sử dụng cả các giá trị thực tế và giá trị công cụ (điểm thi ban đầu có tương quan với hạng tử sai số, vì thế gây ra sai lệch). Phương pháp thứ hai, sử dụng hồi quy quantile cũng rất có ích và ngày càng trở nên phổ biến. Cụ thể, trong khi phương pháp hồi quy bình phương nhỏ nhất chỉ cho biết tác động trung bình của chương trình sách giáo khoa, thì hồi quy theo khoảng phân nhóm cho phép có sự khác nhau trong tác động của chương trình phụ thuộc vào vị trí của học sinh trong hàm phân phối.

Đánh giá này cũng đặc biệt có ích trong việc cung cấp một

phương pháp xem xét các sai lệch lựa chọn và sai lệch tiêu hao. Nguyên nhân tiềm tàng chủ yếu gây ra các vấn đề sai lệch này là sự khác nhau về tỷ lệ lên lớp và tỷ lệ lưu ban của các trường nhận được sách giáo khoa và các trường so sánh. Chẳng hạn, trẻ em có thể được lên lớp từ lớp hai (lớp không nhận được sách giáo khoa) lên lớp ba (lớp nhận được sách giáo khoa) tại các trường nhận được sách giáo khoa. Tỷ lệ lên lớp khác nhau làm sai lệch theo chiều hướng giảm đối với kết quả tại các lớp phải nhận các học sinh kém nhất, và có thể làm sai lệch theo chiều hướng tăng đối với kết quả ở những lớp mà các học sinh này từ đó ra đi. Cả hai tác động này đều được đưa vào trong đánh giá bằng cách ước lượng lại mô hình theo hai cách: đó là bỏ qua tất cả những học sinh lưu ban ở hai nhóm trường học và bỏ qua những học sinh yếu nhất ở mỗi lớp học. Tính vững vàng của kết quả trong cả hai phương pháp này đã xác nhận tác động của chương trình.

Cuối cùng, để minh họa tầm quan trọng của các tác động thứ cấp, đánh giá này lượng hoá tác động của việc cung cấp sách giáo khoa tới nguồn tài chính đóng góp của cha mẹ học sinh cho nhà trường. Nghiên cứu này nhận thấy chương trình không làm suy giảm phần đóng góp của cha mẹ học sinh- số tiền đóng góp không ICS cho các trường so sánh là 465 đôla và cho các trường có sách giáo khoa là 267 đôla (giá trị trung bình của sách giáo khoa ICS là 485 đôla). Nghiên cứu này sử dụng phân tích hồi quy đơn giản, điều tra và khẳng định cho giả thuyết: tại các trường nhỏ, sự gây ra hiệu ứng thóai giảm của chương trình tới đóng góp của cha mẹ học sinh xảy ra ở mức độ lớn hơn so với tại các trường lớn.

**Ai là người tiến hành.** Một tổ chức phi lợi nhuận của Hà Lan, International Christelijk Steunfonds, đã tài trợ cho dự án. Đánh giá do một giáo sư của Học viện công nghệ Massachusetts (ông Kremer) và hai nhà kinh tế của Ngân hàng Thế giới (ông Paul Glewwe và bà Sylvie Moulin) tiến hành. Một số chi phí do National Science Foundation, và Uỷ ban nghiên cứu của Ngân hàng Thế giới chi trả.

## V. Kết quả

Kết quả của đánh giá này trái ngược căn bản với kết quả trong các đánh giá tương tự về những chương trình cung cấp sách giáo khoa. Kết quả cơ bản là tính trung bình, không có tác động đáng kể từ việc cung cấp sách giáo khoa tới kết quả học tập, mặc dù tác động này là có ý nghĩa đối với các học sinh học khá. Kết quả này có tính vững chắc khi sử dụng các kỹ thuật ước lượng khác nhau và khi cắt giảm dữ liệu.

## VI. Các bài học

Bài học có ích nhất từ đánh giá này là tầm quan trọng của việc sử dụng các kỹ thuật kinh tế lượng khác nhau để kiểm tra tính vững chắc của các kết quả thực nghiệm. Mặc dù công việc thu thập dữ liệu đã được tiến hành rất tốt, tác động dự kiến của chương trình vẫn hầu như không thay đổi khi sử dụng các giả định kinh tế lượng và phát biểu mô hình khác nhau. Việc sử dụng phương pháp hồi quy quantile và các hạng tử tương tác cũng là một biện pháp rất có ích nhằm phân tích tác động tới các nhóm nhỏ khác nhau trong tổng thể. Cuối cùng, cần tìm hiểu và xác định tác động thứ cấp- trong trường hợp này là khả năng xảy ra sự lấn át tài trợ.

## VII. Nguồn

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 1998. "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya." Development Resrach Group (DECRG), World Bank, Washington, D.C. Processed.

## Phụ lục 1.8: Đánh giá dự án khuyến nông của Kênia

### I. Giới thiệu

**Mô tả dự án.** Dự án Khuyến nông Quốc gia đầu tiên (NEP-I) ở Kênia đã đưa vào sử dụng hệ thống quản lý dịch vụ khuyến nông Đào tạo và tham quan (T &V), năm 1983. Dự án này có mục tiêu kép là phát triển thể chế và cung cấp dịch vụ khuyến nông cho nông dân với mục đích nâng cao năng suất nông nghiệp. NEP-II được triển khai sau đó, vào năm 1991 và có mục đích củng cố những lợi ích nhận được từ NEP-I bằng cách tăng cường sự liên hệ trực tiếp với nông dân, cải thiện tính phù hợp của các thông tin và kỹ thuật khuyến nông, nâng cấp trình độ của cán bộ và nông dân, và tăng cường phát triển thể chế.

**Đánh giá tác động.** Hoạt động của hệ thống khuyến nông Kênia gây ra nhiều tranh cãi và là một phần trong cuộc tranh luân rông rãi về chi phí - hiệu quả của T&V áp dụng trong khuyến nông. Bất chấp cường độ của tranh luận, vai trò quan trọng của dịch vụ khuyến nông trong chiến lược phát triển của Ngân hàng Thế giới đối với châu Phi, và số vốn đầu tư lớn đã được thực hiện, có rất ít những nỗ lực nghiêm túc nhằm do lường tác động của khuyến nông sử dụng T&V. Trong trường hợp của Kênia, cuộc tranh luận trở nên gay gắt hơn do một nghiên cứu trước đó ước tính mức lợi tức dự kiến rất cao trong khi không có những kết quả rõ ràng có tính thuyết phục trên thực tế, và ngành nông nghiệp Kênia đạt được các kết quả kém cỏi trong những năm gần đây.

Sự bất đồng (giữa Phòng đánh giá Hoạt động và Khu vực châu Phi của Ngân hàng Thế giới) về kết quả hoạt động của NEP-I đã làm trì hoãn đánh giá đối với chương trình này. Đánh giá này áp dụng một phương pháp thực nghiệm chắc chắn nhằm đánh giá tác động của chương trình tới hoạt động nông

nghiệp. Sử dụng khung quản lý dựa trên kết quả, đánh giá xem xét tác động của dịch vụ khuyến nông tới năng suất và hiệu suất nông nghiệp. Đánh giá này cũng xây dựng các chỉ tiêu phản ánh kết quả của chương trình (như sự nhận thức và áp dụng kỹ thuật mới của nông dân) và các đầu ra (ví dụ, tần xuất và chất lượng các liên hệ với nông dân) nhằm đánh giá hoạt động của hệ thống khuyến nông, và xác nhận tác động thực sự hay tiềm tàng của chương trình.

## II. Thiết kế đánh giá

Chiến lược đánh giá minh họa các kỹ thuật tốt nhất bằng cách sử dụng nhiều phương pháp đánh giá nhằm đánh giá sự thực thi, kết quả, và tác động của chương trình tới năng suất và hiệu quả nông nghiệp. (Đánh giá này không nghiên cứu tác động tới phúc lợi hộ gia đình, phúc lợi này có thể bị tác động của một số nhân tố ngoài phạm vi hoạt động của T&V.) Đánh giá sử dụng các phương pháp định lượng và định tính, nhờ đó có thể bổ sung cho các kết quả thực nghiệm về tác động chương trình bằng sự đánh giá của người thụ hưởng và các cuộc phỏng vấn của nhân viên, trong đó nêu bật những vấn đề thực tế xảy ra trong tiến trình thực hiện chương trình. Nghiên cứu này cũng áp dụng phương pháp đánh giá có điều kiện nhằm suy đoán mức độ sẵn sàng thanh toán cho dịch vụ khuyến nông của nông dân. [Phương pháp đánh giá có điều kiện suy đoán các giá trị sử dụng và không sử dụng của cá nhân đối với những hàng hoá và dịch vụ công cộng và tư nhân khác nhau. Những người được phỏng vấn trả lời các câu hỏi, nêu rõ mức độ sẵn sàng thanh toán của mình để tránh một sự thay đổi giả định trong cung cấp hàng hoá hay dịch vụ- tức là, kết quả "có điều kiện". Trong trường hợp này, các nông dân được hỏi, họ sẵn sàng trả bao nhiêu tiền để tiếp tục nhận được các dịch vụ khuyến nông nếu như chính phủ ngừng cung cấp các dịch vụ này cho họ.]

Sự đánh giá định lượng có tính phức tạp do hệ thống T &V được áp dụng trên quy mô toàn quốc, vì vậy không thể so sánh

giữa nhóm tham dự chương trình và nhóm không tham dự chương trình (nhóm đối chứng). Do đó, phương pháp đánh giá tìm cách khai thác dữ liệu về sản xuất nông nghiệp của hộ gia đình trước khi có dự án để có thể so sánh phần nào giữa trước và sau dự án bằng phương pháp dữ liệu lặp. Để thực hiện điều này, dữ liệu hiện có về các hộ gia đình được bổ sung bằng một đợt điều tra hoàn toàn mới để tạo ra một nhóm lặp. Trong nghiên cứu này, không thể tiến hành đánh giá về người thụ hưởng, nhưng đánh giá đã rút ra được những kết luận có liên quan từ hai đánh giá người thụ hưởng gần đây ở Kenya. Nghiên cứu này đáng chú ý do nó được xây dựng trên những nguồn dữ liệu có từ trước ở Kenya (điều tra hộ gia đình, đánh giá sự tham dự, v.v), được bổ sung bằng một nỗ lực thu thập dữ liệu có tính toàn diện nhằm phục vụ cho đánh giá.

### **III. Thu thập dữ liệu và các kỹ thuật phân tích**

Phương pháp đánh giá sử dụng những nguồn dữ liệu định tính và định lượng sẵn có. Đánh giá định lượng căn cứ chủ yếu vào điều tra hộ gia đình do Phòng đánh giá hoạt động của Ngân hàng Thế giới tiến hành vào năm 1998. Điều tra này tạo ra dữ liệu lặp bằng cách đến thăm lại những hộ gia đình từng tham gia một đợt điều tra hộ gia đình vào năm 1990 do Phòng kỹ thuật châu Phi tiến hành. Điều tra năm 1990 lại căn cứ vào một mẫu nhỏ trong Điều tra chỉ tiêu của hộ gia đình nông thôn vào năm 1982. (Ba điều tra này tạo ra dữ liệu lặp cho khoảng 300 hộ gia đình. Dữ liệu được điều tra bao gồm các đặc điểm về nhân khẩu học của hộ gia đình, các đặc điểm của nông trại, dữ liệu đầu vào- đầu ra trong sản xuất nông nghiệp; điều tra năm 1990 và 1998 cũng thu thập thông tin liên quan tới dịch vụ khuyến nông, kể cả sự nhận thức và áp dụng các thông điệp khuyến nông.) Những dữ liệu này được bổ sung bằng một cuộc điều tra về cán bộ khuyến nông, một số lần đánh giá lại hoạt động khuyến nông do Bộ Nông nghiệp tiến hành hay đặt hàng, và các cuộc thảo luận cá nhân và nhóm tập trung với các cán

bộ khuyến nông. Nghiên cứu này cũng căn cứ vào hai đánh giá người thụ hưởng gần đây. Đó là nghiên cứu của tổ chức ActionAid Kenya vào năm 1997, nghiên cứu này suy đoán quan điểm của người sử dụng và người sử dụng tiềm năng về các dịch vụ khuyến nông của Kenya; và Đánh giá Đói nghèo Tham dự, do Ngân hàng Thế giới, Tổ chức Phát triển Hải ngoại Anh, Quỹ Y học và Nghiên cứu châu Phi, UNICEF và chính phủ Kenya phối hợp tiến hành, trong đó nêu ra những câu hỏi về các dịch vụ công cộng, bao gồm cả khuyến nông.

Phân tích này đánh giá cả tiến trình thực hiện và kết quả của chương trình T&V của Kenya. Nghiên cứu này đánh giá sự phát triển thể chế căn cứ vào các dữ liệu thứ cấp và định tính - điều tra cán bộ, phỏng vấn, và đánh giá của bộ về dịch vụ khuyến nông. Chất lượng và số lượng dịch vụ cung cấp được đánh giá bằng cách kết hợp giữa kết quả của đánh giá tham dự (đánh giá người thụ hưởng), điều tra cán bộ, và do lường phạm vi, bản chất và tần xuất mối liên hệ giữa các cán bộ khuyến nông và nông dân căn cứ vào một cuộc điều tra năm 1998 của OED. Dữ liệu điều tra cũng được sử dụng để đánh giá kết quả của chương trình, kết quả này được đo bằng sự nhận thức của nông dân và việc áp dụng các khuyến nghị trong khuyến nông.

Kết quả của chương trình - tác động thực sự của chương trình tới sản xuất nông nghiệp ở Kenya- được đánh giá trên cơ sở mối quan hệ giữa việc cung cấp dịch vụ khuyến nông với các thay đổi trong năng suất và hiệu suất ở cấp nông trại. Trên cơ sở dữ liệu lặp của các hộ gia đình, người ta ước lượng các tác động của chương trình bằng cách sử dụng phân tích bao phủ dữ liệu (data envelopment analysis), một kỹ thuật phi tham số, nhằm đo lường các thay đổi trong hiệu quả và năng suất của nông dân theo thời gian, và phân tích kinh tế lượng để đo lường tác động của việc cung cấp dịch vụ khuyến nông tới sản xuất nông nghiệp. Phương pháp đánh giá có điều kiện được sử dụng để ước đoán trực tiếp mức độ sẵn sàng thanh toán của nông dân cho các dịch vụ khuyến nông.

#### IV. Kết quả

Sự phát triển thể chế của NEP-I và NEP-II là tương đối hạn chế. Sau 15 năm, hiệu quả của dịch vụ khuyến nông vẫn không được cải thiện đáng kể. Mặc dù gần đây đã có sự xem xét lại đối với các phương pháp khuyến nông, nhưng nhìn chung, chương trình khuyến nông thiếu tầm nhìn chiến lược cho sự phát triển trong tương lai. Việc quản lý hệ thống tiếp tục yếu kém, và các hệ thống thông tin hầu như không tồn tại. Chất lượng và số lượng cung cấp dịch vụ yếu kém. Nhiều người thụ hưởng và các cán bộ khuyến nông đều cho biết các cuộc tham quan diễn ra không thường xuyên và ít có hiệu quả. Mặc dù vẫn còn những nhu cầu chưa được đáp ứng, nhưng dịch vụ khuyến nông công cộng vẫn tập trung vào các thông điệp kinh tế nông nghiệp đơn giản và cơ bản. Thế nhưng, phương pháp được sử dụng- đó là tần xuất liên hệ cao với một số ít các nông dân- lại phù hợp với việc cung cấp các thông tin phức tạp về kỹ thuật hơn. Kết quả là , một hệ thống cung cấp dịch vụ tốn kém và thiếu hiệu quả. Hoạt động khuyến nông tác động rất ít tới sự thay đổi nhận thức và việc chấp nhận các khuyến nghị, điều này cho thấy tác động hạn chế của chương trình. Về tác động thực sự của chương trình tới sản xuất và hiệu quả sản xuất nông nghiệp, các dữ liệu cho thấy, dịch vụ khuyến nông có tác động tích cực nhỏ tới hiệu suất kỹ thuật, nhưng không có tác động tới hiệu suất phân bón hay hiệu suất kinh tế toàn bộ. Hơn nữa, không thể xác định được tác động có ý nghĩa từ việc cung ứng dịch vụ khuyến nông đến năng suất nông trại bằng những dữ liệu hiện có. Mặc dù vậy, các dữ liệu cho thấy khuyến nông có tác động tương đối lớn hơn tại những khu vực trước kia có năng suất thấp, những nơi có chênh lệch về tri thức lớn nhất. Các kết quả như vậy phù hợp với các kết quả từ đánh giá có điều kiện. Đại đa số nông dân, kể cả những người tham dự và chưa tham dự chương trình, đều sẵn sàng trả tiền cho những lời tư vấn, thực tế này cho thấy tồn tại một nhu cầu chưa được đáp ứng. Tuy nhiên, giá trị ước tính của dịch vụ, tính trên những gì nó cung

cấp, thấp hơn nhiều so với chi phí mà hiện nay chính phủ phải bỏ ra để cung cấp.

## V. Ứng dụng chính sách

Đánh giá về Dịch vụ khuyến nông của Kenya nổi bật, nhờ các kết luận chính sách có tính thực tế được rút ra từ kết quả đánh giá, trong đó nhiều kết quả liên quan tới thiết kế của các dự án khuyến nông trong tương lai. Thứ nhất, đánh giá cho thấy nhu cầu cần nâng cao tính định hướng đối tượng của các dịch vụ khuyến nông, tập trung vào những lĩnh vực và nhóm người tại đó chênh lệch giữa phương pháp trung bình và phương pháp tốt nhất là cao nhất, và vì vậy, tác động của khuyến nông cũng có thể cao nhất. Hơn nữa, các lời khuyên cần phải thích hợp để đáp ứng được nhu cầu của nông dân, và có chú ý tới sự khác biệt về mặt kỹ thuật và điều kiện kinh tế giữa các địa phương. Việc thực hiện thành công mức độ định hướng dịch vụ này đòi hỏi phải thường xuyên có những luồng thông tin phù hợp, đáng tin cậy và đúng lúc, và cần phải có một hệ thống theo dõi và đánh giá để tạo ra các phản hồi thường xuyên từ phía những người thụ hưởng dịch vụ.

Để nâng cao hiệu quả của chương trình, sự hiện diện với tần suất ít hơn của các cán bộ khuyến nông, nhưng với độ bao phủ rộng hơn, có thể sẽ có hiệu quả về chi phí hơn. Lý do là, không có đầy đủ những sáng tạo về mặt kỹ thuật để đảm bảo mật độ tham quan cao trong khi những người hiện nay không có khả năng tiếp cận dịch vụ khuyến nông lại có nhu cầu đối với dịch vụ này. Phương pháp tiếp cận bao bọc của chương trình đối với việc cung cấp dịch vụ, dựa chủ yếu vào một phương thức duy nhất (thăm các nông trại) để truyền tải các thông điệp đơn giản chuẩn tắc, cũng làm hạn chế hiệu quả của chương trình. Các chương trình khuyến nông qua radio đang ngày càng phổ biến, các nông dân trẻ có học vấn cao hơn, và những nhà cung cấp khác (các tổ chức phi chính phủ) đang bắt đầu xuất hiện ở nông thôn Kenya. Một cách tiếp cận đa dạng và linh hoạt đối với

cung cấp dịch vụ, đặc biệt là cách tiếp cận sử dụng các phương tiện truyền thông có chi phí thấp, có nhiều khả năng nâng cao được tính hiệu quả về chi phí của chương trình.

Cuối cùng, các kết quả chính cho thấy cần phải cải cách thể chế. Cũng giống như các dịch vụ khác, có thể đạt được hiệu quả cao hơn trong cung ứng các dịch vụ khuyến nông nhờ các thoả thuận thể chế hợp lý hơn. Các thể chế cần tập trung chính vào khách hàng (nông dân). Sự phân quyền trong thiết kế chương trình, kể cả việc sử dụng một cơ chế tham dự cho phép có tiếng nói của nông dân (ví dụ sự chia sẻ chi phí và các tổ chức nông dân) cần phải trở thành một phần không thể tách rời trong cơ chế cung cấp dịch vụ. Sự bền vững về tài chính của dịch vụ khuyến nông là rất quan trọng. Quy mô và mật độ của dịch vụ cần dựa trên chênh lệch hiện có về mặt công nghệ và tri thức và tốc độ của luồng công nghệ mới. Sự hoàn trả chi phí, cho dù chỉ là một phần, có một số ưu điểm: nó cung cấp động cơ phù hợp, giải quyết được các vấn đề về tính trách nhiệm, kiểm soát được chất lượng, làm cho dịch vụ có tính định hướng theo nhu cầu và nhanh nhạy hơn, đồng thời cũng hỗ trợ về mặt tài chính cho dịch vụ này. Các thể chế phân quyền hiện nay vẫn chưa được khai thác ở Kenya, cũng như trong nhiều chương trình khuyến nông ở châu Phi và trên thế giới.

## VI. Chi phí và tổ chức quản lý việc đánh giá

**Chi phí.** Tổng ngân sách dành cho đánh giá là 250.000 đôla, bao gồm chi phí thu thập và xử lý dữ liệu điều tra của hộ gia đình (65.000 đôla- mặc dù số tiền này chỉ là một ước tính thấp hơn chi phí thực tế); chi phí điều tra cán bộ khuyến nông, dữ liệu và báo cáo của nhà tư vấn (12.500 đôla); chi phí thu thập các dữ liệu khác (12.500 đôla); và chi phí cho một nhà phân tích (8.000 đôla). Cần cộng thêm vào chi phí cán bộ khoảng 100.000 đôla (không được phản ánh trong chi phí chính thức) cho các công việc xử lý số liệu, phân tích và viết báo cáo để phản ánh đầy đủ chi phí của nghiên cứu này.

**Tổ chức quản lý.** Để duy trì tính khách quan và tách bạch công việc điều tra ra khỏi cơ quan khuyến nông của chính phủ và Ngân hàng Thế giới, điều tra hộ gia đình do Viện Tegemeo thuộc trường Đại học Egerton, một viện nghiên cứu độc lập ở Kenya thực hiện. Việc phân tích do Madhur Gautam của Ngân hàng Thế giới đảm nhận.

## VII. Các bài học

- Kết hợp giữa đánh giá dựa trên lý thuyết và khung mẫu dựa trên kết quả tạo ra một cơ sở vững chắc để đánh giá tác động của chương trình, đặc biệt là khi có nhiều nhân tố có thể tác động đến kết quả dự kiến. Thiết kế của đánh giá này cung cấp cho việc đo lường các chỉ số cơ bản ở các giai đoạn quan trọng trong vòng đời dự án, liên kết các đầu vào dự án với các kết quả dự kiến nhằm thu thập đầy đủ bằng chứng về tác động.
- Đánh giá thực nghiệm đòi hỏi phải có sự giám sát liên tục và có cường độ cao. Có thể đơn giản hóa đáng kể đánh giá nhờ một hệ thống theo dõi và đánh giá hoạt động tốt và có chất lượng cao, đặc biệt là khi có dữ liệu gốc tốt. Những nguồn lực đầy đủ cho các hoạt động này hiếm khi sẵn có. Đánh giá cũng nhận được lợi ích to lớn từ sự tiếp cận một số dữ liệu trong giai đoạn trước khi có dự án (mặc dù những dữ liệu này cũng hạn chế) và các dữ liệu độc lập để có thể so sánh được.
- Kiểm chứng chéo các kết luận sử dụng các phương pháp phân tích và nguồn dữ liệu khác nhau có vai trò quan trọng để thu thập được những bằng chứng đáng tin cậy. Dữ liệu không hoàn hảo và các khó khăn về thực thi đặt giới hạn cho mức độ tin tưởng mà các phương pháp riêng lẻ trả lời cho các câu hỏi đánh giá then chốt. Các đánh giá định tính và định lượng bổ sung cho nhau một cách mạnh mẽ. Kinh nghiệm từ đánh giá này chỉ ra rằng, ngay cả khi không có đánh giá sự tham dự của người thụ hưởng, thì vẫn có thể đưa những câu

hỏi được thiết kế hợp lý vào trong điều tra để thu thập các thông tin định tính và định lượng. Những thông tin này có thể rất có ích để bổ sung cho đánh giá định lượng.

- Nếu như được áp dụng hợp lý, đánh giá có điều kiện có thể là một công cụ có ích, đặc biệt trong việc đánh giá giá trị của một dịch vụ công cộng hiện đang tồn tại. Các kết quả của việc ứng dụng phương pháp này trong đánh giá đều rất đáng khích lệ, và các câu trả lời có vẻ đều hợp lý.

### VIII. Nguồn

World Bank. 1999. *World Bank Agricultural Extension Projects in Kenya: An Impact Evaluation*. Operations Evaluation Department, Report no. 19523. Washington, D.C.

Ngoài ra, còn có các tài liệu làm việc dưới đây có sẵn ở Phòng đánh giá hoạt động của Ngân hàng Thế giới:

*The Efficacy of the T&V system of Agricultural Extension in Kenya: Results from a Household Survey*

Awareness and Adoption of Extension Messages

Reconsidering the Evidence on Returns to T&V Extension in Kenya

Farmer Efficiency and Productivity Change in Kenya: An Application of the Data Envelopment Analysis

The Willingness to Pay for Extension Services in Kenya: An Application of the Contingent Valuation Method

## Phụ lục 1.9: Tác động của chương trình đào tạo lại của Mêhicô tới việc làm và tiền lương (PROBECAT)

### I. Giới thiệu

Tình huống này có phân bát thường do có đến ba đánh giá về chương trình đã được tiến hành- đánh giá thứ nhất của Ngân hàng Thế giới, sử dụng dữ liệu năm 1992 (Revenga, Riboud, và Tan 1994); đánh giá thứ hai do Bộ Lao động Mêhicô thực hiện, sử dụng dữ liệu năm 1994 (STPS 1995); và đánh giá thứ ba là sự cập nhật kết quả của Ngân hàng Thế giới (Wodon và Minowa 1999). Các phương pháp sử dụng cho hai đánh giá đều tương đối giống nhau, và chúng cho các kết quả tương tự. Việc tăng cường về mặt phương pháp trong đánh giá thứ ba dẫn đến những kết quả và kết luận chính sách tương đối khác. Thực tế là, các kết quả khác nhau đáng kể giữa đánh giá đầu và đánh giá thứ ba, làm nổi bật tầm quan trọng của phương pháp và dữ liệu được sử dụng, và cần thận trọng trong việc diễn giải các kết quả khi tiến hành đánh giá chương trình.

**Mô tả dự án.** PROBECAT (*Programa de Becas de Capacitacion para Trabajadores*) là một chương trình đào tạo ngắn hạn của Mêhicô có mục tiêu nâng cao thu nhập và việc làm cho những người thất nghiệp và bị sa thải. PROBECAT do các văn phòng việc làm của nhà nước thực hiện. Những người được đào tạo nhận được mức tiền lương tối thiểu trong thời kỳ đào tạo kéo dài từ một đến sáu tháng, và sau đó văn phòng việc làm địa phương sẽ tìm kiếm việc làm cho họ. Ban đầu, chương trình này có quy mô nhỏ (khoảng 50.000 người tham dự), nhưng trong những năm gần đây, quy mô chương trình đã tăng lên rất nhiều, và đáp ứng hơn 500.000 người mỗi năm.

**Các đặc điểm nổi bật của các đánh giá.** Các đặc điểm nổi bật như sau:

- Đánh giá năm 1994 lý thú vì bốn lý do: sử dụng dữ liệu hiện có một cách sáng tạo; xây dựng một nhóm đối chứng tương hợp; thừa nhận chính thức tính nhiều mặt của các kết quả chương trình, đặc biệt là với các nhóm công nhân khác nhau; và phân tích rõ ràng về chi phí- lợi ích. Các kết quả của đánh giá này tương đối tích cực khi xét đến tác động của chương trình tới những người thụ hưởng.
- Đánh giá năm 1995 là sự lặp lại phương pháp luận của đánh giá năm 1994, sử dụng các dữ liệu cập nhật hơn. Các kết quả cũng có tính tích cực khi xem xét tác động của chương trình. Do thiết kế và kết quả của đánh giá năm 1995 tương tự với đánh giá năm 1994, nên dưới đây chúng tôi sẽ không thảo luận về đánh giá năm 1995 nữa.
- Đánh giá năm 1999 được thực hiện như là một phần trong đánh giá tình trạng đói nghèo ở Mêhico, sử dụng cùng dữ liệu như đánh giá năm 1995, nhưng bằng một phương pháp kinh tế lượng khác. Các đơn vị đối chiếu được sử dụng để kiểm soát tính nội sinh của biến tham dự chương trình cho thấy, không có tác động của chương trình tới khả năng có việc làm và tiền lương sau đào tạo. Mặc dù điều này không có nghĩa là chương trình không có lợi ích, nó cho thấy, chương trình đóng vai trò như một mạng lưới an toàn tạm thời cho những người thất nghiệp hơn là một chương trình đào tạo việc làm.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Trong đánh giá năm 1994, các tác giả ước lượng tác động của đào tạo tới (a) xác suất có việc làm sau 3, 6 và 12 tháng; (b) thời gian để ra khỏi tình trạng thất nghiệp; (c) tác động tới thu nhập hàng tháng, số giờ làm việc mỗi tuần và tiền lương theo giờ ; và (d) lợi tức của đầu tư.

Đánh giá năm 1999 xem xét cùng những câu hỏi này, trừ số

giờ làm việc trong tuần và tiền lương mỗi giờ. Do đánh giá này nhận thấy không có tác động của chương trình tới việc làm và thu nhập hàng tháng, nên lợi tức của đầu tư cũng bằng không. Tuy vậy, chương trình này vẫn có thể đóng vai trò như một mạng lưới an toàn.

Thiết kế của cả hai đánh giá đều có tính sáng tạo trong việc tạo dựng nhóm đối chứng. Trong cả hai trường hợp, đánh giá kết hợp giữa điều tra lực lượng lao động theo nhóm lặp, được gọi là Encuesta Nacional de Empleo (ENEU), với việc điều tra một nhóm lặp bao gồm những người được đào tạo trong cùng thời kỳ. Như vậy, tiêu chí lựa chọn của chương trình đã được sử dụng để xác định nhóm đối chứng từ ENEU. Mặc dù không có cách nào thay cho việc kết hợp các cuộc điều tra do hạn chế về dữ liệu, có thể phê phán việc xây dựng mẫu chung (nhóm đối chứng và nhóm tham dự) như đã được thảo luận trong đánh giá năm 1999 như sau:

- Trong việc sử dụng những cá nhân thất nghiệp trong ENEU để tạo thành nhóm đối chứng, người ta giả định không có cá nhân nào trong ENEU được hưởng lợi từ chương trình. Giả định này không đúng, do mỗi cá nhân trong ENEU đều có xác suất tham dự PROBECAT nhất định. Thật may mắn là, do chương trình có quy mô nhỏ cho tới năm 1993, chỉ có một số rất nhỏ các cá nhân trong nhóm đối chứng là có thể tham dự chương trình (dữ liệu cho đánh giá năm 1999 là dữ liệu cho 1993-94);
- Việc kết hợp hai mẫu ngẫu nhiên (những người được đào tạo trong PROBECAT và những cá nhân thất nghiệp trong ENEU) không phải là một mẫu ngẫu nhiên, do đó, khi không có những đặc điểm tiêu chuẩn của phần dư thì các kết quả hồi quy có thể không cho các ước lượng có tham số nhất quán, đặc biệt là vì mô hình được sử dụng nhạy cảm với giả định về phân phối chuẩn hai biến. Khi không có dữ liệu tốt hơn thì không thể làm được gi nhiều trong trường hợp này.

Sự khác nhau chủ yếu giữa các đánh giá năm 1994 và 1999 là như sau:

- Trong đánh giá năm 1994, các tác giả cố gắng giải quyết vấn đề sai lệch lựa chọn xảy ra do sự lựa chọn phi ngẫu nhiên những người được đào tạo của PROBECAT bằng cách ước lượng một mô hình probit về xác suất tham dự chương trình. Nhóm so sánh được giới hạn là những cá nhân có nhiều khả năng tham dự. Trong đánh giá năm 1999, các tác giả cho rằng, phương pháp này không loại trừ hết vấn đề nội sinh. Do đó, họ sử dụng một biến công cụ để kiểm soát tính nội sinh của sự tham dự chương trình.
- Khi ước lượng thu nhập trong đánh giá năm 1994, trong khi kiểm soát được việc tham dự PROBECAT, thì đánh giá vẫn chưa tính đến sai lệch do lựa chọn mẫu phát sinh từ quyết định đi làm. Trong nghiên cứu năm 1999, cả hai vấn đề trong lựa chọn mẫu này đều được quan tâm đến.

### III. Dữ liệu

Trong đánh giá năm 1994, dữ liệu về những người được đào tạo được thu thập từ một cuộc điều tra hồi tưởng tiến hành với 881 đàn ông và 845 phụ nữ, những người này được đào tạo vào năm 1990. Dữ liệu này được bổ sung bằng dữ liệu lặp về 371 đàn ông và 189 phụ nữ được rút ra từ một đợt điều tra hộ gia đình đối với 16 khu vực thành thị chính ở Mêhicô. Điều tra này là một phần trong điều tra về lực lượng lao động định kỳ hàng quý có tên là ENEU, do cơ quan thống kê Mêhicô tiến hành. Các tác giả đã khai thác cấu trúc nhóm quay vòng của điều tra để điều tra về những công nhân bị thất nghiệp trong quý ba năm 1990, và sau đó theo dõi những công nhân này trong một năm. Nhóm này được bổ sung bằng một nhóm thất nghiệp trong quý tư năm 1990, và nhóm này được theo dõi trong chín tháng nữa. Phương pháp này cũng được sử dụng trong đánh giá năm 1999, nhưng với các dữ liệu cập nhật hơn.

#### IV. Các kỹ thuật kinh tế lượng

Các kỹ thuật kinh tế lượng quan trọng được sử dụng là phân tích tồn tại (các mô hình thời kỳ) cho xác suất làm việc và hồi quy Heckman cho tiền lương. Sau đây là các kỹ thuật kinh tế lượng được sử dụng trong đánh giá năm 1999. Những điểm khác biệt với đánh giá năm 1994 được nêu bật.

**Tác động của PROBECAT tới khoảng thời gian tìm kiếm việc làm.** Trong phân tích tồn tại (survival analysis), hàm tồn tại  $S(t)$  thể hiện khoảng thời gian thất nghiệp sau khi được đào tạo (đo bằng tháng). Với  $S(t)$  cho trước, hàm rủi ro  $\lambda(t)$  chỉ khả năng bị thất nghiệp (hay rủi ro tiếp tục bị thất nghiệp) ở thời điểm  $t$  trong số những cá nhân hiện không có việc làm vào thời điểm này bằng  $\lambda(t) = -d(\log S(t))/dt$ . Có thể mô tả đường cong tồn tại như là một hàm số của sự tham dự chương trình P, các đặc điểm cá nhân X, các đặc điểm của nhà nước Z, do đó  $\lambda = \lambda(t; X, Z, P)$ . Trong mô hình rủi ro tỷ lệ của Cox, nếu  $i$  được dùng để chỉ hộ gia đình và  $j$  để chỉ khu vực mà hộ gia đình này sinh sống, thì chúng ta có

$$\lambda(t; X, Z, P1, P2) = \lambda_0(t) \exp(\gamma' X_{ij} + \delta' Z_j + \mu P_{ij}). \quad (1)$$

Cox đề xuất ước lượng khả năng tối đa (maximum likelihood estimation) mô hình này, trong đó hàm cơ sở  $\lambda_0(t)$  không cần thiết phải được xác định. Nếu  $\mu$  dương và có ý nghĩa thống kê thì chương trình có tác động tích cực tới việc làm. Có thể mô tả sự khác biệt giữa các đánh giá năm 1994 và năm 1996 như sau:

- Trong đánh giá năm 1994, các tác giả hồi quy theo hàm probit đối với sự tham dự chương trình và loại trừ khỏi nhóm đối chứng các cá nhân có xác suất tham dự chương trình thấp. Sau đó họ chạy phương trình (1) mà không tiếp tục kiểm soát đối với hiện tượng nội sinh.
- Trong đánh giá năm 1999, các tác giả cũng chạy hồi quy probit cho sự tham dự chương trình, nhưng sử dụng mức độ sẵn

sàng của chương trình ở cấp địa phương (có được từ dữ liệu hành chính) như là một nhân tố quyết định bổ sung cho sự tham dự (nhưng không phải là nhân tố quyết định kết quả với điều kiện có sự tham dự của cá nhân.) Sau đó họ chạy phương trình (1), không phải với giá trị thực sự của biến tham dự, mà là với giá trị dự đoán (chỉ số) có được từ phương trình probit giai đoạn một. Đây là một thủ tục biến công cụ. ý tưởng này tiếp nối các công trình về đánh giá chương trình sử dụng đặc tính phân quyền của Ravallion và Wodon (2000) và Cord và Wodon (1999). Các tác giả so sánh những kết quả của mình với các phương pháp khác, chỉ ra rằng, các phương pháp khác có sai lệch trong giá trị của các ước lượng tham số do chưa kiểm soát đầy đủ đối với hiện tượng nội sinh.

**Tác động của PROBECAT tới thu nhập hàng tháng.** Để thực hiện phân tích này, người ta sử dụng một mô hình có kiểm soát đối với sự lựa chọn mẫu trong lực lượng lao động và sự tham dự chương trình trong đánh giá năm 1999 (đánh giá năm 1994 chỉ kiểm soát sự tham dự chương trình). Chúng ta ký hiệu logarit của tiền lương dự kiến của một cá nhân là  $\log w$ . Tiền lương này không bằng không khi và chỉ khi nó lớn hơn tiền lương dự trữ của cá nhân (nếu không, cá nhân này sẽ lựa chọn không làm việc). Chúng ta ký hiệu chênh lệch không quan sát được giữa tiền lương dự kiến của một người và mức tiền lương dự trữ của anh ta là  $\Delta^*$ . Tiền lương dự kiến của người này được xác định bởi một số biến riêng lẻ (vectơ  $E$ , chủ yếu gồm trình độ giáo dục và kinh nghiệm quá khứ của cá nhân này) và các biến địa điểm  $Z$ , cộng với sự tham dự chương trình  $P$ . Chênh lệch giữa tiền lương dự kiến của một người với tiền lương dự trữ được xác định cũng bởi những biến tương tự trên, cộng thêm số trẻ em, việc gia đình này có phải là chủ gia đình hay không, và việc người này đã lập gia đình chưa. Mô hình do vậy là:

$\Delta_{ij}^* = \phi_\Delta E_{ij} + \pi_\Delta D_{ij} + \eta_\Delta Z_j + \alpha_\Delta P_{ij} + v_{ij}$  with  $\Delta_{ij} = 1$  if  $\Delta_{ij}^* > 0$ , và  
bằng 0 nếu  $\Delta_{ij}^* < 0$  (2)

$\text{Log } w_{ij}^* = \phi_w E_{ij} + \eta_w Z_j + \alpha_w P + \kappa_{ij}$  với  $\text{Log } w = \log w^*$  nếu  $\Delta = 1$   
và bằng 0 nếu  $\Delta = 0$ . (3)

Cũng như trong mô hình tồn tại, để kiểm soát tính nội sinh của sự tham dự chương trình trong đánh giá năm 1999, có thể ước lượng trước hết hàm probit cho sự tham dự chương trình bằng cách sử dụng sự sẵn có của chương trình ở cấp độ địa phương như là nhân tố quyết định sự tham dự của cá nhân. Sau đó phương trình trên được ước lượng bằng cách sử dụng giá trị dự đoán (chỉ số) của sự tham dự chương trình thay cho các giá trị thực sự. Trong đánh giá năm 1994, mô hình này không kiểm soát đối với quyết định tham gia thị trường lao động trong phương trình (2) ở trên. Phương trình này được thay thế bằng hàm probit về sự tham dự chương trình được ước lượng mà không có sự sẵn có của chương trình đóng vai trò biến độc lập. Việc so sánh các mô hình khác nhau một lần nữa cho thấy tồn tại sai lệch khi không sử dụng kỹ thuật biến số công cụ.

## V. Ai là người tiến hành

Đánh giá năm 1994 do Ana Revenga thuộc Phòng Quốc gia châu Mỹ Latinh và vùng Caribê II của Ngân hàng Thế giới, Michelle Riboud thuộc Phòng Quốc gia châu Âu và Trung Á IV của Ngân hàng Thế giới, và Hong Tan thuộc Phòng Phát triển khu vực tư nhân của Ngân hàng Thế giới tiến hành. Đánh giá năm 1999 do Quentin Wodon và Mari Minowa, cũng ở Ngân hàng Thế giới (khu vực Mỹ Latinh) tiến hành.

## VI. Kết quả

Kết quả có được từ những đánh giá này rất khác nhau. Đánh

giá năm 1994 và 1995 nhận thấy có tác động tích cực của chương trình tới việc làm và tiền lương. Đánh giá năm 1999 không nhận thấy tác động tích cực nào trong khi đánh giá này dựa trên cùng bộ dữ liệu sử dụng cho đánh giá năm 1995. Về khía cạnh phân tích chi phí- lợi ích, hai đánh giá đầu tiên cho thấy kết quả tích cực, trong khi đánh giá thứ ba không cho thấy kết quả như vậy. Các kết quả gây thất vọng trong đánh giá thứ ba không phải là điều đáng ngạc nhiên. Hầu hết các chương trình đào tạo lại của các nước trong Tổ chức Hợp tác và phát triển kinh tế (OECD) đều có những tác động hạn chế, và nếu như các chương trình có một số tác động nhất định, thì những tác động này có xu hướng mất đi sau một vài năm (Dar và Gill 1998). Việc PROBECAT có thể không có lợi trong trung hạn và dài hạn đối với những người tham dự chương trình như trong kết luận của đánh giá cuối cùng, không có nghĩa là cần phải bãi bỏ chương trình này. Có thể xem chương trình này như là sự cung cấp mạng lưới an toàn tạm thời ( thông qua khoản thu nhập ở mức tiền lương tối thiểu) thay vì như một chương trình đào tạo. Hoặc là có thể cải thiện chương trình để cung cấp hoạt động đào tạo có những tác động lâu dài hơn.

## VII. Các bài học

Ngoài một số đặc điểm sáng tạo và những hạn chế trong các đánh giá này, bài học quan trọng ở đây là chúng ta cần phải cẩn trọng khi tiến hành đánh giá chương trình và sử dụng kết quả thu được để đưa ra khuyến nghị chính sách. Cần phải luôn nhớ đến việc một đánh giá về sau có thể mâu thuẫn với đánh giá trước đó khi sử dụng các kỹ thuật kinh tế lượng khác nhau. Có rất nhiều trường hợp như vậy đã được nêu ra trong các tài liệu nghiên cứu.

### VIII. Nguồn

- Revenga, Ana, Michelle Riboud, and Hong Tan. 1994. "The Impact of Mexico's Retraining Program on Employment and Wages." *World Bank Economic Review* 8 (2): 247- 77.
- Wodon, Quentin, and Mari Minowa. "Training for the Urban Unemployed: A Reevaluation of Mexico's PROBECAT." World Bank, Government Programs and Poverty in Mexico, Report No. 19214-ME, Vol. II.

## **Phụ lục 1.10: Mêhicô, Chương trình Giáo dục, Y tế, và Dinh dưỡng quốc gia (PROGRESA): Đề xuất đánh giá**

### **I. Giới thiệu**

**Mô tả dự án.** PROGRESA là một chương trình đa lĩnh vực nhằm đấu tranh chống tình trạng đói nghèo tuyệt đối ở Mêhicô bằng cách cung cấp kết hợp các dịch vụ y tế, dinh dưỡng, và giáo dục cho các gia đình nghèo. Chính phủ Mêhicô sẽ hỗ trợ tài chính, bổ sung dinh dưỡng, trợ cấp giáo dục, và cung cấp các dịch vụ y tế cơ bản trong ít nhất ba năm liên tục. Chính phủ dự định mở rộng phạm vi PROGRESA từ quy mô hiện tại gồm 400.000 gia đình thành 1,5 triệu gia đình vào cuối năm 1998, với một khoản chi tiêu 500 triệu đô la.

**Các đặc điểm nổi bật của đánh giá.** Đánh giá này đặc biệt phức tạp vì nó đánh giá cả ba khía cạnh của chương trình: hoạt động, tính hiệu quả trong định hướng đối tượng, và tác động. Đánh giá càng trở nên phức tạp hơn do bản thân các kết quả cũng có tính đa chiều. Do vậy, có nhiều thành phần đánh giá khác nhau: lựa chọn đối tượng thụ hưởng, phương pháp đánh giá, khung phân tích không mang tính thử nghiệm, yêu cầu dữ liệu, tác động tới giáo dục, tác động tới sức khoẻ, tác động tới tiêu dùng thực phẩm và dinh dưỡng, tác động tới chi tiêu cho tiêu dùng và phân bổ nguồn lực trong phạm vi hộ gia đình, tác động tiềm tàng trong vòng thứ hai của chương trình, mô phỏng các thay đổi về lợi ích của chương trình, và các vấn đề về chi phí - hiệu quả và chi phí - lợi ích.

Mặc dù đánh giá này là nội dung phác thảo các ý tưởng hơn là kết quả thực hiện, nhưng có một bài học quan trọng từ đánh giá là, cách thức suy nghĩ và cấu trúc đánh giá trước khi thực

sự tiến hành đánh giá. Đặc biệt, đây là, một phác thảo rất có ích về những vấn đề khái niệm và thực nghiệm cần phải giải quyết trong đánh giá, cũng như cách thức giải quyết các vấn đề này. Một thành phần quan trọng của đánh giá là bề rộng của nó: thay vì chỉ đánh giá đơn giản tác động của chương trình, đánh giá này sẽ giúp chỉ ra rằng kết quả có được là do thành công hay thất bại trong hoạt động và định hướng đối tượng của chương trình.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Các câu hỏi nghiên cứu cơ bản là đánh giá ba khía cạnh trong hoạt động của PROGRESA- các khía cạnh hoạt động, sự định hướng đối tượng, và tác động. Khía cạnh hoạt động của chương trình thường bị bỏ qua mặc dù có thể chuyển các thất bại thành thành công nếu như tiến hành các biện pháp điều chỉnh. Cũng có thể đưa ra một luận điểm tương tự về tính định hướng đối tượng của chương trình: một chương trình có thể thất bại chỉ đơn giản vì định hướng đối tượng kém, chứ không phải vì bản thân chương trình có thiếu sót. Đánh giá tác động có tính chuẩn tắc hơn, mặc dù ngay cả mục tiêu này cũng khá tham vọng vì nó phân tích cả quy mô của tác động và cách thức đạt được tác động này.

Việc theo dõi hoạt động của chương trình là một thủ tục hai bước. Nhóm nghiên cứu xây dựng một lược đồ về trình tự các bước của chương trình. Nhóm này sau đó sử dụng các quan sát, phỏng vấn, nhóm tập trung, và hội thảo với những người liên quan nhằm đánh giá, phân tích, và có thể thay đổi các tiến trình của chương trình.

Một phương pháp hai bước cũng được sử dụng để định hướng đối tượng tới các hộ gia đình trong PROGRESA. Bước đầu tiên là xác định địa phương nào trong khu vực đủ điều kiện nhận được PROGRESA bằng cách sử dụng một chỉ số căn cứ vào mức độ nghèo. Bước thứ hai là xác định xem một gia đình

nào đó trong địa phương này có đủ điều kiện tham dự không, căn cứ vào trao đổi giữa các cán bộ của PROGRESA với những người lãnh đạo địa phương. Nghiên cứu này sẽ giải đáp tính hợp lệ của định hướng đối tượng bằng cách (a) so sánh phân phối mức tiêu dùng của các hộ gia đình trong các hộ gia đình tham dự và không tham dự chương trình tại các địa phương có chương trình, (b) đưa ra một chỉ tiêu ngưỡng về tiêu dùng của hộ gia đình phù hợp với tổng số hộ gia đình mà PROGRESA có thể phục vụ, (c) tiến hành phân tích độ nhạy cảm và đặc trưng của các hộ gia đình tham dự và không tham dự PROGRESA so với các hộ gia đình được lựa chọn và không được lựa chọn khi sử dụng chỉ tiêu ngưỡng này, (d) khai thác khả năng sử dụng tiêu chí hiện tại để dự đoán tiêu dùng, (e) xác định tiêu chí khác nhau căn cứ từ những nguồn dữ liệu khác nhau, và (f) mô phỏng các mô hình có thể cải thiện sự định hướng đối tượng bằng cách sử dụng những tiêu chí khác nhau (Viện Nghiên cứu chính sách lương thực quốc tế 1998, trang 6).

Đối với đánh giá tác động, cần sử dụng cùng hệ thống này, với kết quả là người ta đã tiến hành phân công ngẫu nhiên các địa phương thành 296 nhóm tham dự và 173 nhóm không tham dự, trong đó nhóm thứ nhất có 14.382 gia đình và nhóm thứ hai có 9.202 gia đình. Các gia đình đủ tư cách tham dự ở trong nhóm đối chứng sẽ được tham dự sau ít nhất là một năm.

Các nhà tư vấn lập kế hoạch kiểm định phi ngẫu nhiên có thể có của dự án bằng cách so sánh đặc điểm của nhóm tham dự và nhóm đối chứng. Nếu các nhóm này khác nhau một cách có hệ thống, thì sau đó người ta sẽ sử dụng ba phương pháp phi thử nghiệm là phương pháp hàm kiểm soát, phương pháp tương hợp, và phương pháp hồi quy.

### III. Dữ liệu

Thành phần dữ liệu hoạt động có được từ các quan sát, phỏng vấn, nhóm tập trung, và hội thảo với các bên liên quan. Đánh

giá tập trung chủ yếu vào việc xác định những gì đang diễn ra và tại sao lại diễn ra, mức độ hài lòng với tiến trình thực hiện và các đề xuất để cải thiện. Các dữ liệu này được thu thập ở các địa phương, và cũng căn cứ đáng kể vào các tài liệu hành chính nội bộ của PROGRESA.

Có hai đợt điều tra đã được tiến hành: điều tra dân số vào tháng 12 năm 1997, và điều tra ban đầu vào tháng Ba năm 1998. Tiêu chí chủ yếu để lựa chọn đối tượng là tiêu dùng của hộ gia đình, và mặc dù dữ liệu này không được thu thập trong điều tra dân số, nó đã được thu thập sau đấy trong đợt điều tra vào tháng Ba. Tuy nhiên, biến số này thiếu thông tin về tự tiêu dùng (self-consumption), và mặc dù số liệu này sẽ được thu thập về sau, nhưng nó sẽ bị tác động do việc thực hiện PROGRESA. Các nhà tư vấn dự định làm việc hoàn toàn với những hộ gia đình đủ điều kiện và không đủ điều kiện trong địa phương đóng vai trò so sánh.

Dánh giá tác động dựa trên sự lựa chọn các chỉ số tác động. Chương trình PROGRESA có tác động tới cả chất lượng và số lượng các dịch vụ cung cấp và sự đầu tư vào sức khoẻ, dinh dưỡng, và giáo dục. Nhiều chỉ số đánh giá được đưa ra, căn cứ vào một số kết quả tác động, và mỗi chỉ số lại có một nguồn dữ liệu đi kèm. Phúc lợi hộ gia đình, được đo lường theo tiêu dùng, tiết kiệm, tích luỹ các hàng hoá lâu bền của hộ gia đình, sẽ được đo lường trong các đợt điều tra gốc và điều tra tiếp theo; tình trạng dinh dưỡng và sức khoẻ của trẻ em sẽ được đo lường trong các kỳ điều tra gốc và điều tra tiếp theo đối với một mẫu nhỏ về dinh dưỡng; việc đi học sẽ được nghiên cứu trong một cuộc điều tra ở cấp độ trường học, và thông qua các cuộc điều tra gốc và điều tra tiếp theo; các trang thiết bị y tế có thể được theo dõi thông qua sổ sách của các trung tâm y tế và nhờ điều tra; và tình trạng của phụ nữ có thể được xác định nhờ các cuộc điều tra và kiểm tra với những người liên quan.

Một đặc điểm rất hấp dẫn của đánh giá dự kiến này là sử dụng phương pháp phân tích để đánh giá các thước đo kết quả

hiện tại và thảo luận rộng rãi về các chỉ tiêu kết quả và kiểm soát phù hợp hơn đối với giáo dục, sức khỏe, và tiêu dùng.

Đánh giá dự kiến bao gồm một phân tích về chi phí-lợi ích. Những lợi ích đã được xác định, mặc dù có những khó khăn luôn tồn tại trong việc tiền tệ hoá chất lượng cuộc sống, và sự gia tăng quyền lực. Hai dạng chi phí khác nhau cũng được xác định là chi phí thực hiện chương trình và chi phí chương trình. Chi phí thứ nhất bao gồm các chi phí sàng lọc, chi phí cung cấp cho đối tượng, và chi phí giám sát; chi phí thứ hai bao gồm cả thu nhập bỏ lỡ.

#### IV. Kỹ thuật kinh tế lượng

Các kỹ thuật kinh tế lượng được sử dụng phụ thuộc vào mối quan hệ nào sẽ được ước lượng. Các nhà tư vấn thảo luận về sự phù hợp của mối quan hệ trong hàm sản xuất (chẳng hạn, về thành tích học tập), mối quan hệ với nhu cầu (chẳng hạn, về các dịch vụ y tế hay giáo dục), và mối quan hệ nhu cầu có điều kiện (trong đó một số biến được gia đình quyết định thay vì bởi cá nhân).

Kỹ thuật kinh tế lượng lý thú nhất được sử dụng trong ước lượng hàm chi tiêu Working-Leser có dạng sau:

$$W_j = \alpha_I + \beta_{1j} \ln pce_{exp} + \beta_{2j} \ln size + \sum_k \delta_{kj} dem_k + \sum_s \Theta_{sj} z_s + \beta_{3j} P + e_j$$

trong đó  $W_j$  là phần ngân sách của hàng hóa thứ  $j$ ;  $\ln pce_{exp}$  là log của tổng chi tiêu trên đầu người;  $\ln size$  là log của quy mô hộ gia đình;  $dem_k$  là tỷ lệ của nhóm dân số  $k$  trong hộ gia đình;  $z_s$  là vectơ của biến giả tác động tới địa điểm của hộ gia đình;  $P$  cho biết sự tham dự PROGRESA; và  $e_j$  là hạng tử sai số.

Phương pháp này có nhiều ưu điểm: nó cho phép có sự tham gia của các nhân tố kiểm soát; nó thỏa mãn hạn chế cộng thêm (adding-up constraint); và nó được sử dụng rộng rãi, nhờ đó cho phép có sự so sánh với các nghiên cứu khác. Cuối cùng, có

thể sử dụng mô hình này để xác định những cách thức khác nhau mà qua đó, PROGRESA có thể ảnh hưởng tới chi tiêu: thông qua thay đổi nguồn lực của hộ gia đình ( $\beta_{lj}$  nhân với xu hướng tiêu dùng cận biên, được ước lượng riêng rẽ), thông qua thay đổi trong phân phối thu nhập (bằng cách điều chỉnh nó để đưa vào tỷ lệ phụ nữ là người lớn trong hộ gia đình), và thông qua tác động tạo ra sự tham dự lớn hơn. Điều tra cơ sở và điều tra tiếp theo cho phép sử dụng phương pháp khác biệt trong khác biệt.

Các nhà tư vấn cũng phát hiện ra những vấn đề kinh tế lượng quan trọng có thể gặp phải: hiện tượng đa cộng tuyến, sai số đo lường, các biến bị bỏ sót, hiện tượng đồng thời (simultaneity), và xác định khung thời gian để tác động có thể trở nên nhận thấy được.

## V. Ai sẽ tiến hành

Các cán bộ của Viện nghiên cứu chính sách lương thực quốc tế tiến hành, gồm Gaurav Datt, Lawrence Haddad, John Hoddinott, Agnes Quisumbing và Marie Rudei. Nhóm nghiên cứu gồm Jere Behrman, Paul Gerler và Paul Schultz.

## VI. Các bài học

Bài học đầu tiên ở đây là giá trị của việc xác định được các vấn đề đánh giá, phương pháp luận, và nguồn dữ liệu - và đánh giá có tính phê phán về đánh giá - trước khi đánh giá diễn ra. Phác thảo nội dung thảo luận này rất có giá trị trong việc đưa ra một ví dụ minh họa cho tất cả các khó khăn và cạm bẫy mà người đánh giá có thể gặp phải. Cụ thể là, một số vấn đề thường xảy ra trong đánh giá tác động đã được xác định, đó là (a) các thay đổi chính sách có thể khó dự đoán do hiện tượng thay thế chéo và điều chỉnh hành vi; (b) lợi ích biên và chi phí biên phụ thuộc vào một số nhân tố: ngoại ứng (tạo ra sự chênh lệch giữa giá

trị xã hội và giá trị tư nhân), các tác nhân liên quan (cha mẹ so với trẻ em); (c) tầm quan trọng của các đặc điểm không quan sát được; (d) tầm quan trọng của việc kiểm soát các đặc điểm cá nhân, gia đình, và cộng đồng, và (e) các ước lượng thực nghiệm phụ thuộc vào các đặc điểm kinh tế vĩ mô, thị trường, môi trường chính sách và điều tiết cho trước.

### VIII. Nguồn

International Food Policy Research Institute. 1998. *Programa Nacional de Educación, Salud, y Alimentación (PROGRESA): A Proposal for Evaluation* (Với phụ lục kỹ thuật kèm theo.). Washington, D.C.: IFPRI.

## Phụ lục 1.11: Đánh giá về cải cách trường học ở Nicaragua: Phương pháp kết hợp định lượng và định tính

### I. Giới thiệu

**Mô tả dự án.** Vào năm 1991, chính phủ Nicaragua đã triển khai chương trình cải cách mạnh mẽ hệ thống giáo dục công cộng của mình. Tiến trình cải cách đã phân quyền quản lý trường học (các quyết định về chi tiêu, ngân sách, chương trình học, và phương pháp sư phạm) và chuyển giao trách nhiệm tài trợ cho cấp địa phương.

Các cải cách được tiến hành thành nhiều giai đoạn, bắt đầu bằng một nghị định năm 1991, thiết lập các hội đồng cha mẹ học sinh tại tất cả các trường quốc lập. Sau đó, một chương trình thí điểm được tiến hành vào năm 1993 tại 20 trường trung học được lựa chọn. Chương trình này thành lập các hội đồng cha mẹ học sinh trong các hội đồng quản lý nhà trường với trách nhiệm lớn hơn về nhân lực, ngân sách, chương trình học, và tính sư phạm. Tới năm 1995, các hội đồng quản lý nhà trường hoạt động tại 100 trường trung học và hơn 300 trường tiểu học; các trường này tham dự chương trình thông qua một tiến trình tự lựa chọn bao gồm cả đơn xin tham dự của các giáo viên và hiệu trưởng nhà trường. Quyền tự chủ trường học dự kiến sẽ được áp dụng hầu như toàn bộ vào cuối năm 1999.

Mục tiêu của cải cách ở Nicaragua là tăng cường việc học tập của học sinh bằng cách thay đổi các tiến trình tổ chức trong các trường quốc lập, nhờ đó, việc ra quyết định tại các trường này mang lại lợi ích cho học sinh như là một sự ưu tiên cao nhất. Do việc quản lý nhà trường ngày càng trở nên dân chủ, với sự tham dự rộng rãi hơn và nguồn tài chính tạo ra tại địa phương tăng lên, nên cách thức chi tiêu cũng trở nên hợp lý hơn và

được phân bổ nhằm cải thiện trực tiếp phương pháp sư phạm và tăng cường kết quả của học sinh.

**Đánh giá tác động.** Đánh giá Cải cách tự chủ trường học của Nicaragua là một trong những nỗ lực có hệ thống đầu tiên nhằm đánh giá tác động của sự phân quyền ở trường học tới kết quả học tập của học sinh. Đánh giá này, do Ngân hàng Thế giới và Bộ giáo dục đồng tiến hành, được khởi đầu năm 1995, và sẽ hoàn thành vào cuối năm 1999. Thiết kế đánh giá có tính sáng tạo do nó kết hợp các phương pháp định tính và định lượng, trong đó thành phần định lượng có tính đặc trưng vì bao gồm một phần riêng biệt đánh giá tiến trình ra quyết định ở trường học. Đánh giá cũng minh họa các kỹ thuật "tốt nhất" khi không có dữ liệu gốc và khi việc tiến hành cải cách có lựa chọn (không phải là ngẫu nhiên) đã loại trừ khả năng sử dụng thiết kế đánh giá thử nghiệm.

Mục đích của thành phần định tính trong đánh giá là minh họa liệu thực sự có cải cách trong quản lý và tài trợ như dự kiến tại nhà trường không, và để đánh giá việc các bên liên quan nhìn nhận như thế nào về tiến trình cải cách. Thành phần định lượng bổ sung cho các kết quả này bằng cách trả lời câu hỏi sau: "Liệu các thay đổi trong quản lý và tài trợ của nhà trường có thực sự tạo ra các kết quả học tập tốt hơn cho trẻ em không?" Các kết quả định tính cho thấy, việc thực hiện thành công cải cách phụ thuộc lớn vào bối cảnh và môi trường của nhà trường (ví dụ, mức độ đói nghèo của cộng đồng dân cư), trong khi đó, các kết quả định lượng cho thấy sự gia tăng quyền quyết định của nhà trường trên thực tế đã đi kèm với kết quả học tập của học sinh tốt hơn.

## II. Thiết kế đánh giá

Thiết kế của đánh giá Cải cách quyền tự chủ trường học ở Nicaragua dựa trên "kỹ thuật so sánh tương hợp," trong đó dữ liệu của một mẫu đại diện gồm các trường học tham dự vào

tiến trình cải cách được so sánh với dữ liệu từ một mẫu gồm các trường không tham dự. Mẫu các trường không tham dự được lựa chọn càng giống với những đặc điểm của các trường tham dự càng tốt, và vì vậy tạo ra sự tương phản thực tế. Thiết kế này được lựa chọn vì việc thiếu các dữ liệu gốc đã loại trừ khả năng sử dụng kỹ thuật đánh giá so sánh trước và sau khi có chương trình, và vì các cải cách không được áp dụng một cách ngẫu nhiên cho các trường học nên không thể sử dụng thiết kế đánh giá thử nghiệm (trong đó mẫu các trường được nghiên cứu trong đánh giá sẽ có tính ngẫu nhiên, và vì vậy, có tính đại diện trên toàn quốc).

### III. Các kỹ thuật thu thập dữ liệu và phân tích

Nghiên cứu định lượng căn cứ vào dữ liệu của một mẫu bao gồm 12 trường học, 9 trường tham dự cải cách và 3 trường không cải cách, đại diện cho nhóm đối chứng (Thực ra dữ liệu được thu thập cho 18 trường nhưng chỉ có 12 trường trong số này có mặt trong nghiên cứu định tính do có sự trì hoãn trong việc chuẩn bị bảng điểm và do quyết định tập trung phân tích vào các trường cải cách- những trường cung cấp các tài liệu phù hợp hơn cho phân tích này.) Mẫu 12 trường này được lựa chọn để đại diện cho cả các trường tiểu học và trường trung học, ở nông thôn và ở thành thị, và căn cứ vào dữ liệu từ điều tra định lượng năm 1995, các trường này đại diện cho những mức độ tự chủ khi ra quyết định khác nhau. Có tổng số 82 cuộc phỏng vấn và phiên họp nhóm tập trung đã được tiến hành, tập trung vào việc tìm hiểu xem các hiệu trưởng, thành viên hội đồng nhà trường, cha mẹ học sinh, và giáo viên hiểu và nhìn nhận như thế nào đối với tiến trình phân quyền. Tất cả các cuộc phỏng vấn đều được thực hiện bởi những người Nicaragua bản địa, đã được huấn luyện bằng các mô phỏng phỏng vấn và kiểm tra thử để có thể sử dụng một loạt các câu hỏi có tính định hướng mà không gợi ý câu trả lời cho người được phỏng vấn. Các cuộc phỏng vấn được ghi âm lại, chuyển thể thành văn bản, và sau

đó được tóm tắt trong khoảng từ hai đến bốn trang văn bản. Văn bản ghi lại này sau đó sẽ được phân tích để tìm ra những bằng chứng riêng rẽ và các chủ đề cơ bản này sinh giữa các trường học, người tham dự và giữa các trường cải cách với nhóm đối chứng.

Dữ liệu định lượng được thu thập bao gồm hai phần: điều tra lặp về các trường học được tiến hành trong hai đợt (tháng 11-12 năm 1995 và tháng 4-8 năm 1997) và các kỳ thi của học sinh tại những trường trên, được tiến hành vào tháng 11 năm 1996. Điều tra trường học thu thập dữ liệu về tỷ lệ nhập học, tỷ lệ lưu ban và bỏ học, cơ sở vật chất và nhân lực, việc ra quyết định của nhà trường, và các đặc điểm của hiệu trưởng, giáo viên, học sinh và gia đình học sinh. Phần câu hỏi về quá trình ra quyết định ở nhà trường có tính độc đáo và bao gồm 25 câu hỏi nhằm tìm hiểu xem cải cách có thực sự làm tăng quyền ra quyết định của nhà trường không và điều này diễn ra như thế nào. Điều tra gồm 116 trường trung học (73 trường cải cách và 43 trường không cải cách đại diện cho nhóm đối chứng) và 126 trường tiểu học (80 trường cải cách và 46 trường không cải cách). Nhóm đối chứng được lựa chọn để phù hợp với những đặc điểm của các trường cải cách. Điều tra cũng thu thập dữ liệu về 400 giáo viên, 182 thành viên hội đồng và 3.000 học sinh cùng gia đình, trong đó chọn ngẫu nhiên từ 10 đến 15 học sinh ở mỗi trường. Những học sinh vẫn ở lại trường và có thể theo dõi được tham dự các kỳ kiểm tra trình độ được tiến hành vào cuối năm học 1996, và sau đó được tiến hành vào lần thu thập dữ liệu thứ hai trong năm 1997.

Phân tích dữ liệu định lượng sử dụng các kỹ thuật hồi quy để ước lượng hàm sản xuất của giáo dục. Kỹ thuật này xem xét tác động của cơ chế quản lý trường học (mức độ phân cấp) tới trình độ của học sinh, giữ nguyên các đầu vào của nhà trường, đặc điểm của hộ gia đình và học sinh. Phân tích này do lưỡng tác động của cả sự phân cấp theo luật định và trên thực tế. Sự phân cấp theo luật định chỉ phản ánh trường học có tham gia cải cách về mặt pháp lý hay không, còn sự phân cấp thực tế do

lường mức độ tự chủ thực sự mà trường học có được. Sự phân cấp thực tế được xác định bằng tỷ lệ phần trăm của 25 quyết định quan trọng mà trường học đưa ra. Kết quả này được dự kiến sẽ khác nhau giữa các trường do cải cách được tiến hành theo từng giai đoạn (vì vậy các trường học trong mẫu sẽ ở những giai đoạn khác nhau trong tiến trình cải cách) và do năng lực thực hiện thành công cải cách của các trường khác nhau tuỳ thuộc vào bối cảnh của nhà trường (một kết quả được xác nhận trong phân tích định tính).

#### IV. Các kết quả

Nghiên cứu định tính chỉ ra rằng, các thay đổi chính sách ở cấp trung ương không phải luôn luôn tạo ra những kết quả mạnh mẽ ở cấp địa phương. Nhìn chung, các cuộc cải cách đi kèm với sự tham dự nhiều hơn của cha mẹ học sinh và cải thiện được hoạt động quản lý và lãnh đạo nhà trường. Nhưng mức độ thành công của cải cách khác nhau tại các trường khác nhau. Các nhân tố đặc biệt quan trọng là mức độ đói nghèo của cộng đồng xung quanh (tại các cộng đồng nghèo, việc gia tăng tài trợ cho trường học trong vùng là khó khăn) và mức độ đoàn kết của các cán bộ trường học (khi những cán bộ chủ chốt như các giáo viên không cảm thấy cần tham gia tiến trình cải cách, thì thành công của sự phân quyền cũng sẽ bị hạn chế). Các nhà hoạch định chính sách thường không chú ý tới các bối cảnh địa phương đa dạng khi triển khai các chương trình mới. Các kết quả định tính chỉ ra rằng, trong trường hợp của Nicaragua, mục tiêu gia tăng nguồn tài trợ của địa phương cho các trường học dễ bị sai lệch trên thực tế, đặc biệt ở những cộng đồng nghèo, và vì vậy, cần phải suy nghĩ kỹ về mục tiêu này.

Nghiên cứu định lượng cũng có kết quả: các trường học tiến hành cải cách đã có nhiều quyết định của riêng mình hơn, đặc biệt là những quyết định liên quan đến các vấn đề về sư phạm và nhân sự. Quyền tự chủ theo luật định - việc một trường có ký kết hợp đồng tham dự cải cách hay không - không nhất

thiết có nghĩa là các trường có nhiều quyền ra quyết định hơn hay quyền ra quyết định của các trường đều tương đương nhau. Mức độ tự chủ thực sự có được phụ thuộc vào mức độ đói nghèo của cộng đồng và vào khoảng thời gian trường tham gia tiến trình cải cách. Kết quả hồi quy cho thấy, quyền tự chủ theo luật có liên quan rất ít tới các kết quả học tập của học sinh, trong khi quyền tự chủ thực tế - mức độ phân cấp thực sự mà nhà trường đạt được - liên quan đáng kể với các kết quả học tập tốt hơn của học sinh. (Đây là kết quả sơ bộ, còn có khả năng khai thác hơn nữa dữ liệu lặp mà gần đây đã trở nên sẵn có). Hơn nữa, các mô phỏng chỉ ra rằng, tăng cường phân cấp tại trường học tác động mạnh mẽ tới kết quả học tập của học sinh nhiều hơn sự cải thiện trong các chỉ số khác thường nhận được sự quan tâm chính sách, như đào tạo giáo viên, giảm quy mô lớp học, hay gia tăng số lượng sách giáo khoa.

## V. Ứng dụng chính sách

Các kết quả đánh giá cung cấp bằng chứng cụ thể cho thấy, Cải cách quyền tự chủ trong trường học ở Nicaragua đã thu được những kết quả dễ thấy. Các trường học tham dự cải cách thực sự đã có nhiều quyết định riêng hơn - như vậy là sự phân cấp đang diễn ra trên thực tế chứ không phải chỉ trong lý thuyết - và sự gia tăng quyền ra quyết định ở cấp địa phương này đã cải thiện kết quả học tập của học sinh.

Kết quả cũng chỉ ra những lĩnh vực mà tại đó có thể cải thiện chính sách, và vì vậy, Bộ Giáo dục đã tiến hành một số thay đổi trong chương trình cải cách trường học. Chương trình hiện nay nhấn mạnh nhiều hơn đến vai trò của giáo viên và việc nâng cao các khía cạnh sư phạm trong cải cách. Đào tạo giáo viên là một phần của chương trình này, và việc thiết lập một hội đồng sư phạm hiện đang được xem xét. Hơn nữa, để ứng phó với khó khăn về tài chính của các cộng đồng nghèo, Bộ Giáo dục đã xây dựng một chương trình trợ cấp căn cứ vào

phân bố đối nghịch. Cuối cùng, các lợi ích dễ nhận thấy của đánh giá này đã thúc đẩy Bộ đưa thành phần đánh giá thường xuyên vào trong chương trình cải cách của mình.

## **VI. Chi phí và tổ chức quản lý công tác đánh giá**

**Chi phí.** Tổng chi phí đánh giá xấp xỉ bằng 495.000 đôla, thấp hơn 1,5 phần trăm tín dụng của Ngân hàng Thế giới. (Tổng chi phí này không kể tới chi phí của nhóm đối tác trong nước của Bộ giáo dục Nicaragua). Trong tổng chi phí đánh giá này, 39 phần trăm chi phí là chi tiêu cho sự hỗ trợ kỹ thuật của các nhà tư vấn nước ngoài, 35 phần trăm là chi phí thu thập dữ liệu, 18 phần trăm là chi phí thời gian của cán bộ Ngân hàng Thế giới, và 8 phần trăm là chi phí đi lại.

**Tổ chức quản lý.** Đánh giá này do Bộ Giáo dục Nicaragua và Ngân hàng Thế giới phối hợp tiến hành. Tại Nicaragua, nhóm đánh giá do Patricia Callejas, Nora Gordon, và Nora Mayorga de Caldera ở Bộ Giáo dục chỉ đạo. Tại Ngân hàng Thế giới, đánh giá được tiến hành như là một phần trong dự án nghiên cứu, “đánh giá tác động của các dự án giáo dục liên quan tới phân cấp và tư nhân hoá”, với sự chỉ đạo của Elizabeth King, cùng với Laura Rawlings và Berk Ozler. Với sự điều phối của nhóm nghiên cứu thuộc Ngân hàng Thế giới, Bruce Fuller và Madgalena Rivarola ở Trường Giáo dục Harvard, cùng với Liliam Lopez từ Bộ giáo dục Nicaragua tiến hành đánh giá định tính.

## **VII. Các bài học**

**Giá trị của cách tiếp cận sử dụng kết hợp các phương pháp khác nhau.** Việc sử dụng các kỹ thuật nghiên cứu định tính và định lượng tạo ra sự kết hợp có giá trị của các kết quả có ích và phù hợp về mặt chính sách. Nghiên cứu định lượng cung cấp tổng quan chung, có tính hợp lý về mặt thống kê về các điều

kiên và kết quả của các trường học; nghiên cứu định tính tăng cường các kết quả này bằng cách tìm hiểu tại sao một số kết quả dự kiến của chương trình cải cách lại thành công trong khi những kết quả khác lại không đạt được, và nhờ vậy, đã góp phần chỉ dẫn cho sự điều chỉnh chính sách. Hơn nữa, do có tính trực quan, nghiên cứu định tính dễ tiếp cận hơn, và vì thế lý thú hơn đối với cán bộ của bộ, và điều này góp phần thúc đẩy nhanh chóng công việc tăng cường năng lực và tăng cường sự tín nhiệm đối với tiến trình đánh giá ở bộ.

**Tầm quan trọng của phát triển năng lực địa phương.** Phát triển năng lực ở địa phương thường tổn kém và đòi hỏi sự liên lạc và phối hợp thường xuyên với các đối tác của Ngân hàng Thế giới và các nhà tư vấn bên ngoài. Tuy nhiên, lợi ích của sự phát triển năng lực là sự gia tăng nhanh chóng quyền sở hữu và trách nhiệm của địa phương đối với tiến trình đánh giá, nhờ vậy, khuyến khích sự chấp nhận đối với các kết quả đánh giá, cho dù các kết quả này phản ánh tích cực hay tiêu cực về chương trình. Các kết quả đánh giá cung cấp đầu vào trực tiếp cho cải cách khi cải cách này được tiến hành. Tác động chính sách của đánh giá cũng được tăng cường nhờ một nhóm nghiên cứu địa phương, trong đó những người đánh giá và những người hoạch định chính sách phối hợp làm việc, và nhờ vị Bộ trưởng bộ giáo dục cũng tham gia hội đồng như là một người ủng hộ nhiệt thành đối với tiến trình đánh giá.

### VIII. Nguồn

Các tài liệu sau đây cung cấp thông tin chi tiết về đánh giá cải cách quyền tự chủ của trường học ở Nicaragua:

Fuller, Bruce, and Magdalena Rivarola. 1998. *Nicaragua's Experiment to Decentralize Schools: Views of Parents, Teachers and Directors*. Working Paper Series on Impact Evaluation of Education Reforms, paper no. 5. World Bank, Washington, D.C.

- King, Elizabeth, and Berk Ozler. 1998. *What's Decentralization Got to Do with Learning? The Case of Nicaragua's School Autonomy Reform*. Working Paper Series on Impact Evaluation of Education Reforms, paper no. 9. World Bank, Washington, D.C.
- King, Elizabeth, Berk Ozler, and Laura Rawlings. 1999. *Nicaragua's School Autonomy Reform: Fact or Fiction?* Washington, D.C.: World Bank.
- Nicaragua Reform Evaluation Team. 1996. *Nicaragua's School Autonomy Reform: A First Look*. Working Paper Series on Impact Evaluation of Education Reforms, paper no. 1. World Bank, Washington, D.C.
- Nicaragua Reform Evaluation Team. 1996. *1995 and 1997 Questionnaires, Nicaragua School Autonomy Reform*. Working Paper Series on Impact Evaluation of Education Reforms, paper no. 7. World Bank, Washington, D.C.
- Rawlings, Laura. 2000. "Assessing Educational Management and Quality in Nicaragua." In Bamberger, *Integrating Quantitative and Qualitative Methods in Development Research*. Washington, D.C.: World Bank.

## **Phụ lục 1.12: Cải cách giáo dục môn toán ở bậc tiểu học tại Nicaragua: Nghiên cứu thử nghiệm về tác động của sách giáo khoa và radio tới kết quả học tập**

### **I. Tóm tắt đánh giá**

Hầu hết các nước nghèo đều có các nguồn lực rất hạn chế dành cho giáo dục, vì vậy, việc phân bổ các nguồn lực này một cách có hiệu quả là rất quan trọng. Trong số ba lựa chọn chính sách thông dụng thường có- quy mô lớp học nhỏ, chương trình đào tạo giáo viên dài hơn, và cung cấp sách giáo khoa- thì chỉ có lựa chọn cuối thường được nhận thấy có tác động tích cực đáng kể tới kết quả học tập của học sinh. Đánh giá này định lượng hoá tác động của sự sẵn có sách giáo khoa tới kết quả học môn toán của học sinh lớp một ở Nicaragua.

Thiết kế của đánh giá này là cung cấp sách giáo khoa cho tất cả học sinh trong một tập hợp con, thuộc các lớp học trước kia đóng vai trò làm nhóm đối chứng trong một nghiên cứu đang tiến hành về hiệu quả của chương trình đào tạo qua radio. Một nửa số lớp học nhận được sách giáo khoa, còn nửa kia không nhận được. Tất cả các lớp học đều tham dự một bài kiểm tra trước vào đầu năm học, và một bài kiểm tra sau vào cuối năm học. Nghiên cứu này sau đó sử dụng các kỹ thuật hồi quy đơn giản để so sánh giữa trung bình điểm thi trong kỳ thi sau của các lớp như là một hàm số của điểm thi trong kỳ thi trước và chương trình cung cấp sách giáo khoa.

Một bài học quan trọng ở đây là làm thế nào thiết kế đánh giá một cách kỹ lưỡng: phương pháp lựa chọn ngẫu nhiên được thiết kế đặc biệt tốt và đã kết hợp một cách thông minh với một bài thi để tối đa hoá được khả năng so sánh giữa các lớp. Một bài học khác liên quan tới tính thực tế: đánh giá được thiết kế để nhận biết trước các vấn đề kinh tế chính trị khá nghiêm

trọng. Cuối cùng, đánh giá cho thấy một số ví dụ thực tế về những kiểu quyết định cần phải đưa ra trong điều tra thực địa.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Có hai thành phần rất lý thú trong thiết kế đánh giá. Đó là việc đánh giá này dựa vào một đánh giá tồn tại từ trước và sự tìm hiểu từ trước về môi trường chính trị tại đó đánh giá sẽ được tiến hành. Câu hỏi nghiên cứu quan trọng rất rõ ràng. Đó là đánh giá tác động của sự sẵn có về sách giáo khoa tới kết quả học tập của học sinh lớp một- đặc biệt tập trung vào việc liệu sách giáo khoa có được thực sự sử dụng trong lớp học hay không. Do đã có một chương trình dạy học qua radio (Radio Mathematics), nên câu hỏi nghiên cứu này được mở rộng để so sánh tác động của sự sẵn có sách giáo khoa với dạy học qua radio cũng như với một nhóm đối chứng.

Theo dõi việc sử dụng sách giáo khoa trên thực tế làm cho đánh giá trở nên khó khăn hơn. Nhiều chương trình giáo dục cung cấp tài liệu cho các lớp học, nhưng rõ ràng tác động của việc cung cấp này phụ thuộc vào việc sử dụng. Tuy nhiên, như những người đánh giá chỉ ra, quyết định này có nghĩa là đánh giá "không đánh giá khả năng sách giáo khoa hay các bài học qua radio cải thiện một cách tốt nhất kết quả của học sinh. Thay vì thế, đánh giá cố gắng đánh giá tác động của chúng như chúng được đón nhận tại một nước đang phát triển điển hình" (Jamison 1981 trang 559). Vì vậy, chỉ riêng việc cung cấp sách giáo khoa khó có thể đủ nếu như không có một phương pháp đảm bảo là các giáo viên sử dụng sách giáo khoa như dự kiến.

Danh giá này sử dụng thiết kế lựa chọn ngẫu nhiên, căn cứ theo một đánh giá dự án đã tồn tại từ trước. Trong Dự án Đầu phát thanh Nicaragua, người ta đã tiến hành phân công ngẫu nhiên và kiểm tra các thủ tục đang có để đánh giá tác động của một chương trình giáo dục qua radio. Dự án hiện có này đã phân nhóm tất cả các trường tiểu học tại ba tỉnh của Nicaragua vào các nhóm có radio và nhóm đối chứng sử dụng quy trình

chọn mẫu ngẫu nhiên, trên cơ sở phân lớp dựa theo mức độ đô thị hoá (có khoảng 30 phân trăm số học sinh ở các trường học nông thôn nhưng ở mỗi tầng (stratum), người ta lựa chọn số lớp học như nhau).

Đánh giá của chương trình sách giáo khoa khai thác thiết kế sẵn có này bằng cách lựa chọn các trường tham dự và các trường đối chứng theo cách sau. Trước tiên, những người đánh giá có được danh sách tất cả các trường có lớp học phù hợp với một trong sáu khoản mục (ba tỉnh, nông thôn và thành thị). Tiếp theo, họ phân bổ ngẫu nhiên các trường này vào nhóm tham dự và nhóm đối chứng từ danh sách nêu trên cho từng khoản mục, và sau đó, các trường này được sử dụng theo thứ tự xuất hiện (một trường từ chối tham dự được thay thế bằng trường tiếp theo trong danh sách). Các đề nghị tham dự chương trình của các lớp học trong nhóm đối chứng bị từ chối, và toàn bộ việc sử dụng các tài liệu thử nghiệm đều được các tác giả kiểm soát. Cần lưu ý là thiết kế đánh giá đã giải quyết những khó khăn tiềm tàng về chính trị ngay từ đầu. Nhóm đánh giá công bố dự định của họ từ đầu, nhận được sự chấp thuận chính thức và sự ủng hộ về chính sách, và nhóm này cũng thiết lập các thủ tục một cách rõ ràng và nhất quán cho chương trình.

Nghiên cứu này do vậy đã lựa chọn ngẫu nhiên 88 lớp học: 48 lớp học có radio và 40 thuộc về các trường đối chiếu. Hai mươi trường đối chiếu nhận được sách giáo khoa cho trẻ em trong trường, và các giáo viên nhận được các chỉ dẫn cả trên văn bản và bằng lời cùng với các tài liệu của giáo viên cho các bài kiểm tra. Chương trình radio bao gồm 150 bài học về môn toán, kết hợp với bản ghi cho học sinh và chỉ dẫn bằng lời và trên giấy cho giáo viên.

Một quyết định thú vị được đưa ra là việc cố tình không giám sát nhóm tham dự chương trình. Quyết định này rõ ràng là khó khăn vì sự không giám sát làm cho khó có thể đánh giá được mức độ hữu ích của chương trình. Tuy nhiên, phí tổn của giám sát trên khía cạnh tác động tới hành vi của người tham dự có lẽ là quá lớn. Các cuộc tham quan bất ngờ, một giải pháp

trung dung có thể chấp nhận được, đã không thể được sử dụng trong tình huống này do những náo động về chính trị vào năm đánh giá, và vì vậy, đã được tiến hành vào một năm sau.

Một quyết định thứ hai là, việc các cán bộ dự án là người tổ chức các đợt kiểm tra thay vì giáo viên của các lớp học. Điều này rõ ràng đã làm tăng chi phí hành chính, nhưng lại làm giảm đi sự thiên vị có thể có trong thi cử. Các học sinh tham dự một kỳ kiểm tra trình độ toán trong ba tuần đầu tiên của năm học. Kỳ kiểm tra sau nhằm đo lường kết quả của chương trình, dự kiến được tiến hành trong ba tuần cuối cùng của năm học, nhưng đã được thực hiện sớm hơn hai tuần, do các vấn đề về chính trị. Những học sinh nhận được các điều kiện tương tự nhau trong hai đợt kiểm tra, do các bài kiểm tra có cùng độ dài thời gian và các chỉ dẫn đều được thu băng.

### III. Dữ liệu

Có hai bài học chính được rút ra từ việc thu thập dữ liệu. Thứ nhất là, thường khó tránh khỏi các khó khăn về hậu cần. Mặc dù có sự chuẩn bị cẩn thận, vẫn nảy sinh một loạt các vấn đề liên quan tới việc xây dựng một tập hợp hoàn hảo những so sánh giữa kỳ kiểm tra trước và sau. Mặc dù có tổng cộng 20 lớp học đối chiếu, 20 lớp học có sách giáo khoa, và 47 lớp học có radio nhưng số các điểm kiểm tra trong các kỳ kiểm tra trước và sau khác nhau trong mỗi nhóm do đăng ký muộn, bỏ học, vắng mặt, và không thể kiểm tra vì quá đông. Các thông tin cá nhân về học sinh đã không được thu thập.

Bài học thứ hai là, phương pháp sáng tạo mà những người đánh giá sử dụng cho bài kiểm tra sau để giảm thiểu gánh nặng cho người tham dự mà vẫn đạt được các thông tin cần thiết. Người đánh giá gấp phải một số vấn đề như sau:

- Các bài kiểm tra tiêu chuẩn hoá không được sử dụng ở Nicaragua
- Bài kiểm tra phải đánh giá được mục tiêu của chương trình

- Bài kiểm tra phải đánh giá được kết quả trong mỗi chủ đề để cho phép đánh giá về hiệu quả của chương trình trong từng chủ đề nói riêng, cũng như trong tất cả các chủ đề.

Những người đánh giá sử dụng một thiết kế lựa chọn theo ma trận nhằm giải quyết các vấn đề này. Bài kiểm tra có hai dạng câu hỏi: các câu hỏi dành cho tất cả học sinh trong lớp (40 câu G) và các câu hỏi dành cho tập hợp nhỏ trong số học sinh (44 câu I). Tất cả các câu hỏi I đều được kiểm tra ở mọi lớp học trong khi một phần tư các câu hỏi G được kiểm tra ở mỗi lớp học. Nhờ vậy, các nhà nghiên cứu có thể phân công một cách ngẫu nhiên các đơn vị theo hai khía cạnh: trường học và các dạng bài kiểm tra. Điểm trung bình trong kỳ kiểm tra sau của nhóm tham dự và nhóm đối chứng được tính bằng cách cộng điểm trung bình của mỗi bài kiểm tra, và sai số chuẩn được xác định bằng cách sử dụng phương sai số dư, sau khi đã loại trừ tác động chính của các câu hỏi và của học sinh.

Thông tin về việc sử dụng sách giáo khoa cũng được thu thập vào một năm sau khi tiến hành chương trình tại 19 trong số 20 trường sử dụng sách giáo khoa.

#### **IV. Các kỹ thuật kinh tế lượng**

Cấu trúc của đánh giá có nghĩa là việc so sánh đơn giản giá trị trung bình của các nhóm tham dự và nhóm đối chứng là phù hợp, và trên thực tế, người ta đã làm như vậy. Phương pháp này có thể rất phiền phức nếu như tổng thể bao gồm nhiều lớp và chương trình có nhiều tác động, như ở trong tình huống đánh giá này. Do đó, những người đánh giá cũng sử dụng một phương pháp hồi quy đơn giản. Ở đây, lớp học là đơn vị phân tích, và điểm trung bình kỳ kiểm tra sau của lớp học được hồi quy theo các biến độc lập, gồm có điểm trung bình trong kỳ kiểm tra trước, các biến giả đại diện cho tác động của radio và sách giáo khoa, một biến giả biểu diễn cho thành thị - nông thôn, và điểm thi trung bình của lớp học trong kỳ thi trước.

Một thành phần quan trọng của đánh giá này là liệu các nhóm khác nhau có chịu tác động khác nhau từ cùng một tác động không. Có thể kiểm tra điều này, như đã được tiến hành ở trong đánh giá, bằng cách sử dụng một cách sáng tạo các biến tương tác. Sự khác nhau giữa thành thị và nông thôn được xác định bằng cách cho phép biến giả thành thị - nông thôn tác động đến chương trình; sự khác nhau trong tác động của chương trình căn cứ theo kết quả kiểm định ban đầu được xác định bằng cách cho phép yếu tố điểm thi ban đầu tác động đến chương trình.

## V. Ai là người tiến hành

Ngân hàng Thế giới hỗ trợ cho dự án nghiên cứu, nhưng dự án này đã được kết hợp trong Dự án dạy Toán qua Radio do Cơ quan Phát triển quốc tế Mỹ và Bộ giáo dục Nicaragua cùng thực hiện.

## VI. Kết quả

Các tác giả nhận thấy việc cung cấp sách giáo khoa và dạy học qua radio có tác động quan trọng tới kết quả của học sinh: sự sẵn có của sách giáo khoa đã làm tăng điểm kỳ kiểm tra sau của học sinh lên thêm 3,5 câu trả lời đúng, các bài học qua radio làm tăng 14,9 câu trả lời đúng. Mức tăng này rất đáng kể, vì độ lệch chuẩn của lớp học là 8,3, và của mỗi câu hỏi là 11,8. Các bài học qua radio và sách giáo khoa có hiệu quả hơn ở các trường nông thôn và có thể đóng vai trò quan trọng trong việc làm giảm khoảng cách về chất lượng giáo dục giữa thành thị và nông thôn. Các kết quả này dường như độc lập với trình độ kỹ năng ban đầu của lớp học, như được đo lường bằng điểm kỳ thi trước.

Các tác giả cho rằng, sự khác biệt trong các kết quả dạy học qua radio và cung cấp sách giáo khoa là do sự khác nhau trong

sử dụng sách giáo khoa, đặc biệt là do các giáo viên có trình độ kém.

## VII. Các bài học

Có ba bài học chủ yếu được rút ra từ đánh giá này: tầm quan trọng của chính trị trong các quyết định thiết kế, sự hữu ích của các thiết kế bài thi mang tính sáng tạo, và các khó khăn liên quan tới công việc điều tra thực địa. Thứ nhất, tính kinh tế chính trị của thiết kế lựa chọn ngẫu nhiên được nêu bật trong nghiên cứu này. Rõ ràng đã có những áp lực chính trị khá mạnh đối với chương trình và những áp lực này cần phải được giải quyết từ sớm và với sự hỗ trợ của chính phủ. Thứ hai, các tác giả đã có thể đo lường được nhiều khía cạnh trong kết quả học tập mà không cần phải tiến hành các bài kiểm tra dài dòng nhờ áp dụng một thiết kế bài thi có tính sáng tạo. Cuối cùng, những người đánh giá rõ ràng đã trả lời được một số câu hỏi liên quan tới công việc thực địa. Đó là liệu có thể theo dõi được sự chấp nhận sách giáo khoa trên thực tế hay không, làm thế nào để theo dõi được, và ai là người nên tổ chức kiểm tra kết quả.

## VII. Nguồn

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

## Phụ lục 1.13: Tác động của các chương trình hoàn trả chi phí khác nhau tới khả năng tiếp cận và bình đẳng ở Nigiê

### I. Giới thiệu

**Mô tả dự án.** Khả năng hoàn trả một phần chi phí y tế đóng vai trò quan trọng đối với việc cung ứng các dịch vụ y tế. Tuy nhiên, chúng ta vẫn biết rất ít về tác động của các chiến lược khác nhau tới chất lượng và các kết quả phúc lợi. Đánh giá này ước lượng tác động tới nhu cầu y tế trong hai chương trình hoàn trả chi phí thử nghiệm trong lĩnh vực chăm sóc sức khoẻ cơ bản (không phải là bệnh viện) ở Nigiê. Nigiê là một nền kinh tế nghèo, dựa vào nông nghiệp; chi phí y tế công cộng bằng từ 5 đến 6 phần trăm ngân sách của chính phủ; và hầu hết nguồn tài chính này đều được định hướng sai lầm tới các bệnh viện và nhân sự. Chính phủ mong muốn đánh giá hậu quả của các cơ chế thanh toán khác nhau và đã xem xét hai cơ chế: một cơ chế hoàn toàn trả phí cho dịch vụ và một cơ chế bao gồm thuế cộng với trả phí cho dịch vụ. Cả hai cơ chế này đều được kết hợp với sự tăng cường chất lượng và quản lý. Chính phủ đặc biệt quan tâm tới việc xác định xem nhu cầu chăm sóc sức khoẻ thay đổi như thế nào, đặc biệt ở trong các nhóm dễ tổn thương, và liệu sự cải thiện chất lượng có tính bền vững hay không.

**Các đặc điểm nổi bật của đánh giá.** Các cơ chế thanh toán khác nhau được thực hiện ở ba quận, một cơ chế đóng vai trò thử nghiệm và cơ chế kia để so sánh. Đánh giá sử dụng thiết kế bán thử nghiệm căn cứ vào điều tra hộ gia đình kết hợp với các dữ liệu hành chính về việc sử dụng và chi phí vận hành. Đánh giá này đặc biệt hấp dẫn do nó giải quyết trực tiếp các vấn đề kinh tế chính trị bằng cách đặt ra câu hỏi cho những người được phỏng vấn về mức độ sẵn sàng thanh toán của họ cho các dịch vụ được cải thiện. Điều này thể hiện sự thừa nhận bản thân các

kết quả có ý nghĩa cung không đủ đảm bảo rằng một dự án bền vững thực sự là một đóng góp có giá trị. Một khía cạnh có ích khác là việc đánh giá rõ ràng về tác động của chương trình tới các nhóm mục tiêu khác nhau (trẻ em, phụ nữ, những làng không có cơ sở y tế công cộng và những người nghèo nhất).

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Các câu hỏi chính là, tác động của chương trình tới (a) nhu cầu về các cơ sở chăm sóc y tế công cộng và việc sử dụng các cơ sở này, (b) các nhóm mục tiêu cụ thể (người nghèo, phụ nữ và trẻ em), (c) khả năng tiếp cận tài chính và địa lý, (d) việc sử dụng các dịch vụ thay thế, và (e) tính bền vững của sự cải thiện dịch vụ khi có sự hoàn trả chi phí (các chi phí người bệnh và chi phí thuốc cũng như các nguồn thu và khả năng sẵn sàng thanh toán).

Có ba huyện được lựa chọn trong số các tỉnh khác nhau từ một cơ quan đăng ký hành chính. Mặc dù tất cả các huyện này đều tương tự nhau về các đặc điểm kinh tế, nhân khẩu và xã hội, nhưng chúng khác nhau về mặt dân tộc. Mỗi huyện có một trung tâm y tế, với một trung tâm y tế bà mẹ và trẻ em, một trạm y tế, và một thầy thuốc, cũng như các trạm phát thuốc nông thôn.

Trong hai huyện tham gia dự án có bốn sự cải thiện về mặt chất lượng và quản lý, còn trong huyện đối chiếu thì không có sự cải thiện nào. Cụ thể, các cải thiện bao gồm, việc cung cấp thuốc trước; đào tạo cán bộ trong chẩn đoán và chữa trị; thiết lập một hệ thống quản lý thuốc tại kho và quản lý tài chính cùng với đào tạo cán bộ sử dụng hệ thống này; già tăng năng lực giám sát nhằm củng cố cho quản lý.

Có hai cơ chế định giá khác nhau được sử dụng cùng một lúc. Cơ chế thứ nhất là phí cho mỗi lần sử dụng dịch vụ, với mức phí 200 FCFA (0,66 đôla Mỹ) cho những người sử dụng trên năm tuổi và 100 FCFA (0,33 đôla Mỹ) cho trẻ em dưới năm

tuổi. Cơ chế thứ hai bao gồm một khoản thuế hàng năm bằng 200 FCFA do những người đóng thuế trong huyện trả và mức phí bằng 50 FCFA cho người sử dụng trên năm tuổi, và 25 FCFA cho trẻ em dưới năm tuổi. Thu nhập hàng năm của người dân thấp hơn 300 đô la Mỹ trên đầu người. Cả hai dự án đều có miễn phí cho các nhóm mục tiêu. Nguồn vốn được quản lý ở cấp huyện.

### III. Dữ liệu

Có ba huyện được chọn ra từ dữ liệu hành chính. Người ta tiến hành hai đợt điều tra hộ gia đình, trong đó có một đợt điều tra gốc. Cả hai đợt điều tra đều thu thập thông tin cá nhân và đặc điểm nhân khẩu học của hộ gia đình trong một mẫu được lựa chọn ngẫu nhiên gồm 1.800 hộ gia đình. Điều tra cơ sở thu thập thông tin về 2.833 người ốm đau trong hai tuần trước khi tiến hành điều tra và 1.770 phụ nữ đang mang thai. Điều tra cuối cùng có dữ liệu về 2.710 người ốm đau và 1.615 phụ nữ đang mang thai. Dữ liệu hành chính bao gồm các thông tin tương đối chi tiết về chi tiêu và quản lý thuốc hàng tháng, sự duy trì nhân sự, và thanh toán lệ phí cho việc sử dụng các cơ sở y tế. Thông tin này được thu thập vào một năm trước khi có chương trình, tức là năm cơ sở (từ 5-1992 đến 4-1993), và một năm sau khi có chương trình.

### IV. Các kỹ thuật kinh tế lượng

Nghiên cứu này kết hợp các so sánh giá trị trung bình bằng phương pháp logit đơn giản. Phương pháp này được sử dụng để xác định được sự thay đổi trong việc sử dụng dịch vụ y tế. Cụ thể, phản ứng cá nhân trong tình huống hiện có cơ sở y tế hay không được sử dụng để ( $P_1$ ) xác định mô hình sau:

$$\text{Logit } (P_1) = X\beta + * (A + B).$$

Mô hình này, trong đó có sự kiểm soát đối với các đặc điểm cá nhân X cũng như với các biến giả A và B, được so sánh với

$$\text{Logit } (P_1) = X\exists + *_a A + *_b B.$$

Biến giả A và B có nhiều định nghĩa. Trong dãy hồi quy đầu tiên, A chỉ thời kỳ tiến hành chương trình, B chỉ thời kỳ trước khi có chương trình, và phương trình hồi quy được chạy theo nhóm nhỏ (nhóm mục tiêu đã được xác định) và theo huyện. Trong dãy hồi quy thứ hai, A và B được sử dụng để tạo ra sáu cặp so sánh của huyện với huyện khác trong chương trình. Trong mỗi trường hợp, các tác giả kiểm định giả thuyết ( $*_a + *_b$ ) = \*. tác động của khả năng tiếp cận về mặt địa lý và tài chính được xác định trong ma trận X bằng các thước đo khoảng cách tính theo thời gian đi bộ và nhóm thu nhập. Chúng ta không rõ khoản mục nào đã bị bỏ qua trong mỗi trường hợp. Chúng ta cũng không rõ đã có sự điều chỉnh đối với sai số chuẩn trong các ước lượng do tính phân mảnh của thiết kế mẫu chưa.

Kỹ thuật logit là một phương pháp hiệu quả nhằm giải quyết ba trong số bốn câu hỏi nghiên cứu: dạng thức sử dụng, tác động tới các nhóm nhỏ, và tác động tới khả năng tiếp cận về địa lý và tài chính. Câu hỏi thứ tư - tác động của các thay đổi trong hoàn trả chi phí - được giải quyết nhờ dữ liệu hành chính và phương pháp so sánh đơn giản các giá trị trung bình. Một mối quan tâm hiển nhiên trong phương pháp sau, mặc dù không được thể hiện rõ ràng, là khả năng xảy ra sai lệch (bias) trong việc báo cáo các kết quả sau khi có chương trình. Cụ thể là, đã xảy ra một số hành vi "rủi ro đạo đức" nếu như những người thực hiện chương trình được đánh giá căn cứ vào các phản ứng tích cực đối với chương trình của người sử dụng.

Tác động của chương trình tới việc sử dụng các hệ thống y tế khác được xác định bằng các kỹ thuật kinh tế lượng, và không trình bày ở đây.

## V. Ai là người tiến hành

Bộ Y tế Công cộng đã tiến hành điều tra với sự hỗ trợ về tài chính và kỹ thuật của Tổ chức Phát triển quốc tế Mỹ và Ngân hàng Thế giới. Việc đánh giá do Francis Diop, Abode Yazbeck, và Ricardo Bitran của Abt Associates thực hiện.

## VI. Kết quả

Nghiên cứu này nhận thấy hệ thống kết hợp giữa thuế và phí không chỉ nhận được sự ủng hộ nhiều hơn mà còn tạo ra mức thu trên đầu người lớn hơn là hệ thống hoàn toàn dựa vào phí. Hệ thống phí và thuế cũng có những kết quả tốt hơn trong việc cung ứng khả năng tiếp cận tới các dịch vụ y tế tốt hơn của người nghèo, phụ nữ và trẻ em. Tuy nhiên, do yếu tố địa lý là một rào cản quan trọng đối với khả năng tiếp cận y tế, một hệ thống căn cứ vào thuế sẽ tái phân bổ một cách có hiệu quả chi phí y tế từ những người sống gần các cơ sở y tế sang những người sống cách xa các cơ sở này.

Tại huyện thực hiện hệ thống trả phí cho dịch vụ, số lần đầu đến cơ sở y tế giảm xuống, nhưng có sự gia tăng nhu cầu cho các dịch vụ y tế - so với sự gia tăng mạnh mẽ cả hai biến này tại huyện áp dụng hệ thống thuế cộng phí. Nguyên nhân chủ yếu là do sự gia tăng chất lượng của dịch vụ đi kèm với sự cải thiện chất lượng, và sự gia tăng chất lượng này có vai trò lớn hơn là việc bù đắp cho sự gia tăng chi phí.

Kiểm soát được chi phí- nhất là chi phí thuốc men- đi kèm với cải cách về chất lượng và quản lý cũng chứng tỏ được tính hiệu quả và bền vững. Mức độ hoàn trả chi phí tại huyện sử dụng hệ thống thuế cộng phí tiến đến và vượt qua mức 100 phần trăm trong khi lại thấp hơn đáng kể ở huyện sử dụng hệ thống trả phí cho dịch vụ. Hơn nữa, sự sẵn sàng trả tiền cũng cao hơn ở huyện thứ nhất so với huyện thứ hai.

Kết quả quan trọng là phương án thuế cộng phí có hiệu quả hơn trong việc đạt được các mục tiêu đã định, và cũng được

dân ưa thích hơn. Tuy vậy, đánh giá này cũng chứng minh: thiếu khả năng tiếp cận địa lý tới các cơ sở y tế là một rào cản quan trọng đối với sự sử dụng. Điều này cho thấy có một số vấn đề phân phối liên quan tới hệ thống thuế cộng phí: các hộ gia đình ở cách xa các trung tâm y tế sẽ phải trợ cấp ngầm cho các hộ gia đình sống ở gần các trung tâm này.

### VII. Các bài học

Có một số bài học có ích từ đánh giá này. Thứ nhất là, phương pháp xem xét nhiều khía cạnh, đánh giá tác động của dự án ở nhiều khía cạnh liên quan tới tính bền vững của dự án: không chỉ là tác động tới mức độ hoàn trả chi phí mà còn tới chất lượng và tới phản ứng của nhóm mục tiêu bị tác động. Một bài học khác là, sự quan tâm tới chi tiết trong khi thu thập dữ liệu đối với cả hai công cụ hành chính và điều tra. Điều này đem lại kết quả nhờ khả năng xác định chính xác những thành phần nào của chương trình có kết quả, và tại sao chúng có kết quả. Cuối cùng, phân tích về tác động của chương trình tới nhóm mục tiêu chứng tỏ tính hữu ích đặc biệt đối với các khuyến nghị chính sách.

### VIII. Nguồn

Diop, F., A. Yazbeck, and R. Bitran. 1995. "The Impact of Alternative Cost Recovery Schemes on Access and Equity in Niger." *Health Policy and Planning* 10 (3): 223– 40.

Wouters, A. 1995. "Improving Quality through Cost Recovery in Niger." 10 (3): 257– 70.

## Phụ lục 1.14: Các kết quả đi học trong các trường tiểu học ở Philippin: Đánh giá tác động của bốn thử nghiệm

### I. Giới thiệu

**Mô tả dự án.** Tại hầu hết các nước đang phát triển, tỷ lệ bỏ học cao và đi học không đầy đủ trong giáo dục tiểu học là một vấn đề quan tâm của các nhà hoạch định chính sách. Đây là trường hợp ở Philippin: có gần một phần tư trẻ em Philippin bỏ học trước khi hoàn thành lớp sáu, và những trẻ em rời khỏi nhà trường thường chỉ nếm được chưa đầy một nữa những gì đã được dạy. Chính phủ đã tiến hành một Chương trình Hạn chế bỏ học (DIP) vào năm 1990-92 nhằm giải quyết các vấn đề này. Chương trình này tiến hành bốn thử nghiệm: cung cấp các tài liệu học tập nhiều cấp độ (MLM), ăn trưa tại trường học (SL), và mỗi thử nghiệm đều trên lai kết hợp với mối quan hệ giữa cha mẹ học sinh và giáo viên (PTP). Phương pháp thử nhất cho phép các giáo viên điều chỉnh tiến độ giảng dạy theo nhu cầu khác nhau của học sinh và ít tốn kém hơn nhiều so với việc cho học sinh ăn tại trường. Quan hệ giữa cha mẹ và giáo viên hầu như không tốn kém chi phí, nhưng có thể giúp đỡ việc học hành của học sinh cả ở nhà và ở trường.

**Các đặc điểm nổi bật của đánh giá.** Đánh giá này đáng lưu ý vì nó nhằm mục tiêu rõ ràng là xây dựng năng lực trong nước chủ nhà để đánh giá sẽ trở thành một thành phần không thể thiếu trong các đề xuất mới, và các yêu cầu về dữ liệu sẽ được xem xét trước chứ không phải sau khi thực hiện dự án trong tương lai. Tuy vậy, hậu quả có một số vấn đề xảy ra, và đánh giá cần nêu rõ ràng các kết quả dự kiến. Một đóng góp quan trọng khác của đánh giá là việc kiểm tra tính vững vàng của kết quả bằng các phương pháp kinh tế lượng khác nhau. Cuối cùng, phân tích chi phí - lợi ích được áp dụng vào lúc cuối cùng

à rất quan trọng, do nó chính thức thừa nhận rằng, các kết quả có ý nghĩa vẫn là không đủ: các chương trình ít tốn kém có thể vẫn còn tốt hơn các chương trình tốn kém.

## II. Các câu hỏi nghiên cứu và thiết kế đánh giá

Câu hỏi nghiên cứu quan trọng là đánh giá tác động của bốn chương trình khác nhau tới hiện tượng bỏ học và kết quả học tập của học sinh. Tuy nhiên, thiết kế đánh giá được ràng buộc bởi nhu cầu thực tế cũng như bởi các yêu cầu của chương trình. Nhóm DIP đã tiến hành một quy trình lựa chọn trường học theo ba bước như sau:

- Tại năm vùng trong nước, chọn ra hai huyện ở mỗi vùng là các địa phương có thu nhập thấp. Tại một huyện, các lựa chọn chương trình là đổi chứng, MLM hoặc MLM-PTP; ở huyện còn lại, các lựa chọn là đổi chứng, SL hoặc SL- PTP. Việc phân công hai nhóm chương trình này được thực hiện ngẫu nhiên.
- Tại mỗi huyện, nhóm đánh giá chọn lựa ba trường có (a) tất cả các cấp lớp, với một lớp học ở mỗi cấp; (b) có tỷ lệ bỏ học cao và (c) hiện không có chương trình cho ăn tại trường.
- Ba trường học ở mỗi huyện được phân công thuộc về nhóm đổi chứng hay một trong hai chương trình trên cơ sở lựa chọn ngẫu nhiên.

Mỗi chương trình được phân công ngẫu nhiên cho tất cả các lớp học ở năm trường, và người ta tiến hành các kỳ kiểm tra trước và sau trong cả năm 1991 và 1992 tại tất cả các lớp học ở tất cả 20 trường cũng như ở 10 trường đổi chứng.

## III. Dữ liệu

Thủ tục thu thập dữ liệu có tính chỉ dẫn. Việc thu thập các dữ liệu gốc được bắt đầu vào năm 1990-91, và các chương trình

được từ 29  
mỗi r  
như  
tin th  
một p  
trước  
ở ba  
Dù  
trườ  
fier)  
sinh  
nhận  
nhất  
ban  
có th  
phép

Cáu  
lượn  
quả  
rõ r  
và t  
chứ  
thứ

T  
khá  
trun  
năm  
chiế  
kết  
đổi  
này

được thực hiện năm 1991-92. Thông tin chi tiết được thu thập từ 29 trường học, khoảng 180 giáo viên và khoảng 4000 học sinh mỗi năm trong hai năm này. Mặc dù bảng câu hỏi rất chi tiết nhưng nó đã trở thành không cần thiết. Chỉ có một số ít thông tin thực sự được sử dụng, điều này cho thấy, có thể giảm đi một phần gánh nặng trong quá trình đánh giá. Các kỳ kiểm tra trước và sau cũng được thực hiện vào đầu và cuối mỗi năm học ở ba môn: toán, tiếng Philippin, và tiếng Anh.

Dữ liệu được cấu trúc theo thời gian cho các học sinh và trường học. Không may là những đặc điểm nhận dạng (identifier) về học sinh đã trở nên không có tính đặc trưng cho các học sinh và trường học giữa hai năm này. Cần lưu ý là, người ta chỉ nhận ra vấn đề này sau sáu tháng làm việc cho thấy có sự thiếu nhất quán nội bộ. Việc sử dụng trở lại các đặc điểm nhận dạng ban đầu của Bộ giáo dục Philippin là không thể. May mắn là có thể cứu được dữ liệu cho các học sinh lớp một, nhờ đó cho phép có một số phân tích theo thời gian.

#### IV. Các kỹ thuật kinh tế lượng

Cấu trúc thủ tục chọn mẫu cho thấy có một số vấn đề kinh tế lượng thú vị: đó là phân tích tình trạng bỏ học và phân tích kết quả điểm thi. Trong mỗi trường hợp, có hai tiêu chuẩn so sánh rõ ràng: thứ nhất là, nhóm đối chứng bao gồm các trường học, và thứ hai là điều tra gốc được tiến hành vào năm trước khi có chương trình. Các tác giả giải quyết vấn đề theo những cách thức khác nhau.

Trong phân tích về sự bỏ học, cần sử dụng phương pháp khác biệt trong khác biệt và so sánh thay đổi trong tỷ lệ bỏ học trung bình của mỗi lớp học tham gia chương trình trong hai năm với thay đổi trong tỷ lệ bỏ học trung bình của các lớp đối chiếu. Tuy nhiên, có hai vấn đề lập tức phát sinh. Thứ nhất, các kết quả mặc dù tương đối lớn về quy mô, nhưng chỉ có ý nghĩa đối với chương trình MLM, có thể do quy mô mẫu nhỏ. Vấn đề này không phải là không phổ biến với dạng quy trình này và

có thể thường xuyên xảy ra do thiếu nguồn tài chính để tiến hành các thử nghiệm quy mô lớn ở các nước đang phát triển. Thứ hai, sự kiểm tra ngắn gọn liệu các đặc điểm và kết quả của học sinh trên thực tế có giống nhau không giữa các trường học trong năm trước, khi chưa có chương trình cho thấy có một số đặc điểm khác biệt đáng kể. Hai vấn đề này dẫn đến việc các tác giả phải kiểm tra tính vững vàng của kết quả bằng kỹ thuật hồi quy dạng logit có kiểm soát các đặc điểm cá nhân (PC) và nền tảng gia đình (FB). Những kết quả cơ bản không thay đổi. Tuy nhiên, kỹ thuật hồi quy đã không cho thấy một nguyên nhân cơ bản gián tiếp quan trọng khác dẫn đến sự bỏ học, đó là kết quả học tập yếu kém. Điều này tự nhiên dẫn đến phân tích thứ hai, tập trung vào kết quả học tập.

Một số các mối lo ngại khác về kinh tế lượng được nêu ra trong đánh giá về tác động của chương trình INTER tới kết quả học tập của cá nhân I tại trường s vào thời điểm t (AP<sub>ist</sub>), được tác giả mô hình hoá dưới dạng sau:

$$AP_{ist} = \delta_0 + \delta_1 AP_{ist-1} + \delta_2 PC_i + \delta_3 FB_i + \delta_4 LE_{st} + \delta_5 CC_i + \delta_6 INTER_{jt} + \epsilon$$

trong đó LE là môi trường học tập và CC là tình trạng của lớp học.

Vấn đề thứ nhất trong số các vấn đề này là, tính đến việc có thể tồn tại hiện tượng tương quan phân mảnh (clustered coorelation) trong các sai số của học sinh trong cùng lớp học và trường học. Thứ hai là, cố gắng nắm bắt được hiện tượng phương sai không đồng nhất. Và thứ ba là, một vấn đề có liên quan- đó là sai lệch do lựa chọn (selection bias).

Vấn đề thứ nhất được giải quyết bằng cách áp dụng điều chỉnh Huber-White đối với sai số chuẩn. Vấn đề thứ hai, về nguyên tắc, có thể được giải quyết ở cấp độ cá nhân bằng cách sử dụng sự khác biệt trong điểm thi như là một biến độc lập. Tuy nhiên, các tác giả cho rằng, phương pháp này không phù hợp vì nó giả định trước là giá trị của  $\delta_1$  bằng 1, và giả định này không được chứng thực trong các kiểm tra. Các tác giả sử

dụng phương trình có độ trễ trong biến phụ thuộc, nhưng điều này lại làm phát sinh một vấn đề khác: đó là hiện tượng sai lệch trong một biến nội sinh. Có thể giải quyết vấn đề này bằng cách thay điểm kỳ kiểm tra trước của mỗi môn bằng điểm kiểm tra trước trong các môn khác. Tuy nhiên, các tác giả nhận thấy việc làm giảm sai lệch như vậy đã phải trả giá vì nó làm giảm tính hiệu quả, và do đó đã cho các kết quả sử dụng cả phương pháp bình phương nhỏ nhất và phương pháp biến công cụ. Các tác giả sử dụng cả tác động cố định của trường học và giáo viên để kiểm soát hiện tượng phương sai không đồng nhất không quan sát được trong LE và CC.

Vấn đề thứ ba cũng là vấn đề thường xuyên gặp phải trong các tài liệu nghiên cứu và vẫn chưa có giải pháp được chấp nhận hoàn toàn cho nó. Đó là sai lệch do lựa chọn. Rõ ràng, do tỷ lệ bỏ học khác nhau nên kết quả học tập của một cá nhân phụ thuộc vào quyết định không bỏ học. Mặc dù vấn đề này thường được giải quyết bằng thủ tục Heckman hai giai đoạn, nhưng vẫn còn có sự không hài lòng lầm với cách làm này, ít nhất vì ba lý do: sự nhạy cảm của phương pháp này với giả định phân phối chuẩn, việc chọn lựa và sự đầy đủ các biến phù hợp để sử dụng trong giai đoạn một, và sự phụ thuộc thường xuyên vào sự đồng nhất (identification) thông qua tính phi tuyến trong giai đoạn một. Không may là vẫn chưa có sự nhất trí về một phương án khác phù hợp. Krueger đề xuất một phương án: phân bổ cho những người bỏ học thứ tự của họ ở kỳ thi trước và đưa chúng vào trong phương trình hồi quy. Do đó các tác giả đã nêu ra cả ba tập hợp kết quả theo các phương pháp: hồi quy đơn giản kết quả theo chương trình, phương pháp Krueger, và thủ tục Heckman.

## V. Ai là người tiến hành

Việc thu thập dữ liệu do Phòng Giáo dục tiểu học thuộc Bộ Giáo dục, Văn hoá và Thể thao Phillipin tiến hành. Việc phân

tích do một cán bộ Ngân hàng Thế giới và hai nhà nghiên cứu tiến hành.

## VI. Kết quả

Nghiên cứu này đánh giá tác động của các chương trình tới tình trạng bỏ học ở các lớp từ lớp một đến lớp sáu và tới kết quả điểm thi ở lớp một sử dụng phương pháp khác biệt trong khác biệt, các kỹ thuật biến công cụ và phương pháp lựa chọn Heckman. Tác động của việc cung cấp các tài liệu đa trình độ - đặc biệt cùng với sự phối hợp giữa cha mẹ và giáo viên - tới tỷ lệ bỏ học và sự cải thiện kết quả học tập có tính vững chắc trước các tiêu chuẩn khác nhau và cũng có chi phí - hiệu quả cao. Tác động của bữa ăn trưa tại trường nhìn chung là yếu. Một thành phần thú vị của nghiên cứu này là phân tích chi phí lợi ích - nhờ đó cho thấy các kết quả có ý nghĩa vẫn là chưa đủ! Cụ thể, việc tính toán các chi phí trực tiếp và gián tiếp (chi phí cơ hội) của chương trình đã đưa đến kết luận: phương pháp MLM vừa hiệu quả vừa tiết kiệm chi phí.

Tuy vậy, sự thiếu hiệu quả của kế hoạch ăn tại trường có thể đã được phóng đại. Có thể có một cách tiếp cận có trọng điểm hơn đối với chương trình ăn tại trường phù hợp. Hơn nữa, do khoảng thời gian giữa khi thực hiện và đánh giá chương trình tương đối ngắn nên đánh giá không thể phản ánh được tác động lâu dài của các chương trình.

## VII. Các bài học

Có một số bài học trong tiến trình đánh giá này. Bài học quan trọng thứ nhất là các khó khăn nằm trong các chi tiết. Rất nhiều thông tin theo thời gian có thể bị mất đi nếu như không có đủ thông tin, chẳng hạn tính duy nhất của các đặc điểm nhân dạng không còn. Bài học thứ hai là có rất ít thông tin được thu thập từ các điều tra chi tiết được sử dụng, và do đó, có thể giảm bớt đáng kể gánh nặng cho những người trả lời phỏng vấn. Thứ ba,

nghiên cứu này nhấn mạnh giá trị của các cách tiếp cận kinh tế lượng khác nhau và lợi ích của việc tìm ra tính nhất quán trong các phương pháp khác nhau. Thứ tư, nghiên cứu này là mẫu mực trong việc sử dụng phân tích chi phí-lợi ích, cả trong việc nhận ra và đánh giá chi phí của các chương trình khác nhau. Cuối cùng, mặc dù rõ ràng đã có những sai sót trong nghiên cứu, nhưng các tác giả đã lưu ý là động cơ chủ yếu của nghiên cứu là phát triển năng lực đánh giá ở Philippin. Việc DIP được thực hiện và đánh giá có nghĩa là có thể phát triển năng lực này trong Bộ giáo dục.

### VIII. Nguồn

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

## **Phụ lục 1.15: Đánh giá tác động tới đói nghèo của các dự án đường bộ nông thôn ở Việt Nam**

### **I. Giới thiệu**

**Mô tả dự án.** Đường bộ nông thôn được Ngân hàng Thế giới và các nhà tài trợ khác xem là một công cụ nhằm làm giảm đói nghèo. Dự án giao thông nông thôn ở Việt Nam I được tiến hành vào năm 1997 với sự tài trợ của Ngân hàng Thế giới cho việc thực hiện dự án trong vòng từ ba đến năm năm. Mục đích của dự án là nâng cao mức sống ở các vùng nghèo bằng cách tái thiết những con đường và cầu đang tồn tại và nâng cao khả năng tiếp cận thị trường. Ở mỗi tỉnh tham dự, người ta nhận diện các dự án tái thiết bằng chỉ tiêu chi phí thấp nhất (số người dân sẽ được hưởng lợi và chi phí của dự án). Tuy nhiên, trong một nỗ lực nhằm tăng cường tính định hướng theo tình trạng đói nghèo, 20 phần trăm nguồn tài trợ cho mỗi tỉnh sẽ được dành cho các vùng miền núi, thưa thớt dân cư, nơi các dân tộc ít người sinh sống. Các dự án ở những địa phương này không hoàn toàn đủ tiêu chuẩn theo tiêu chí chi phí thấp nhất.

**Đánh giá tác động.** Mặc dù có sự nhất trí chung về tầm quan trọng của đường nông thôn, nhưng vẫn có rất ít bằng chứng cụ thể về mức độ và bản chất của các lợi ích có được từ cơ sở hạ tầng này. Mục tiêu của Đánh giá tác động Đường nông thôn ở Việt Nam là xác định xem phúc lợi của các hộ gia đình thay đổi như thế nào tại các xã có dự án đường nông thôn so với các xã không có. Vấn đề quan trọng đối với đánh giá là tách biệt một cách thành công tác động của đường sá khỏi vô số các nhân tố khác đang thay đổi ở nông thôn Việt Nam ngày nay như là kết quả của quá trình chuyển đổi sang nền kinh tế thị trường hiện đang diễn ra.

Đánh giá được khởi đầu đồng thời với việc chuẩn bị dự án,

vào đầu năm 1997, và vẫn đang được tiến hành. Hiện vẫn chưa có kết quả của đánh giá. Đánh giá này có sức hấp dẫn vì nó là một trong những nỗ lực toàn diện đầu tiên nhằm đánh giá tác động của một dự án đường nông thôn tới các kết quả phúc lợi. Đó là đánh giá xem liệu dự án có thực sự làm giảm đói nghèo hay không. Thiết kế đánh giá được cải tiến trên cơ sở các đánh giá cơ sở hạ tầng trước đây bằng cách kết hợp những thành phần sau: (a) thu thập dữ liệu điều tra gốc và điều tra tiếp theo, (b) bao gồm những đơn vị đối chiếu phù hợp để các kết quả có tính vững vàng trước các nhân tố không thể quan sát mà có tác động tới cả sự tham dự và kết quả của chương trình, và (c) được tiến hành theo sau dự án trong một khoảng thời gian đủ dài (thông qua các đợt thu thập dữ liệu nối tiếp nhau) để nắm bắt được đầy đủ tác động phúc lợi của dự án.

## II. Thiết kế dự án

Thiết kế của Đánh giá tác động Đường nông thôn ở Việt Nam tập trung vào các dữ liệu điều tra gốc (trước khi có chương trình) và điều tra tiếp theo (sau khi có chương trình) cho một mẫu gồm các xã tham dự và không tham dự dự án. Có thể xác định các đơn vị đối chiếu phù hợp trong số các xã không tham dự nhờ kỹ thuật so sánh tương hợp. Dữ liệu cơ sở cho phép so sánh trước và sau ("phản thân") về các chỉ tiêu phúc lợi tại những xã có dự án và các xã thuộc nhóm đối chứng. Về lý thuyết, nhóm đối chứng, được lựa chọn theo kỹ thuật so sánh tương hợp, giống như nhóm dự án về các đặc điểm quan sát được và không quan sát được, nhờ thế có thể quy các kết quả của những cộng đồng tham dự chương trình cho vai trò của dự án.

## III. Thu thập dữ liệu và kỹ thuật phân tích

Dữ liệu được thu thập để đánh giá bao gồm các điều tra cấp

xã và cấp hộ gia đình, cùng với cơ sở dữ liệu cấp huyện, cấp tỉnh và cấp dự án. Các điều tra gốc và tiếp theo ở cấp xã và cấp hộ gia đình được tiến hành vào năm 1997 và 1999, và các kỳ điều tra thứ ba và thứ tư được tiến hành cách quãng hai năm như dự kiến. Mẫu điều tra bao gồm 100 xã tham gia dự án và 100 xã không tham gia, ở 6 trong số 18 tỉnh có dự án. Các xã có dự án được lựa chọn ngẫu nhiên từ một danh sách tất cả các xã có đề xuất dự án tại mỗi tỉnh. Sau đó, một danh sách được xây dựng bao gồm tất cả các xã còn lại tại các huyện có đề xuất dự án, và các xã đối chiếu được rút ra một cách ngẫu nhiên từ danh sách này. (Lý tưởng là các đơn vị đối chứng chỉ khác nhau tham gia dự án ở chỗ là chúng không tham gia chương trình. Và vì lý do hậu cần, nên điều tra thực địa được giới hạn tại những khu vực nhất định. Nhóm đối chứng do vậy được lựa chọn ở huyện lân cận, hay ở chính huyện có xã tham gia dự án. Các huyện thường lớn và khó có tác động lan truyền từ xã có dự án và xã không có dự án, nhưng cũng cần phải kiểm tra kỹ lưỡng điều này.) Kỹ thuật tương hợp mức độ xu hướng dựa trên đặc điểm của các xã sẽ được sử dụng để kiểm định sự lựa chọn nhóm đối chứng, và bất kỳ đơn vị đối chứng nào có các đặc điểm khác thường so với các xã tham gia dự án sẽ bị loại ra khỏi mẫu. Một mô hình logit về sự tham dự của các xã vào dự án sẽ được ước lượng và sử dụng để đảm bảo là xã đối chứng có mức độ xu hướng tương tự (giá trị dự đoán từ mô hình logit).

Cơ sở dữ liệu các xã được rút ra từ dữ liệu hành chính được thu thập hàng năm của các xã, bao gồm dữ liệu về tình hình nhân khẩu, sử dụng đất và hoạt động sản xuất, và được bổ sung bằng một kỳ điều tra cấp xã nhằm phục vụ cho đánh giá này. Điều tra bao gồm những đặc điểm chung, cơ sở hạ tầng, việc làm, nguồn sinh sống, nông nghiệp, đất đai và các tài sản khác, giáo dục, y tế, các chương trình phát triển, các tổ chức cộng đồng, tài chính của xã và giá cả. Những dữ liệu này sẽ được sử dụng để xây dựng một số chỉ số cấp xã về phúc lợi và

để kiểm định tác động của các chương trình theo thời gian.

Mục tiêu chính của điều tra hộ gia đình là nắm bắt được thông tin về khả năng tiếp cận các phương tiện và dịch vụ khác nhau của hộ gia đình, và khả năng tiếp cận này thay đổi như thế nào theo thời gian. Bảng câu hỏi hộ gia đình áp dụng cho 15 hộ gia đình được lựa chọn ngẫu nhiên ở mỗi xã, bao gồm các vấn đề về việc làm, tài sản, hoạt động sản xuất và việc làm, giáo dục, sức khoẻ, thị trường, tín dụng, hoạt động của cộng đồng, khả năng tiếp cận các chương trình an toàn xã hội và chống nghèo đói, và giao thông. Do năng lực điều tra hạn chế ở trong nước, vẫn chưa có một cố gắng nào nhằm thu thập những dữ liệu phức tạp cần thiết để tạo ra một chỉ số phúc lợi của hộ gia đình (ví dụ thu nhập hay tiêu dùng). Tuy nhiên, có một số câu hỏi được đưa vào trong điều tra cũng tương tự như các câu hỏi trong Điều tra mức sống dân cư ở Việt Nam. Sử dụng thông tin này và các thông tin khác về đặc điểm của hộ gia đình chung cho cả hai kỳ điều tra, người ta sẽ sử dụng các kỹ thuật hồi quy để ước lượng vị trí của mỗi hộ gia đình trong phân phối phúc lợi trên toàn quốc. Một cơ sở dữ liệu cấp huyện có quy mô nhỏ cũng được chuẩn bị nhằm cung cấp bối cảnh cho các dữ liệu cấp xã, bao gồm dữ liệu về dân số, sử dụng đất, kinh tế và cả chỉ số xã hội. Các điều tra này sẽ được lặp lại theo lịch trình điều tra các xã.

Các thông tin sẵn có được sử dụng để tạo ra hai cơ sở dữ liệu bổ sung. Một cơ sở dữ liệu lớn cấp tỉnh được tạo ra để giúp hiểu được về sự lựa chọn các tỉnh tham gia dự án. Cơ sở dữ liệu này bao gồm tất cả các tỉnh của Việt Nam, và gồm có dữ liệu rộng rãi về nhiều biến kinh tế xã hội. Cuối cùng, một cơ sở dữ liệu ở cấp dự án cho mỗi khu vực dự án được điều tra cũng được xây dựng nhằm kiểm soát phạm vi dự án và phương pháp thực hiện trong đánh giá tác động của dự án.

Dữ liệu gốc sẽ được sử dụng để mô hình hoá sự lựa chọn các khu vực dự án bằng cách tập trung vào các quá trình kinh tế, xã hội, và kinh tế chính trị. Các kỳ điều tra sau sẽ được sử

ụng để hiểu được các lợi ích có thể đo lường được ở cấp xã, hụ thuộc vào sự lựa chọn này. Phương pháp phân tích sẽ là phương pháp "khác biệt kép", và phương pháp tương hợp. Phương pháp tương hợp sẽ được sử dụng để lựa chọn đối tượng đối chứng lý tưởng trong số mẫu gồm 100 xã không tham gia dự án. Các kết quả trong những xã tham gia dự án sẽ được so sánh với kết quả của những xã đối chiếu, cả trước và sau khi khởi công các dự án đường. Tác động của chương trình sau đó sẽ được xác định như là sai khác giữa các kết quả tại các khu vực dự án sau khi có chương trình và trước khi có chương trình, trừ đi sai khác trong các kết quả tương ứng tại khu vực đối chứng. Phương pháp này cung cấp một ước lượng không chênh lệch (unbiased estimate) của tác động dự án với sự hiện diện của các nhân tố không thay đổi theo thời gian và không quan sát được mà có tác động tới cả việc lựa chọn khu vực dự án và kết quả của dự án. Kết quả sẽ được tăng cường hơn nữa nhờ các tập hợp dữ liệu về các chỉ số kết quả và biến giải thích rất phong phú. Các chỉ số kết quả sẽ được xem xét bao gồm, năng suất nông nghiệp cấp xã, sự đa dạng hóa nguồn thu nhập, cơ hội việc làm, sử dụng và phân phối đất, sự sẵn có của hàng hoá, dịch vụ và trang thiết bị, và sự giàu có và tình trạng phân phối tài sản.

#### **IV. Chi phí và tổ chức quản lý công tác đánh giá**

**Chi phí.** Tổng chi phí của đánh giá tính đến nay là 222.500 đôla hay 3,6 phần trăm tổng chi phí dự án. Chi phí này bao gồm 202.500 đôla chi phí dành cho kỳ thu thập dữ liệu đầu tiên và 20.000 đô la dành cho việc nghiên cứu. Phí tổn về thời gian và di lại của cán bộ Ngân hàng Thế giới không được đưa vào các chi phí này.

**Tổ chức quản lý.** Đánh giá do cán bộ của Ngân hàng Thế giới

Đòn  
kiến  
và  
nướ  
tạo  
cần,

Var  
P

Dominique van de Walle thiết kế. Một nhà tư vấn độc lập có kiến thức về kinh tế học và nghiên cứu trong lĩnh vực đói nghèo và phát triển nông thôn được tuyển làm người theo dõi trong nước cho nghiên cứu này. Nhà tư vấn này tuyển dụng và đào tạo những người giám sát nhóm, tổ chức các hoạt động hậu cần, và giám sát tất cả các quá trình thu thập dữ liệu.

## VII. Nguồn

Van de Walle, Dominique. 1999. *Assessing the Poverty Impact of Rural Road Projects*. World Bank, Washington, D.C. Processed.

## Phụ lục 2

# Mẫu điều khoản tham chiếu

### Ví dụ I: Dự án dinh dưỡng và phát triển trẻ em ở Uganda

#### Điều khoản tham chiếu cho hằng tư vấn tham gia hỗ trợ đánh giá dự án

##### I. Cơ sở

Chính phủ Uganda đã vay tín dụng từ Hiệp hội Phát triển Quốc tế để tài trợ cho chi phí của Dự án dinh dưỡng và phát triển trẻ em. Dự án này tập trung vào việc nâng cao chất lượng cuộc sống cho trẻ dưới sáu tuổi và tăng cường năng lực chăm sóc trẻ em của các gia đình và cộng đồng. Cụ thể, dự án sẽ nhằm mục tiêu đạt được sự phát triển của trẻ em thông qua cải thiện tình trạng dinh dưỡng, sức khoẻ, tâm lý, và nhận thức của trẻ em dưới sáu tuổi ở Uganda.

##### II. Tính cần thiết phải đầu tư vào sự Phát triển của trẻ em

Đầu tư vào phát triển trẻ em (ECD) có những lợi ích hữu hình không chỉ cho trẻ em và cha mẹ chúng mà còn cho toàn bộ cộng đồng và quốc gia. Sự tăng trưởng về mặt thể chất và phát triển về mặt tinh thần xảy ra trong thời kỳ trẻ sơ sinh và trẻ em: ở độ tuổi hai tuổi, bộ não của một đứa trẻ đã phát triển gần như hoàn toàn. Năng lực nhận thức cũng phát triển đáng kể khi đứa trẻ bốn tuổi. Sự tăng trưởng và phát triển đầy đủ về mặt thể chất và tinh thần trong thời kỳ trẻ em nâng cao khả năng sẵn

sàng đi học, cải thiện tình trạng lưu ban khi đi học và đóng góp cho sự phụ thuộc vào vốn con người. Các trẻ em có nguồn gốc xuất thân bất lợi có thể hưởng nhiều lợi ích từ sự chăm sóc trẻ, nhờ vậy thu hẹp chênh lệch và tình trạng bất bình đẳng đi kèm với đói nghèo.

Sức khoẻ và dinh dưỡng tốt cũng như sự kích thích trí óc là điều kiện quan trọng để đưa trẻ phát triển các cấu trúc khái niệm vững vàng trong cuộc sống sau này. Sự phối hợp giữa dinh dưỡng, sức khoẻ và kích thích trí óc là rất quan trọng nên chỉ có thể đạt được các tác động tích cực tới sự tăng trưởng và phát triển của trẻ em nhờ một phương pháp tiếp cận tổng hợp.

### III. Các mục tiêu và chiến lược của dự án

Mục tiêu phát triển của dự án là cải thiện sự tăng trưởng và phát triển của trẻ em dưới sáu tuổi về mặt dinh dưỡng, sức khoẻ, tâm lý và nhận thức. Việc đạt được các mục tiêu này vào cuối thời kỳ thực hiện 5 năm sẽ được đo lường bằng các chỉ số sau: (a) giảm một phần ba số trẻ em chưa đi học thiếu cân theo mức năm 1995 tại các huyện có dự án; (b) giảm một phần tư tỷ lệ trẻ em còi cọc khi bước vào trường tiểu học theo mức năm 1995 tại các huyện có dự án, (c) cải thiện sự phát triển về tâm lý và nhận thức của trẻ em, (d) giảm tỷ lệ lưu ban và bỏ học ở các lớp thấp tại các trường tiểu học, và (e) phát triển năng lực kinh doanh và tăng cường quyền lực kinh tế của các bà mẹ và những người chăm sóc trẻ em.

Dự án này hỗ trợ Chương trình hành động quốc gia vì trẻ em Uganda và Kế hoạch hành động nhằm triệt tiêu đói nghèo. Dự án này đặc biệt tăng cường sự sẵn sàng cho đi học của trẻ em, và vì vậy đóng góp cho việc đạt được mục tiêu giáo dục phổ cập tiểu học. Chiến lược chủ yếu của dự án là nâng cao năng lực của các gia đình và cộng đồng trong việc chăm sóc trẻ em chưa đến độ tuổi đi học tốt hơn (trẻ từ không đến sáu tuổi), thông qua tăng cường kiến thức về tăng trưởng và phát triển của trẻ em, kiến thức làm cha mẹ, dinh dưỡng và chăm sóc sức khoẻ, và các hoạt động tạo ra thu nhập cho phụ nữ.

#### **IV. Các tiếp cận của dự án**

Dự án này là một chương trình được ưu tiên ở địa phương, diễn ra theo quá trình thay vì như một chương trình hoàn chỉnh. Các đầu vào được đưa vào cộng đồng theo từng giai đoạn như là kết quả của tiến trình lập kế hoạch tham dự nhằm đảm bảo sự sở hữu và tính bền vững. Chương trình sẽ có sự tham gia hợp tác giữa chính phủ và các tổ chức phi chính phủ- kể cả các tổ chức phi chính phủ trong nước và quốc tế- và cộng đồng. Do là một chương trình đa lĩnh vực, liên quan đến sức khoẻ, dinh dưỡng, giáo dục trẻ em, chăm sóc trẻ em, tiết kiệm và tạo thu nhập, nên sẽ có sự tham gia của nhiều bộ ngành trong chính phủ và của các tổ chức phi chính phủ nhằm cung cấp một dịch vụ tổng hợp được định hướng vì sự phát triển của trẻ em. Dự án sẽ hỗ trợ cho nhiều lựa chọn- theo một danh mục chương trình- có liên quan tới nhu cầu của trẻ em trước tuổi đi học và gia đình của chúng.

#### **V. Các thành phần của dự án**

**Thành phần dự án 1- Chương trình hợp nhất về chăm sóc trẻ em của cộng đồng.** Thành phần này hỗ trợ các mục tiêu của chính phủ: (a) cải thiện sự nhận thức của cha mẹ trẻ em về những khía cạnh quan trọng về chăm sóc trẻ em, sự tăng trưởng và phát triển của trẻ em thông qua giáo dục cha mẹ, theo dõi sự tăng trưởng của trẻ em và khuyến khích, đào tạo và làm tăng sự nhạy cảm của trẻ em; và (b) tăng cường quyền lực của cộng đồng nhằm hỗ trợ trẻ em thông qua việc xây dựng năng lực, các kỹ năng tạo ra thu nhập, và các chương trình hỗ trợ. Mục tiêu là giảm tình trạng suy dinh dưỡng (cân nặng thấp theo độ tuổi) của trẻ em xuống một phần ba vào cuối thời kỳ 5 năm tại các huyện có dự án và gia tăng mức độ sẵn sàng của trẻ em khi đi học tiểu học, và nhờ vậy, đóng góp cho mục tiêu giáo dục phổ cập tiểu học. Kế hoạch của chính phủ là sẽ thực hiện chương trình này ở tất cả các huyện; tuy vậy, chương trình sẽ

được thực hiện trong giai đoạn một ở 25 huyện do chính phủ lựa chọn căn cứ vào mức độ suy dinh dưỡng, tỷ lệ tử vong trẻ sơ sinh và tỷ lệ nhập học tiểu học. Dự án bao gồm các thành phần có liên quan với nhau sau đây:

(a) *Giáo dục cha mẹ.* Thành phần phụ này sẽ làm tăng sự hiểu biết của cha mẹ và những người chăm sóc trẻ em về các khía cạnh quan trọng trong chăm sóc trẻ em, sự tăng trưởng và phát triển của trẻ, bao gồm cả sự phát triển về dinh dưỡng, sức khoẻ, nhận thức và tâm lý của trẻ. Một loạt các năng lực có liên quan của cha mẹ trẻ em sẽ được tăng cường. Xây dựng các kỹ năng và kiến thức của cha mẹ sẽ cải thiện sự phát triển về sức khoẻ, tâm lý và phúc lợi của trẻ em, và cuối cùng sẽ làm tăng khả năng sẵn sàng đón nhận giáo dục ở cấp tiểu học. Chương trình sẽ huy động các nhóm bà mẹ (và cha mẹ) ở cấp độ cộng đồng, với sự trợ giúp của các tài liệu dự án bằng tiếng địa phương, sự giám sát về kỹ thuật và truyền thông. Các tài liệu giảng dạy được đơn giản hóa dành cho người lớn có trình độ học vấn thấp đã được thử nghiệm thành công ở Uganda. Dự án sẽ chú trọng nâng cao các hoạt động chăm sóc trẻ em có tác dụng khuyến khích sự tăng trưởng và phát triển phù hợp của trẻ em, kể cả dinh dưỡng và sức khoẻ trẻ em (nuôi con hoàn toàn bằng sữa mẹ và các biện pháp cai sữa- đặc biệt là thời kỳ giới thiệu các thực phẩm cai sữa, cũng như các loại thực phẩm, chuẩn bị thực phẩm, khuyến khích sự tăng trưởng của trẻ em và tẩy giun), phát triển tâm lý, khuyến khích nhận thức và sự hỗ trợ của xã hội, tăng cường vệ sinh và chăm sóc sức khoẻ tại nhà.

Các công việc trên sẽ được tăng cường và hỗ trợ bằng một hoạt động có quy mô hơn (ngày trẻ em) được tổ chức tại các địa phương nhằm cho phép các cộng đồng tiếp cận được một số dịch vụ liên quan đến trẻ em tại một địa điểm. Một nghiên cứu về tác động của cung ứng thuốc tẩy giun albendazole cho trẻ em ở những khu dân cư nhất định cũng sẽ được tiến hành trong các ngày trẻ em của các khu dân cư và sẽ đo lường tác động của chương trình tới cân nặng trẻ em sáu tháng một lần.

*(b) Xây dựng năng lực của cộng đồng và tăng cường năng lực chăm sóc trẻ em.* Thành phần phụ này bao gồm hai hoạt động có liên quan với nhau: (a) xây dựng năng lực cộng đồng được tiến hành thông qua các cuộc hội thảo về lập kế hoạch và tăng cường sự nhạy cảm của cộng đồng, và (b) huấn luyện về tinh thần kinh doanh nhằm gia tăng thu nhập của các bà mẹ và những người chăm sóc trẻ em.

**Thành phần dự án 2- Các khoản trợ giúp của cộng đồng cho Phát triển của trẻ em.** Có hai hình thức trợ giúp cho các cộng đồng:

*(a) Trợ cấp hỗ trợ cộng đồng-* các khoản trợ cấp cho các cộng đồng căn cứ vào sự đóng góp tương xứng của cộng đồng. Các khoản trợ cấp và các khoản đóng góp từ cộng đồng sẽ được sử dụng để hỗ trợ cho các chương trình vì sự phát triển của trẻ em nằm trong khuôn khổ hướng dẫn và danh mục chương trình trong tài liệu hướng dẫn thực hiện dự án. Để nhận được trợ cấp, các cộng đồng sẽ phải đóng góp phần của mình, có thể dưới hình thức hàng hoá, công việc hay dịch vụ. Ví dụ về việc sử dụng các khoản trợ cấp này là xây dựng và vận hành các trung tâm chăm sóc trẻ em của cộng đồng, các trung tâm chăm sóc trẻ em tại nhà, hay sản xuất và bán ra thị trường các thực phẩm cai sữa. Thành phần trợ cấp này sẽ được thực hiện tại 25 huyện ở trong thành phần 1.

*(b) Trợ cấp cho đổi mới-* các khoản trợ cấp này được cung cấp cho các cộng đồng để giải quyết các vấn đề liên quan đến trẻ em. Trợ cấp đổi mới sẽ hỗ trợ cho việc thực hiện chương trình ngoài danh mục các chương trình được mô tả trong phần (a) trợ cấp hỗ trợ cộng đồng ở trên. Như cái tên cho thấy, khoản trợ cấp “đổi mới” này sẽ được sử dụng để hỗ trợ các cộng đồng ở những cấp khác nhau trong việc thực hiện “các ý tưởng đổi mới” nhằm cải thiện cuộc sống của trẻ em ở trong cộng đồng của mình. Trợ cấp đổi mới sẽ được các cộng đồng tiếp cận theo cách thức giống như trợ cấp hỗ trợ cộng đồng. Các cộng đồng sẽ chuẩn bị các kế hoạch sau khi đã có được sự tham gia rộng rãi vào tiến trình lập kế hoạch. Các kế hoạch này sẽ được một

uỷ ban cấp tiểu hạt sàng lọc, và sau đó sẽ được xúc tiến tài trợ của dự án.

**Thành phần dự án thứ 3- Chương trình hỗ trợ quốc gia vì sự phát triển của trẻ em.** Thành phần này bao gồm các hoạt động chương trình trung tâm và các sáng kiến chính sách nhằm hỗ trợ các chương trình cấp huyện trong thành phần 1 và 2, và đảm bảo về chất lượng cho hoạt động của các dự án tiền tiêu tại cấp độ cộng đồng. Thành phần này bao gồm (a) theo dõi và đánh giá chương trình, (b) hỗ trợ ngăn ngừa hiện tượng thiếu dinh dưỡng, (c) phát triển chương trình học ECD, (d) huấn luyện những người giảng dạy chương trình ECD, và (e) thông tin, giáo dục, và truyền thông.

## VI. Tổ chức thực hiện

Việc thực hiện dự án là trách nhiệm của chính phủ Uganda với sự hỗ trợ của các tổ chức phi chính phủ trong khuôn khổ phân quyền và uỷ quyền tới cấp thấp hơn như được nêu ra trong các chính sách quốc gia. Cộng đồng (LC-1) là đơn vị vận hành cung cấp dịch vụ, mặc dù cấu trúc phối hợp cũng có sự tham gia tại cấp độ (LC-2), tiểu khu (LC-3), và huyện (LC-5).

Thêm vào đó, dự án hy vọng sử dụng sự nhạy cảm và tham vấn của những người liên quan, huy động cộng đồng, lập kế hoạch của cộng đồng với sự tham dự rộng rãi, xây dựng năng lực bền vững ở tất cả các cấp, cùng với tăng cường sự hỗ trợ từ các chương trình và cơ cấu quốc gia hiện đang tồn tại. Các cơ cấu chính trị và thể chế cần được sử dụng như là những cơ cấu đa lĩnh vực. Tính minh bạch và trách nhiệm giải trình cũng cần phải có ở tất cả các cấp.

## VII. Phối hợp dự án

**Cấp quốc gia.** Một uỷ ban chỉ đạo dự án bao gồm các bộ chủ

quản, các nhà tài trợ, và đại diện của các tổ chức phi chính phủ sẽ chịu trách nhiệm chỉ dẫn chung cho việc thực hiện dự án.

Uỷ ban này, do thư ký thường trực trong Bộ tài chính, Kế hoạch và Phát triển kinh tế lãnh đạo, sẽ chỉ dẫn cho dự án về các vấn đề chính sách, đánh giá lại, và chấp thuận chất lượng và hiệu quả của việc thực hiện dự án. Uỷ ban chỉ đạo dự án cũng sẽ đưa ra các đề xuất cải thiện ngân sách hàng năm của huyện và soạn thảo các kế hoạch của dự án.

Một văn phòng điều phối dự án nhỏ (PCO), bao gồm một nhà điều phối, một phó chủ tịch điều phối, một kế toán viên có năng lực, và một cán bộ hỗ trợ phụ, sẽ trực thuộc bộ phận phụ trách các dịch vụ xã hội của Bộ tài chính, Kế hoạch và Phát triển kinh tế, và sẽ chịu trách nhiệm điều phối hàng ngày các hoạt động của dự án ở cấp quốc gia.

**Cấp huyện.** Một uỷ ban đa lĩnh vực đang tồn tại- uỷ ban điều phối cấp huyện (DCC) - sẽ do Trưởng Văn phòng điều hành (CAO) chỉ định đảm nhận trách nhiệm điều phối dự án ở cấp huyện. CAO cũng lựa chọn một người trong số các viên chức chính phủ chịu trách nhiệm điều phối các hoạt động liên quan tới các tổ chức phi chính phủ và các hoạt động khác trong dự án. NGO dẫn đầu sẽ là một thành viên của DCC. Tại các huyện không có tổ chức phi chính phủ nào có đủ năng lực và kỹ năng phù hợp hoặc khi không thể cung cấp các tổ chức phi chính phủ này để chúng đảm nhận vai trò NGO dẫn đầu thì việc thực hiện sẽ được tiến hành thông qua chính quyền huyện.

**Cấp tiểu hụt.** Một uỷ ban khu vực đang tồn tại tương tự như ở cấp huyện sẽ chịu trách nhiệm điều phối các hoạt động dự án ở cấp tiểu hụt. Uỷ ban điều phối tiểu hụt cũng tạo điều kiện cho các mối liên kết giữa các cơ cấu đang tồn tại và các cơ cấu của dự án, và cùng với NGO dẫn đầu trong huyện, chấp thuận kế hoạch công việc hàng năm và yêu cầu tài trợ dự án trong tiểu hụt của NGO/CBO (CBO là các tổ chức dựa trên cộng đồng – Community-Based Organizations).

### VIII. Đánh giá tác động dự án

Chính phủ Uganda quan tâm tới việc đánh giá tác động các khía cạnh khác nhau của dự án nhằm đảm bảo tính hiệu quả của dự án, và hướng dẫn cho thiết kế của các dự án NECD sau này. Hơn nữa, do Ngân hàng Thế giới xem dự án này như là một nguồn thông tin cho các nước khác liên quan tới dịch vụ NECD, nên Ngân hàng Thế giới đã đưa dự án này vào trong đánh giá tại ba nước về các chương trình ECD, và sẽ cung cấp trợ giúp kỹ thuật trên cơ sở viện trợ cho PCO nhằm hỗ trợ cho các hoạt động nghiên cứu khoa học. Cụ thể là, hai nghiên cứu nhằm đánh giá tác động của các dự án cụ thể sẽ được tiến hành như là một phần trong dự án chung. Hai nghiên cứu này là:

- Nghiên cứu ngày sức khoẻ trẻ em trong xã nhằm đánh giá tầm bao phủ của các biện pháp tẩy giun được thực hiện trong các ngày sức khoẻ trẻ em ở cấp xã và tác động của chúng tới sự gia tăng cân nặng của trẻ em dưới sáu tuổi. Nghiên cứu này sử dụng thiết kế thử nghiệm lựa chọn ngẫu nhiên.
- Nghiên cứu điều tra sử dụng phương pháp điều tra gốc và điều tra lại để đánh giá (a) tác động của việc sử dụng các biện pháp tẩy giun và các hoạt động dự án tới sự phát triển nhận thức, sức khoẻ và dinh dưỡng của trẻ dưới sáu tuổi; (b) tác động của trình độ giáo dục của những người chăm sóc trẻ em và chiến dịch truyền thông qua phương tiện thông tin đại chúng tới kiến thức, thái độ và thực tế nuôi dưỡng trẻ của những người chăm sóc trẻ chủ yếu; và (c) tác động của đào tạo quản lý ở cấp cơ sở, các hoạt động tạo ra thu nhập và hình thành nhóm tiết kiệm tín dụng, và việc cung cấp trợ cấp của cộng đồng tới phúc lợi của hộ gia đình và của cộng đồng.

Hãng tư vấn được lựa chọn sẽ cung cấp hỗ trợ về kỹ thuật và hậu cần cho các nghiên cứu nêu trên và sẽ được mời tham dự với tư cách người thực hiện nghiên cứu địa phương vào trong quá trình thiết kế, thu thập dữ liệu và phân tích cần thiết để hoàn thành hai nghiên cứu đánh giá tác động. Hãng này sẽ

là đối tác chính của PCO, các nhà nghiên cứu trong nước, và các nhà nghiên cứu từ Ngân hàng Thế giới và Đại học Oxford, những người sẽ tiến hành đánh giá tác động.

## IX. Tổng quan về nghiên cứu

**Nghiên cứu thứ nhất:** tác động của tẩy giun vào các Ngày của Trẻ em ở cấp xã. Đã có một số nghiên cứu chỉ ra tác động của việc sử dụng các loại thuốc tẩy giun tới trẻ em trong độ tuổi đi học. Tuy nhiên, chỉ có một thử nghiệm có quy mô lớn, lựa chọn mẫu ngẫu nhiên là chỉ ra tác động lớn tới cân nặng của trẻ em chưa đi học. Điều này làm nảy sinh câu hỏi liệu có một tác động như vậy đối với trẻ em châu Phi không. Do đó, dự án NECD sẽ bao gồm một nghiên cứu có mẫu được lựa chọn ngẫu nhiên về tác động của việc sử dụng các biện pháp tẩy giun cho trẻ em ở 25 xã trong các ngày trẻ em và đo lường tác động của chương trình được tiến hành sáu tháng một lần này tới cân nặng. Các dữ liệu sẽ được thu thập từ các xã này và từ 25 xã đối chứng là các xã cũng tổ chức ngày sức khoẻ trẻ em nhưng không tiến hành tẩy giun trẻ em một cách định kỳ. Nếu việc tẩy giun được tiến hành thành công và chứng tỏ hiệu quả tích cực tới trẻ em Uganda, thì chương trình này sẽ được khuyến nghị áp dụng cho tất cả các huyện.

Bởi vì đây là một thí nghiệm khoa học có sự kiểm soát nên PCO lựa chọn các xã sẽ tiến hành các biện pháp tẩy giun từ một danh sách các xã sẽ tổ chức ngày chăm sóc trẻ em (danh sách này sẽ được các NGO hoạt động ở trong các huyện cung cấp). PCO cũng sẽ lựa chọn các xã đóng vai trò nhóm đối chứng. Thiết kế thử nghiệm này là rất quan trọng đối với thành công của đánh giá.

Hàng được lựa chọn sẽ đảm bảo các NGO địa phương chịu trách nhiệm tổ chức ngày sức khoẻ trẻ em tại các xã đều nhận thức được lý do của thiết kế thử nghiệm và tuân thủ chính sách này. Mỗi đứa trẻ từ 12 tháng tuổi trở lên và dưới sáu tuổi ở 25 xã được chỉ định sẽ nhận được 400 miligam thuốc albendazole

có thể nhai được và đã đăng ký sở hữu. Thuốc albendazole này sẽ được cung cấp sáu tháng mỗi lần; trong trường hợp các NGO lựa chọn tổ chức ngày sức khoẻ trẻ em một cách thường xuyên hơn, thì việc tẩy giun vẫn phải được tiến hành theo lịch trình sáu tháng chứ không được thường xuyên hơn.

Trẻ em ở các xã sử dụng albendazole cũng như trẻ em ở 25 xã đối chứng sẽ được cân nặng vào ngày trẻ em và số cân của chúng sẽ được ghi lại ở trên thẻ sức khoẻ của chúng và trong sổ đăng ký của xã. Những trẻ em quá nhỏ để có thể đứng trên cân một mình, sẽ được cân ở trên tay của mẹ sau khi đã đặt số đo ở cân bằng không, khi bà mẹ đứng lên cân một mình. Các số cân sẽ được ghi lại tính theo đơn vị phần mười (0,1) kilôgam. Các NGO địa phương chịu trách nhiệm về dữ liệu trong các sổ đăng ký của xã mặc dù hằng được thuê sẽ làm việc cùng với các NGO để đảm bảo hệ thống thu thập dữ liệu phù hợp với tất cả các mục tiêu của nghiên cứu.

Hàng tư vấn sẽ chuyển số liệu về cân nặng này vào một bảng kê khai được thiết kế trên cơ sở phối hợp với các cố vấn kỹ thuật của Ngân hàng Thế giới và Đại học Oxford. Việc chuyển dữ liệu này sẽ được tiến hành sáu tháng một lần sau ngày trẻ em ở trong khu vực có dự án. Ngoài mã số của trẻ (là sự kết hợp duy nhất giữa mã số của xã, mã số của làng, và mã số cá nhân được ghi trên thẻ của trẻ và sổ đăng ký của cộng đồng), dữ liệu trên bảng kê khai sẽ bao gồm giới tính của trẻ, ngày sinh của trẻ, được lấy ra từ thẻ sức khoẻ của trẻ, hoặc nếu như không rõ ngày sinh thì là tuổi của trẻ căn cứ vào sổ đăng ký của xã; ngày cân nặng trẻ; và việc trẻ có uống albendazole không. Các dữ liệu này sẽ được đưa vào máy tính ở Kampala. Mã số cá nhân sẽ tạo ra cơ sở kết hợp dữ liệu từ những thời kỳ khác nhau. Do đó phải ghi lại mã số mỗi lần chuyển dữ liệu, và mã số của mỗi đứa trẻ phải không thay đổi trong toàn bộ thời kỳ.

Các tình huống và điều kiện của địa phương vào ngày sức khoẻ trẻ em có thể ngăn cản các bà mẹ tham dự cũng được ghi lại. Những tình huống này bao gồm dữ liệu về tình trạng mùa

màng và điều kiện thời tiết, cả hai yếu tố này đều có thể ngăn cản các bà mẹ tham dự. Bất kỳ một phương pháp và cơ hội đặc biệt nào được sử dụng để quảng cáo ngày của trẻ cũng sẽ được ghi lại bởi vì những hình thức quảng cáo khác nhau có thể tác động đến sự tham dự. Cũng cần ghi lại ước lượng số trẻ em tham dự ngày của trẻ đến từ các xã khác, và những đứa trẻ không nhận được mã số trong ngày này.

Thử nghiệm này sẽ kéo dài trong hai năm. Do đó, hằng tư vấn sẽ ghi nhận lại dữ liệu năm lần ở mỗi xã. Điều này có nghĩa là hằng tư vấn sẽ thu thập dữ liệu vào đầu thời kỳ dự án và vào 6, 12, 18, và 24 tháng sau khi bắt đầu dự án.

Một bản sao toàn bộ dữ liệu sẽ được gửi tới PCO sau sáu tháng. Các bản sao dữ liệu này sẽ được xem như các dịch vụ có thể cung cấp của nghiên cứu thứ nhất trong dự án. Phân tích sơ bộ sẽ được tiến hành tại trường Đại học Oxford theo định kỳ nửa năm một lần. Tuy nhiên, hằng tư vấn được yêu cầu để cử một người đại diện, người sẽ tham dự vào việc thực hiện phân tích chính vào cuối năm thứ hai. Người đại diện này sẽ được cung cấp các chi phí đi lại và ăn ở để tiến hành phân tích tại Oxford. Nguồn tài chính cho việc đi lại này được hạch toán riêng, và do đó, không cần đưa vào trong hợp đồng mời thầu.

**Nghiên cứu thứ hai:** Tác động chung của các chương trình NECD. Các điều tra hộ gia đình và điều tra cộng đồng sẽ thu thập thông tin cơ sở và thông tin tiếp theo cần thiết để đánh giá tác động của các hoạt động dự án. Các điều tra này sẽ có một số phần, trong đó sẽ đo lường:

- Sự phát triển và tăng trưởng về nhận thức của trẻ em dưới sáu tuổi nhờ biện pháp tẩy giun và nhờ các hoạt động dự án nói chung. Nghiên cứu Thứ hai sẽ đánh giá sự tăng trưởng theo thời gian và các kết quả phát triển về tâm lý và nhận thức trong một nhóm trẻ em tại các cộng đồng tham gia dự án (có và không có biện pháp tẩy giun) so với một nhóm trẻ em ở các cộng đồng không tham gia dự án. Cả hai nhóm này sẽ được theo dõi trong hai năm, hoặc lâu hơn. Nghiên cứu

này sẽ bổ sung cho nghiên cứu về biện pháp tẩy giun ở cấp xã, do nó giúp hiểu rõ hơn về quyết định đưa trẻ em tham gia ngày của trẻ và đo lường liệu sự tham gia đó có dẫn đến sự gia tăng trong các chỉ số phát triển nhận thức của trẻ không. Hơn nữa, bằng cách bao gồm cả các cộng đồng không nhận được bất kỳ dịch vụ ECD nào, nghiên cứu này sẽ đánh giá, liệu các dịch vụ cung cấp có làm cải thiện tình trạng dinh dưỡng và phát triển nhận thức hay không.

- Các thay đổi về kiến thức, thái độ và thực tế nuôi dưỡng trẻ em của những người chăm sóc trẻ nhờ nội dung giáo dục cha mẹ của dự án và chiến dịch truyền thông đại chúng.
- Cải thiện tình trạng sức khoẻ và dinh dưỡng của trẻ dưới sáu tuổi nhờ các hoạt động theo dõi tăng trưởng, y tế dự phòng và giáo dục dinh dưỡng, tẩy giun, và các hoạt động của dự án nói chung.
- Phúc lợi của hộ gia đình có được nhờ sự trợ cấp cộng đồng, đào tạo quản lý ở cấp cơ sở, các hoạt động tạo thu nhập, và hình thành nhóm tiết kiệm tín dụng.
- Các đặc điểm của cộng đồng và các thay đổi do hoạt động của dự án mà có thể tác động tới tình trạng của trẻ em trong thời kỳ của dự án.

**Lựa chọn mẫu.** Cơ sở cho nghiên cứu này sẽ là điều tra gốc được thu thập vào thời điểm cung cấp dịch vụ lần đầu cho cộng đồng và điều tra tiếp theo thu thập dữ liệu từ cùng các hộ gia đình này hai năm sau khi điều tra lần đầu. Một phần ba của mẫu được rút ra từ 25 xã trong nhóm tham dự (tham gia tẩy giun) và một phần ba khác từ nhóm đối chứng được nghiên cứu trong Nghiên cứu thứ nhất.Thêm vào đó, một phần ba mẫu sẽ được lấy từ các làng trong 25 xã ở cùng huyện với nhóm tham dự nhưng không nhận được các dịch vụ do dự án NECD cung cấp. Ba mươi hộ gia đình sẽ được chọn ra ở mỗi xã. Như vậy sẽ có 750 hộ gia đình cho mỗi lớp (tổng cộng là 2.250) tham gia vào lần điều tra đầu tiên. Do dự kiến có hiện tượng thu hẹp mẫu, lần điều tra lại dự kiến sẽ có ít hơn từ 5 đến 10 phần trăm.

Để thu thập mẫu trong các xã tham dự và đối chứng, tất cả các hộ gia đình ở mỗi xã (trung bình có khoảng 700 hộ gia đình trong một xã) sẽ được liệt kê, có thể do một thành viên trong cộng đồng. Danh sách này sẽ bao gồm tên của chủ hộ, chỉ số về vị trí của hộ gia đình và số trẻ em dưới sáu tuổi trong hộ. Danh sách này sẽ phục vụ cho hai mục đích. Thứ nhất, một mẫu bao gồm 30 hộ gia đình có ít nhất một đứa trẻ dưới sáu tuổi sẽ được lựa chọn một cách ngẫu nhiên. Thứ hai, tổng số trẻ em dưới sáu tuổi sẽ đóng vai trò là ước lượng cho mức độ bao phủ tiềm tàng tính theo số trẻ em tham dự ngày của trẻ em, và do đó góp phần xác định được tỷ lệ tham dự.

Bởi vì dự án NECD sẽ có ít mối liên hệ với các cộng đồng hiện không có hoạt động của NECD nên việc lựa chọn các hộ gia đình không nhận được dịch vụ ECD cần sử dụng cách chọn mẫu theo cụm (cluster sampling) để làm giảm chi phí liệt kê mẫu. Cụ thể, một tiểu hạt không có trong dự án cần được lựa chọn tương ứng với một tiểu hạt nằm trong nhóm tham dự, nhất là một tiểu hạt có vị trí gần gũi. Tất cả các xã trong số những tiểu hạt này cần phải được liệt kê, và người ta sẽ rút ra ngẫu nhiên 25 xã trong danh sách toàn bộ. Ở mỗi xã, sẽ lựa chọn ra hai làng cũng trên cơ sở từ một danh sách tất cả các làng ở trong xã. Bước này làm giảm số lượng các làng cần phải tiến hành điều tra dân số. Điều tra dân số - tương tự như điều tra được sử dụng trong các xã tham dự và xã đối chứng - sẽ tạo thành danh sách các hộ gia đình được sử dụng để rút ra một mẫu bao gồm 30 hộ gia đình cho mỗi xã. Nhóm này tạo thành tầng thứ ba của điều tra.

Điều tra gốc ban đầu cần được tiến hành vào giữa năm 1999. Thời điểm này căn cứ vào nhu cầu tìm hiểu về các tiểu hạt và các xã tại đó các NGO sẽ hoạt động để có được thiết kế mẫu dự kiến. Thời điểm này cũng căn cứ vào giả định là, việc lựa chọn và đào tạo các NGO dẫn đầu sẽ không được hoàn tất cho tới cuối năm 1998.

Tuy nhiên, sự phát triển và thử nghiệm trước bảng câu hỏi cần được tiến hành sớm hơn nhiều thời điểm này (đầu năm

1999) để có thể sẵn sàng thực hiện điều tra ngay khi các NGO đã xác định được các xã mà chúng sẽ làm việc. Do điều tra gốc cần được tiến hành trước khi lấy giun lần đầu, nên thời điểm lý tưởng của điều tra gốc sẽ đồng thời với thời điểm tổ chức cộng đồng ban đầu sẽ đưa đến ngày của trẻ em. Do mẫu gồm 30 gia đình ở mỗi xã tương đối nhỏ so với tổng thể dân cư, nên khó có khả năng là việc thu thập dữ liệu điều tra sẽ làm gián đoạn các hoạt động khác hay là một gánh nặng lớn cho cộng đồng. Việc thu thập dữ liệu ở nhóm đối chứng (nhóm có hoạt động của NGO nhưng không có chương trình lấy giun và nhóm không có cả hai) cần được tiến hành đồng thời với thu thập dữ liệu ở nhóm tham dự.

*Các công cụ điều tra.* Bảng câu hỏi cơ bản sẽ được sử dụng cho dự án điều tra là bảng câu hỏi hộ gia đình (thu thập dữ liệu ở cấp hộ gia đình và cá nhân), và bảng câu hỏi cộng đồng.

## X. Điều tra hộ gia đình

Dữ liệu hộ gia đình sẽ được thu thập theo một lịch trình được mã hoá trước. Lịch trình này sẽ được soạn thảo căn cứ vào mô hình của Điều tra mức sống dân cư đã được sử dụng ở hơn 30 nước. Bản thảo lần đầu sẽ được các nhà nghiên cứu từ Ngân hàng Thế giới cung cấp. Tuy nhiên, công cụ này sẽ được điều chỉnh cho phù hợp với nhu cầu đặc thù của dự án và điều kiện của địa phương bằng cách sử dụng các nhóm tập trung và một thủ tục kiểm tra trước do hằng tư vấn tiến hành. Bảng câu hỏi hộ gia đình sẽ bao gồm các phần nhằm thu thập dữ liệu về:

1. Các đặc điểm xã hội nhân khẩu học: Một bảng về các cá nhân sống ở hộ gia đình trong vòng 12 tháng qua, tuổi tác và giới tính của họ, cũng như tình hình đi học và loại công việc (nếu có) của họ. Các mã hoá sẽ cung cấp thông tin về cha mẹ của các trẻ em, nếu như có – nếu không có thì sẽ xác định liệu họ còn sống không. Một danh sách chi tiết về các tài sản sẽ được thu thập đóng vai trò một chỉ số thể hiện tình trạng kinh tế xã hội.

2. Kiến thức, thái độ và thực tế: Bảng câu hỏi cũng sẽ thu thập thông tin về kiến thức, thái độ và thực tế chăm sóc trẻ của những người chăm sóc chủ yếu.
3. Dữ liệu về nhân trắc học: Cân nặng của trẻ sẽ được ghi chép lại theo đơn vị phần mươi (0,1) kilôgam gần nhất của tất cả trẻ em dưới sáu tuổi bằng một cân điện tử sẽ được cung cấp.Thêm vào đó, các dữ liệu về chiều cao của trẻ từ hai đến sáu tuổi cũng sẽ được thu thập. Kiểm tra trước sẽ được sử dụng để xác định liệu có thể thu thập dữ liệu về cân nặng của các bà mẹ trẻ em không (nếu họ sống ở trong hộ gia đình).
4. Đánh giá nhận thức: Hàng tư vấn sẽ làm việc cùng với các nhà tư vấn nghiên cứu trong nước và quốc tế trong việc thu thập toàn bộ dữ liệu thực địa. Trong điều tra gốc, một kiểm nghiệm được công nhận quốc tế về sự phát triển nhận thức sẽ được tiến hành với các trẻ em trong độ tuổi từ 4 đến 5,99 tuổi. Kiểm nghiệm này cũng sẽ được tiến hành với cùng nhóm tuổi ở kỳ hai của điều tra, nhờ đó cho phép so sánh các nhóm trẻ. Hơn nữa, một tập hợp con các trẻ em trong độ tuổi từ 6 đến 7,99 vào thời điểm lần điều tra thứ hai cũng sẽ tham gia vào kiểm nghiệm này. (Bảng phụ lục 2.1 tóm tắt công việc này).

Hơn nữa, sự đánh giá kiến thức dựa vào nội dung cụ thể từ chương trình và sự đánh giá năng động có thể được xây dựng trong vòng điều tra thứ hai. Việc đưa những thước đo này vào sẽ được đánh giá trong vòng đời của dự án. Cuối cùng, một thước đo kết quả tại trường sẽ được xây dựng để đánh giá kiến thức đạt được trong năm đầu tiên đi học và được áp dụng cho một tập hợp con bao gồm các trẻ em có tuổi lớn hơn trong kỳ điều tra lại. Các kiểm nghiệm hiện có có thể cũng sẽ được điều chỉnh.

5. Sức khoẻ trẻ em: Số liệu về bệnh tật (bao gồm cả số bệnh và dạng triệu chứng bệnh, mức độ nghiêm trọng, độ dài thời gian), dạng thức tiếp cận và việc sử dụng các dịch vụ sức khoẻ, vệ sinh v.v.

6. Kinh tế hộ gia đình: Phương pháp tốt nhất để thu thập thông tin này sẽ được khai thác triệt để trong giai đoạn kiểm tra trước, và do các cố vấn từ PCO đánh giá trước khi kết thúc bảng câu hỏi. Các biến có thể bao gồm chi tiêu cho lương thực; các hoạt động chăn nuôi; tiêu dùng các sản phẩm làm tại nhà; chi tiêu không phải là lương thực; đặc điểm của nhà ở; tồn kho các hàng hoá lâu bền; việc làm; hoạt động kinh tế; thu nhập; đất đai, cây trồng và động vật; thu nhập từ các hoạt động của dự án; các doanh nghiệp gia đình; sở hữu tài sản; thông tin tín dụng và tiết kiệm về số tiền và hàng hoá vay và cho vay, nếu như tiền và hàng này được vay mượn trong 12 tháng trước đó; tiết kiệm và nợ ròng vào ngày phỏng vấn; thông tin về các khoản nợ, kể cả lịch trình trả nợ, lý do vay mượn và số lượng các khoản nợ từ cùng nguồn; và địa điểm của các khoản tiết kiệm, nếu có, kể cả ở ngân hàng, ngân hàng tiết kiệm mua nhà, ngân hàng tiết kiệm nông thôn, v.v. Thông tin này sẽ là một phần trong điều tra gốc và điều tra cuối cùng.

## XI. Điều tra cộng đồng

Các bảng câu hỏi của cộng đồng sẽ được sử dụng để thu thập thông tin về các điều kiện chung của địa phương cho tất cả các hộ gia đình sống ở trong vùng. Phương pháp tốt nhất để thu thập thông tin này sẽ được khai thác triệt để trong giai đoạn kiểm tra trước và đánh giá kết hợp với các cố vấn từ PCO trước khi kết thúc bảng câu hỏi. Các biến có thể bao gồm:

1. Thông tin nhân khẩu: số hộ gia đình, tổng số dân, số dân dưới sáu tuổi, các nhóm dân tộc và tôn giáo;
2. Thông tin kinh tế, kể cả các hoạt động kinh tế chủ yếu và dạng thức di cư để tìm kiếm việc làm;
3. Cơ sở hạ tầng: sự tiếp cận đường sá, điện, đường ống nước, chợ, ngân hàng, và giao thông công cộng. Điều kiện về cơ sở hạ tầng của địa phương, như đường sá, nguồn nhiên

liệu và nước, mức độ sẵn có của điện, và các phương tiện truyền thông;

4. Các đặc điểm và thực tiễn sản xuất nông nghiệp ở địa phương; các loại cây trồng trong cộng đồng, mức độ thường xuyên và thời điểm chúng được trồng và được thu hoạch, sản phẩm thu hoạch thường được bán như thế nào, và các dữ liệu định tính về mưa, điều kiện khí hậu và mùa vụ;
5. Giáo dục: số lượng và loại hình các trường mẫu giáo, các thoả thuận ECD chính thức và phi chính thức, khoảng cách tới trường, số lượng các lớp học, tỷ lệ nhập học (tổng cộng và chia theo giới), việc tham dự lớp học, kết quả điểm thi, các dịch vụ y tế và dinh dưỡng được cung cấp ở trường (chẳng hạn, các chương trình sức khoẻ ở trường, ăn trưa tại trường);
6. Sức khoẻ: loại cơ sở y tế và khoảng cách và thời gian đi lại tới cơ sở y tế gần nhất trong số một vài dạng cơ sở y tế (bệnh viện, hiệu thuốc, trạm y tế v.v). Khoảng cách và thời gian đi lại tới cán bộ y tế gần nhất trong số một vài dạng (bác sĩ, y tá, dược sỹ, bà đỡ, cán bộ y tế cộng đồng, v.v); và
7. Các biến khác: số lượng và kiểu các NGO/CBO địa phương, các dự án hay chương trình có liên quan tới trẻ em khác (ví dụ, chương trình tiêm chủng của chính phủ), và các dự án phát triển cộng đồng khác.

**Dự kiến về cán bộ điều tra.** Cán bộ điều tra có thể bao gồm:

- Cán bộ điều tra cơ bản: gồm người quản lý điều tra, người quản lý thực địa, người quản lý dữ liệu, và cán bộ nhập dữ liệu, những người này chịu trách nhiệm trong công việc giám sát chung thực địa, phối hợp và theo dõi quá trình thu thập, nhập dữ liệu và các hoạt động quản lý dữ liệu.
- Cán bộ điều tra thực địa: các hoạt động thực địa sẽ được thực hiện bởi các nhóm bao gồm một giám sát viên, hai (hoặc ba) cán bộ phỏng vấn chịu trách nhiệm về bảng câu hỏi chính và các kết quả đo lường nhân trắc học, và một lái xe. Số lượng tương tự của các chuyên gia sẽ tham dự vào việc thực hiện

các kiểm định về sự phát triển nhận thức đối với trẻ em sẽ được lựa chọn và đào tạo trên cơ sở phối hợp với các chuyên gia trong nước và quốc tế.

- Điều phối viên lần thi điểm ngẫu nhiên: điều phối viên sẽ hỗ trợ trong việc phát triển các công cụ thu thập dữ liệu, huấn luyện về thiết kế thử nghiệm cho các NGO trong nước chịu trách nhiệm tổ chức ngày của trẻ em tại các xã về các công việc thiết kế thử nghiệm, thu thập dữ liệu và chuyển dữ liệu. Người này sẽ giám sát việc nhập dữ liệu, quản lý tập hợp các dữ liệu nghiên cứu, và sẽ tham dự vào công việc phân tích chính sẽ được thực hiện ở cuối nghiên cứu này.

**Tổ chức công việc thực địa.** Hàng tư vấn sẽ tham dự vào việc soạn thảo các công cụ thực địa trước khi kiểm tra thử điều tra và sẽ chịu trách nhiệm chính cho kỳ kiểm tra thử. Sau khi kiểm tra thử, bảng câu hỏi sẽ được thiết kế lại (trên cơ sở phối hợp với các nhà nghiên cứu của Ngân hàng Thế giới) và sau đó sẽ được dịch sang tiếng địa phương.

Hàng tư vấn sẽ làm việc cùng với các nhà tư vấn nghiên cứu trong nước và quốc tế được PCO lựa chọn để kết hợp các kiểm định về sự phát triển nhận thức của trẻ em vào trong công việc thu thập dữ liệu thực địa chung. Nhà nghiên cứu ECD địa phương, với sự hỗ trợ của các nhà tư vấn quốc tế, sẽ lựa chọn và thích nghi hoá thử nghiệm về nhận thức và sẽ huấn luyện những người tiến hành thử nghiệm.

Công việc thực địa được tổ chức như sau. Cách tổ chức này dựa trên kinh nghiệm quốc tế và được thiết kế để đảm bảo kiểm soát chất lượng. Một số hình thức biến thể của phương pháp này có thể được chấp nhận trên cơ sở tham vấn các nhà nghiên cứu từ Ngân hàng Thế giới, dựa trên kinh nghiệm của hàng tư vấn và các cố vấn khác của PCO và các thông tin thu nhận được trong kiểm tra thử.

Công việc thực địa sẽ được tổ chức thành các nhóm nhỏ bao gồm một giám sát viên, hai (hoặc ba) cán bộ phỏng vấn chịu trách nhiệm về bảng câu hỏi chính và các do lường nhân trắc

học, và số lượng tương tự các chuyên gia thực hiện kiểm tra sự phát triển nhận thức của trẻ em. Các cán bộ này sẽ được huấn luyện về kiểm tra tâm lý tại Kampala do nhà nghiên cứu của ECD kết hợp với các cố vấn quốc tế. Chương trình huấn luyện sẽ bao gồm việc thảo luận về mục tiêu nghiên cứu, đánh giá lại từng bước trong phỏng vấn, huấn luyện thực tế tại văn phòng, tập dượt trên thực địa, và tổng kết kinh nghiệm sau khi tập dượt.

Sau khi các nhóm đã được huấn luyện, cần duy trì các nhóm này trong toàn bộ đợt điều tra nếu có thể. Tuy nhiên, bởi vì có một số cán bộ có thể chứng tỏ không phù hợp với công việc thực địa, nên cần phải huấn luyện thêm một số cán bộ bổ sung. Nhìn chung, không nên thuê cán bộ làm việc trong vài ngày chỉ ở một xã và sau đó lại thuê cán bộ mới ở xã khác, bởi vì, nếu làm như vậy, dự án sẽ chỉ có các cán bộ thiếu kinh nghiệm. Tất cả cán bộ cần nhận được sự đào tạo mới vào đầu đợt điều tra lại.

Trong khi thực hiện kiểm tra nhận thức, nên để trẻ em ở lại một mình với người phỏng vấn càng nhiều càng tốt. Một người khác (người lớn hay trẻ em) không bao giờ nên trả lời câu hỏi dành cho đứa trẻ đang được phỏng vấn. Tuy nhiên, trong kỳ điều tra lại, việc kiểm tra một nhóm nhỏ gồm các trẻ em tám tuổi có thể được thực hiện theo nhóm nếu như điều này thuận tiện.

Giám sát viên sẽ chịu trách nhiệm đảm bảo người phỏng vấn thực sự có tiến hành điều tra tại hộ gia đình đã được lựa chọn cho mẫu mà không thay thế bởi một hộ nào khác, và tất cả trẻ em trong độ tuổi phù hợp đều được kiểm tra mức độ phát triển nhận thức. Hơn nữa, giám sát viên sẽ kiểm tra lại mỗi bảng câu hỏi sau khi hoàn thành (trước khi nhóm đánh giá chuyển sang một làng khác) để đảm bảo cho bảng câu hỏi không có chỗ trống và kiểm chứng các thông tin có vẻ không nhất quán.

Hãng tư vấn sẽ nhập tất cả các dữ liệu điều tra sau khi thu thập, càng sớm càng tốt. Các bản sao về dữ liệu của hộ gia đình và trẻ em và thang bậc xếp hạng cùng với tài liệu cần thiết để tiếp cận dữ liệu, sẽ được cung cấp cho PCO dưới dạng máy tính

vào cuối thời kỳ điều tra gốc. Bảng câu hỏi ban đầu cần được hăng tư vấn lưu giữ do nhu cầu tiếp cận các dữ liệu gốc trong quá trình phân tích.

Dữ liệu cấp độ trẻ em phải bao gồm các mã số nhận diện chính xác. Các mã số này có thể được so sánh tương đương với các mã số của điều tra hộ gia đình. Mặc dù các mã số cá nhân và hộ gia đình riêng biệt được cung cấp cho PCO không nhất thiết phải bao gồm tên của hộ gia đình hay địa điểm chính xác của họ, nhưng thông tin này cần được hăng lưu trữ dưới hình thức nhất định để có thể thăm lại hộ gia đình này vào một thời điểm sau này. Do một bước trong phân tích sẽ liên kết các cá nhân trong lần điều tra lại với các kết quả kiểm định của họ trong phương án gốc, nên tất cả các mã số cá nhân và hộ gia đình cần phải được giữ nguyên trong ba đợt điều tra.

## XII. Các công việc cụ thể cho chuyên gia điều tra

Hăng tư vấn sẽ tham dự vào các hoạt động sau đây trên cơ sở kết hợp với PCO, các nhà nghiên cứu địa phương, và các nhà nghiên cứu từ Ngân hàng Thế giới và Đại học Oxford, cùng với các NGO thực hiện chương trình:

- Chỉnh sửa các chương trình công việc.
- Phát triển và thích nghi hóa các công cụ thu thập dữ liệu và các tài liệu hỗ trợ, kể cả các tài liệu liệt kê, bảng câu hỏi, chỉ dẫn mã hóa, tài liệu hướng dẫn người phỏng vấn và người giám sát, tài liệu hướng dẫn vận hành, tài liệu hướng dẫn nhập dữ liệu, và các thủ tục thực địa.
- Chỉnh sửa các bản thảo khác nhau của các tài liệu, trình bày, dịch tài liệu, dịch ngược, và kiểm tra thực địa. Cung cấp bản sao trên giấy và bản trên máy tính của tất cả các tài liệu cho PCO. Chuyển bảng câu hỏi tới các nhà nghiên cứu của Ngân hàng Thế giới để xem xét lại và chỉnh sửa trước khi tiến hành kiểm tra thí điểm.
- Liệt kê nơi cư trú và cập nhật bản đồ. Trách nhiệm liệt kê

các hộ gia đình và nơi cư trú ở mỗi xã được lựa chọn, bao gồm việc có được các bản đồ cơ sở, chuẩn bị các tài liệu liệt kê, liên hệ với các viên chức địa phương để thông báo cho họ về hoạt động liệt kê, xác định ranh giới, vẽ bản đồ, liệt kê các hộ gia đình một cách có hệ thống, đạt được các thông tin sơ bộ về các hộ gia đình, kể cả tên của chủ hộ, địa điểm hộ gia đình, và số lượng trẻ em dưới sáu tuổi trong hộ; các thủ tục về tài liệu vào thời điểm thiết kế mẫu, vào cuối thời kỳ công việc hiện trường, và khi hoàn thành hồ sơ dữ liệu.

- Chuẩn bị khung chọn mẫu (cùng với chuyên gia chọn mẫu), huấn luyện các cán bộ thực hiện mẫu đã chọn, giám sát giai đoạn thực hiện nhằm đảm bảo chất lượng của mẫu được lựa chọn, và cung cấp báo cáo chi tiết nêu rõ tất cả các bước liên quan trong thiết kế và thực hiện mẫu.
- Với sự tư vấn của Ngân hàng Thế giới, tham dự vào việc xác định một chiến lược phù hợp cho xác định nhóm đối chứng (tức là các xã không tham gia dự án).
- Lựa chọn và huấn luyện các cán bộ thực địa. Hoạt động này bao gồm tất cả các công việc cần thiết nhằm phát triển các tài liệu huấn luyện và các tài liệu hướng dẫn cho tất cả những người tham gia công việc tại thực địa. Tham dự các khoá huấn luyện là yêu cầu bắt buộc đối với tất cả cán bộ phỏng vấn, người giám sát cán bộ phỏng vấn, người nhập dữ liệu, và người thực hiện do lường nhân trắc học.
- Công việc thực địa, kể cả các công việc hậu cần phục vụ công tác thu thập dữ liệu và đạt được sự chấp thuận của hộ gia đình và cá nhân; lưu giữ bản đăng ký các hộ gia đình tham gia nghiên cứu.
- Cung cấp các báo cáo tiến độ: Hàng tư vấn sẽ chuẩn bị báo cáo tiến độ công việc tại thực địa (sáu tháng một lần) gửi cho PCO và Ngân hàng Thế giới. Hàng này cũng sẽ chuẩn bị một bản mô tả cơ bản về điều tra. Bản mô tả này cần bao gồm nội dung điều tra, kế hoạch và việc thực hiện chọn mẫu và các kỹ thuật thực địa được sử dụng. Một bảng câu hỏi đầy đủ và tài liệu cơ bản cũng được kèm theo như là phần phụ lục.

- Phát triển một chương trình nhập dữ liệu sử dụng phần mềm có thể kiểm tra phạm vi và tính chất quan của dữ liệu, và cung cấp báo cáo chỉ rõ những dữ liệu bị thiếu, các dữ liệu nằm ngoài phạm vi cho phép, và các câu trả lời thiếu nhất quan.
- Làm sạch dữ liệu (data cleaning), nhập dữ liệu, quản lý cơ sở dữ liệu và các kế hoạch lập bảng. Bao gồm việc phát triển một chương trình nhập dữ liệu, các tài liệu hướng dẫn nhập dữ liệu, huấn luyện người nhập dữ liệu, kiểm tra chất lượng dữ liệu, và hướng dẫn sử dụng dữ liệu. Đồng thời, mã hoá các câu hỏi có kết mở, kiểm chứng dữ liệu, kiểm tra dữ liệu nhân trắc học so với các bảng tham chiếu tiêu chuẩn.
- Đảm bảo thi hành thoả thuận chính sách sử dụng dữ liệu: Hàng tư vấn và nhà nghiên cứu tham gia vào tiến trình thu thập và phân tích dữ liệu sẽ ký kết một bản ghi nhớ với PCO, trong đó nêu rõ chính sách liên quan tới các vấn đề như sự tiếp cận dữ liệu, người sử dụng dự kiến, các thủ tục để có được bản sao của tập hợp dữ liệu và tài liệu nghiên cứu, và các quy tắc xuất bản và bản quyền.
- Tiến hành phân tích dữ liệu: Hàng tư vấn sẽ tiến hành phân tích dữ liệu mang tính chất tìm hiểu (ví dụ, phân tích về tần số, các bảng phân trăm, các bảng chéo) của những biến điều tra cơ bản và sự tương quan của chúng. Hàng này sẽ tiến hành xây dựng mô hình thống kê hiện đại của tác động sau đợt điều tra thứ hai và thứ ba nhằm xác định những tiến triển chung trong các chỉ tiêu xã hội (ví dụ, dinh dưỡng, sức khoẻ, thu nhập, và sự phát triển cộng đồng) và các nhân tố sẽ giải thích cho sự thay đổi hay cho sự ít thay đổi của các chỉ tiêu này.
- Cung cấp báo cáo phân tích: Hàng tư vấn sẽ báo cáo về các kết quả sau đợt điều tra thứ hai và thứ ba trên cơ sở phân tích các chỉ số xã hội và tương quan của chúng. Hàng này sẽ phối hợp với PCO và Ngân hàng Thế giới trong Nghiên cứu ngày sức khoẻ trẻ em ở các xã và trong thu thập dữ liệu về tác động tới sự phát triển nhận thức, nhưng sẽ không

chịu trách nhiệm về báo cáo kết quả cuối cùng của các nghiên cứu này.

Các công việc cụ thể trong điều tra cộng đồng bao gồm:

- Làm việc với các cố vấn từ PCO trong việc phát triển bảng câu hỏi cộng đồng và tích cực khám phá từ giai đoạn kiểm tra trước về phương pháp tốt nhất nhằm thu thập thông tin này;
- Làm việc chặt chẽ với các cơ quan thực hiện (các NGO dẫn đầu và địa phương) trong việc thu thập dữ liệu của cộng đồng;
- Liên lạc với các viên chức địa phương và những người lãnh đạo cộng đồng để giải thích về phương pháp đánh giá tác động dự án và đạt được sự nhất trí của xã về nghiên cứu điều tra và nghiên cứu về ngày sức khoẻ trẻ em;
- Phỏng vấn những người cung cấp thông tin quan trọng và đạt được các bản đồ, danh sách, và các sổ sách khác của cộng đồng;
- Đạt được danh sách về các cơ sở y tế và giáo dục (các trường mẫu giáo và tiểu học), bao gồm vị trí địa lý, khu vực phục vụ, và hình thức của cơ sở (ví dụ, của tư nhân hay của nhà nước);
- Đạt được thông tin về nhân khẩu của cộng đồng, bao gồm số lượng hộ gia đình và số dân theo giới tính và độ tuổi; và
- Đạt được các dữ liệu khác cần thiết trong bảng câu hỏi cộng đồng.

Các công việc cụ thể cho nghiên cứu về ngày của trẻ em bao gồm:

- Tham dự vào việc xây dựng nghị định thư về nghiên cứu;
- Phát triển các công cụ thu thập dữ liệu;
- Huấn luyện các NGO địa phương chịu trách nhiệm tổ chức ngày của trẻ em trong các xã về thiết kế thử nghiệm;
- Giám sát dữ liệu thu thập được trong ngày của trẻ em;
- Chuyển thể dữ liệu;
- Nhập và quản lý dữ liệu, và

- Tham dự vào phân tích chính được thực hiện vào cuối nghiên cứu.

**Bảng phụ lục 2.1 Quy mô mẫu dự kiến cho đánh giá tác động của Dự án dinh dưỡng và phát triển trẻ em ở Uganda**

Khoản mục	Tẩy giun và có giáo dục cho phụ huynh	Không tẩy giun và có giáo dục cho phụ huynh	Không tẩy giun và không giáo dục cho phụ huynh	Tổng số
Số xã	25	25	25	
Thời gian	Cơ sở Đợt hai*	Cơ sở Đợt hai	Cơ sở Đợt hai	
Số hộ gia đình	750	700	750	2250
Số trẻ em được cân nặng vào ngày trẻ em <sup>b</sup>	5000	5000	5000	20000
Số trẻ em có độ tuổi từ 0 đến 5,99 được tẩy giun tại nhà (có nghĩa là hai trẻ ở mỗi gia đình) <sup>c</sup>	1500	1395	1500	11580
Số trẻ em tham gia kiểm tra về nhận thức: kiểm tra tất cả trẻ em trong độ tuổi từ 4 đến 5,99 tại các hộ gia đình	500 <sup>d</sup>	465 <sup>e</sup>	500	465
Số trẻ em độ tuổi từ 6,0 đến 7,99 tham gia kiểm tra nhận thức và nhân trắc học	Tập con <sup>f</sup>	- Tập con	-Tập con	Tập con
Tỷ lệ nhập học	25 công đồng	25 công đồng	25 công đồng	25 công đồng

a. Giả định giảm đi một phần nhỏ do mẫu bị tiêu hao bằng 8 phần trăm trong hai năm.

b. Giả định có khoảng 200 trẻ em tham dự vào mỗi ngày của trẻ em.

c. Giả định có hai trẻ em ở mỗi gia đình, nhưng các gia đình vẫn được tuyển lựa nếu như có bất kỳ đứa con nào dưới sáu tuổi. Gia đình ở đây chỉ một cặp bao gồm một bà mẹ (hoặc người thay thế) và đứa trẻ.

d. Đây là số tối đa; số thực có thể khác tuỳ theo số trẻ em trong độ tuổi từ bốn đến năm tuổi trên thực tế.

e. Giả định mức tiêu hao mẫu 8 phần trăm trong hai năm: chỉ những trẻ em có bố mẹ được phỏng vấn mới được tham gia kiểm tra.

f. Con số này sẽ là một tập hợp con của các trẻ em trong độ tuổi này có cha mẹ được phỏng vấn. Chúng sẽ có liên kết với điểm số trước đó. Mặc dù số trẻ em được kiểm tra tăng lên trong đợt điều tra thứ hai, thời gian cho phỏng vấn có thể giảm đi, bởi vì sẽ không cần phải đánh giá lại nhiều thông tin nữa. Cũng có khả năng là quy mô của nhóm này sẽ bị giảm đi.

*Nguồn:* Tài liệu Dự án của Ngân hàng Thế giới.

**Nghiên cứu tính hợp lệ.** Ngoài các nghiên cứu trên, sẽ có một nghiên cứu theo chiều dài thời gian bổ sung để kiểm tra tính hợp lệ của thước đo trước khi đi học cho kết quả học tập vào cuối năm thứ nhất tại trường. Trong điều tra gốc, người ta sẽ tiến hành kiểm tra hai trẻ em ở mỗi cộng đồng trong độ tuổi từ 6,0 đến 6,9 (vẫn chưa đi học), với  $N = 150$ . Các em này sẽ được xác định sau khi kiểm tra và sẽ tham gia vào một kỳ kiểm tra kết quả ở trường sau hai năm, khi chúng ở độ tuổi từ 8,0 đến 8,99.

**Lịch trình công việc.** Thời gian biểu dự kiến cho lịch trình công việc sẽ như sau:

Tháng 1. Bắt đầu tiến trình xây dựng các chỉ số phát triển nhận thức trên cơ sở kết hợp với nhà tư vấn quốc tế và phù hợp với điều khoản tham chiếu. Tiến trình này có thể kéo dài tới sáu tháng.

Tháng 2. Kiểm tra thử lần đầu và chỉnh sửa lại bảng câu hỏi.

Tháng 5. Bắt đầu lên danh sách các hộ gia đình để lựa chọn mẫu. Bước này phụ thuộc vào việc lựa chọn các NGO dẫn đầu

và NGO địa phương. Chưa thể thực hiện được bước này cho tới khi PCO và các NGO lựa chọn được những xã sẽ tổ chức ngày của trẻ em và sau đó lựa chọn địa điểm cho chương trình tẩy giun đầu tiên. Đồng thời, cần phải dịch bảng câu hỏi và kiểm tra lần nữa tại thực địa.

Tháng 7. Bắt đầu thu thập dữ liệu tại ngày trẻ em cho nghiên cứu về tẩy giun. Các dữ liệu sẽ được thu thập theo định kỳ sáu tháng. Giống như ở trên, thời điểm của bước này phụ thuộc vào việc lựa chọn các NGO dẫn đầu và NGO địa phương.

Tháng 8. Huấn luyện các cán bộ thực địa về điều tra hộ gia đình, và bắt đầu điều tra. Điều tra này cần kéo dài từ ba đến bốn tháng, tùy thuộc vào số nhóm điều tra được tuyển dụng. Việc nhập số liệu cần diễn ra đồng thời với quá trình thu thập dữ liệu.

Tháng 14. Phân tích ban đầu đối với dữ liệu gốc. Đây sẽ là một quá trình tiếp diễn.

Tháng 20. Cán bộ của hãng tư vấn sẽ đến trường Đại học Oxford để tham dự vào sự phân tích dữ liệu ban đầu.

Tháng 20-36. Thu thập dữ liệu cho đợt hai của nghiên cứu về biện pháp tẩy giun.

Các điều tra hộ gia đình giữa kỳ và cuối cùng sẽ được tiến hành sau điều tra gốc hai năm và bốn năm.

### Sự hỗ trợ hãng tư vấn

Sẽ không có sự hỗ trợ cụ thể nào đối với hãng tư vấn để thực hiện công việc được giao. Các hãng được khuyến cáo cần nêu rõ tất cả các yêu cầu để thực hiện công việc một cách có hiệu quả trong đề xuất của mình.

**Ví dụ II: Đánh giá tác động  
của đường bộ nông thôn: Điều tra cơ sở  
ở Việt Nam năm 1997 \***

**Điều khoản tham chiếu: Nghiên cứu tác động  
của điều tra gốc về đường bộ nông thôn**

**I. Bối cảnh nghiên cứu**

Nghiên cứu này nhằm mục đích đánh giá tác động tới mức sống dân cư của dự án giao thông nông thôn ở Việt Nam do Ngân hàng Thế giới tài trợ. Dự án này đã được thực hiện ở 15 tỉnh nghèo trong vòng từ 3 đến 5 năm, kể từ năm 1997. Trọng tâm chính của nghiên cứu sẽ là việc các nhân tố quyết định mức sống người dân thay đổi theo thời gian như thế nào ở các xã có dự án đường so với các xã không có. Việc này đòi hỏi phải thu thập dữ liệu gốc trước dự án đối với cả hai khu vực: khu vực có dự án và khu vực không có dự án. Cũng cần có thêm một số đợt thu thập các dữ liệu sau khi có chương trình với tần suất hai năm một kỳ. Một cơ sở dữ liệu chi tiết cấp xã sẽ được tạo ra một phần bằng cách lấy ra từ các sổ sách được thu thập hàng năm ở cấp xã. Các sổ sách này sẽ được bổ sung bằng dữ liệu cấp xã trong quá khứ và việc thu thập nhiều loại dữ liệu bổ sung quan trọng khác. Một điều tra cấp huyện có thời gian ngắn sẽ giúp xác định bối cảnh của các dữ liệu cấp xã. Cuối cùng, sẽ có từ 10 đến 15 hộ gia đình được lựa chọn một cách ngẫu nhiên từ danh sách các hộ gia đình cấp xã, và một bảng câu hỏi ngắn gọn cho hộ gia đình sẽ được sử dụng. Nghiên cứu này sẽ được tiến hành ở 6 tỉnh trong số 15 tỉnh được hưởng lợi từ dự án. Sáu tỉnh này sẽ đại diện cho sáu khu vực

vực  
khoả  
đượ  
đồng  
Tư  
khoả

Chợ

1  
lý c  
tra s  
có n  
hàn  
2  
xã n  
hoặ  
các  
Mộ  
tỉnh  
chợ  
dan  
tỉnh  
tuy  
khü  
ngâ  
kết  
xã  
tha  
tươ

\* Điều khoản tham chiếu này do Dominique van de Walle chuẩn bị.

vực địa lý của Việt Nam. Một mẫu ngẫu nhiên bao gồm khoảng 200 xã tham gia dự án và không tham gia dự án, sẽ được chọn. Sáu nhóm điều tra sẽ được thành lập để điều tra đồng thời ở mỗi tỉnh. Điều tra này sẽ được bắt đầu vào tháng Tư và kết thúc vào khoảng tháng Tám. Dữ liệu sẽ có vào khoảng tháng Mười hoặc tháng Mười một.

## II. Thiết kế điều tra

**Chọn mẫu.** Việc chọn mẫu sẽ được tiến hành ở ba cấp độ.

1. Cấp tỉnh: Mười lăm tỉnh có dự án nằm ở sáu khu vực địa lý của Việt Nam. Tiêu chí để lựa chọn tỉnh sẽ tiến hành điều tra sẽ là: (a) mỗi khu vực địa lý sẽ chọn ra một tỉnh và (b) khi có nhiều hơn một tỉnh tham gia dự án ở mỗi khu vực thì sẽ tiến hành lựa chọn ngẫu nhiên.

2. Cấp xã: Mục đích là điều tra 200 xã hoặc nhiều hơn, các xã này được lựa chọn một cách ngẫu nhiên. Có khoảng một nửa hoặc ít hơn (không nhiều hơn) trong số những xã này phải là các xã có dự án đường, các xã còn lại đóng vai trò đối chứng. Một danh sách sẽ được rút ra từ các xã không có dự án ở sáu tỉnh (hoặc sẽ có một danh sách cho mỗi tỉnh) và người ta sẽ chọn mẫu ngẫu nhiên. Cũng tương tự, người ta sẽ lấy ra một danh sách từ tất cả các xã hưởng lợi nhờ dự án đường tại sáu tỉnh (hoặc trong từng tỉnh). Có thể có nhiều hơn một xã cho mỗi tuyến đường, và tất cả các xã này sẽ đều được đưa vào trong khung chọn mẫu. Từ khung này, mẫu sẽ được chọn một cách ngẫu nhiên. Mẫu sẽ không nhất thiết gồm cả hai xã được liên kết với nhau nhờ một dự án đường. Nếu việc tiếp cận tới các xã được chọn mẫu nhất định là không thể, thì xã này sẽ được thay thế bằng một xã khác ở trong huyện có những đặc điểm tương tự.

3. Cấp hộ gia đình: ở mỗi xã được chọn mẫu, sẽ tiến hành

điều tra hộ gia đình đối với 15 hộ gia đình. Các hộ này (cộng với một vài hộ thay thế có thể) sẽ được lựa chọn ngẫu nhiên từ danh sách các hộ gia đình của xã. Sau khi lựa chọn, những người điều tra sẽ hỏi chính quyền xã là các hộ này thuộc về nhóm hộ rất nghèo, nghèo, trung bình, không nghèo hay giàu.

### III. Tiến trình điều tra

Sáu chuyên gia điều tra sẽ được tuyển để tiến hành điều tra tại sáu tỉnh. Sau khi được huấn luyện và kiểm tra bảng câu hỏi trên thực địa, các chuyên gia này sẽ bắt đầu điều tra đồng thời ở mỗi tỉnh. Ở các huyện, những người điều tra sẽ cần có ít nhất một cán bộ địa phương từ ban quản lý dự án của huyện để giúp đỡ trong việc liên lạc với chính quyền địa phương, và trong một số trường hợp, để tìm kiếm những hướng dẫn viên và phiên dịch viên phù hợp tại các vùng dân tộc thiểu số. Nếu cần thiết, dự án sẽ thuê các trợ lý điều tra hay sự hỗ trợ từ ban quản lý dự án cấp tỉnh.

Mỗi người điều tra sẽ thu thập dữ liệu ở trung bình 35 xã, ở các huyện có các xã này, và với khoảng 15 hộ gia đình ở mỗi xã. Sẽ cần từ ba đến bốn ngày cho mỗi xã. Khoảng thời gian sử dụng trên thực địa sẽ từ 100 đến 140 ngày (bốn đến năm tháng). Tổng thời gian sẽ là sáu tháng.

Trong thời kỳ điều tra, giám sát viên sẽ đến thực địa ở tất cả sáu tỉnh để giám sát việc thu thập dữ liệu và để đảm bảo chất lượng dữ liệu tốt.

Dữ liệu được thu thập sẽ được làm sạch và được nhập vào máy tính bằng một chương trình nhập dữ liệu.

Bảng phụ lục 2.II.1 cung cấp thời gian biểu dự tính của nghiên cứu này.

**Bảng phụ lục 2.II.1 Thời gian biểu cho nghiên cứu  
danh giá tác động ở Việt Nam**

	Tháng 1	Tháng 2	Tháng 3	Tháng 4	Tháng 5	Tháng 6	Tháng 7	Tháng 8	Tháng 9
Thiết kế bảng câu hỏi	****	****							
Điều tra thử trên thực địa			****						
Kiểm điểm lại bảng câu hỏi				****					
Thích nghi chương trình nhập dữ liệu, dịch và in bảng câu hỏi					****				
Thuê và huấn luyện những người điều tra					****				
Điều tra tại thực địa					****	***	***	***	***
Kiểm tra dữ liệu								***	***
Thực hiện vào dữ liệu									****

#### IV. Các vấn đề khác

**Trang thiết bị.** Trang thiết bị mua sắm trong dự án sẽ thuộc về dự án chừng nào nghiên cứu còn tiếp tục (qua các kỳ điều tra trong tương lai), nhưng khi không được nhóm nghiên cứu sử dụng, thì sẽ được lưu giữ tại Đơn vị quản lý dự án để cho nhóm này sử dụng.

**Giải ngân.** Ngân sách cho nghiên cứu (đã trừ khoản thanh toán cho người điều tra chủ yếu, người này sẽ nhận được các khoản

trả dần hàng tháng) sẽ được giải ngân thành ba đợt. Đợt thứ nhất, sau khi ký hợp đồng, sẽ bao gồm 20 phần trăm tổng số tiền. Đợt thứ hai với số tiền bằng 50 phần trăm tổng ngân sách sẽ được giải ngân một khi các bảng câu hỏi cho xã, hộ gia đình, và huyện đã hoàn tất và được người quản lý công việc của Ngân hàng Thế giới chấp thuận. Công việc này dự kiến sẽ diễn ra vào cuối tháng Ba. Đợt thứ ba và là cuối cùng, sẽ được giải ngân vào cuối tháng Bảy hoặc ở giai đoạn giữa thông qua việc thu thập dữ liệu. Chi tiết về ngân sách dự tính được chỉ ra trong bảng phụ lục 2.II.2.

### Bảng phụ lục 2.II.2 Ngân sách nghiên cứu dự kiến

	Số	Thời gian	Số tiền (đôla Mỹ)	Tổng số đôla Mỹ
1. Nhà điều tra chính	1	9 tháng	1.000	9.000
2. Các chuyên gia điều tra	6	6 tháng	400	14.400
3. Chi phí đi lại cho sáu người điều tra, sáu người dẫn đường địa phương và phiên dịch	12	125 ngày	8	12.000
4. Ô tô và phương tiện khác cho sáu nhóm điều tra	6	125 ngày	40	30.000
Thuê xe cho nhà điều tra chính	1	30	50	1.500
5. Vé máy bay Hà Nội - thành phố Hồ Chí Minh - Hà Nội	6			
Cho các cán bộ điều tra (các tỉnh phía nam) 3 người			200	1.200
Cho nhà điều tra chính	3 chuyến			

(Xem tiếp trang sau)

**Bảng phụ lục 2.II.2 (Tiếp)**

	<i>Số</i>	<i>Thời gian</i>	<i>Số tiền</i> <i>(đôla Mỹ)</i>	<i>Tổng số</i> <i>đôla Mỹ</i>
6.	Chi trả cho đào tạo các cán bộ điều tra	12		1.338
	Đi lại tới thực địa	1 tuần	50	
	Công tác phí	3 ngày/3 ôtô	50	
		3 ngày	8	
7.	Kiểm tra bảng câu hỏi ở thực địa (các xã ở miền Nam và miền Bắc)	1	2 tuần	2.000
8.	Làm sạch và nhập dữ liệu	2	2 tháng	200
9.	Các tài liệu điều tra			2.000
10.	Phương tiện liên lạc (máy fax, điện thoại, email, Xerox)			2.000
11.	Trang thiết bị			5.000
	Máy tính (PMU 18)	1	1.700	
	Máy in (PMU 18)	1	1.000	
	Máy fax (cho nhóm nghiên cứu)	1	500	
	Máy tính xách tay (cho nhóm nghiên cứu)	1	1.800	
12.	Phiên dịch (bảng câu hỏi, tài liệu hướng dẫn, tài liệu vận hành) 200 trang		8/trang	1.600
13.	In ấn, photocopy			800
14.	Chi phí bất thường			1.362
	<b>Tổng số</b>			<b>85.000</b>

## **Điều khoản tham chiếu: Giám sát viên điều tra hoặc người điều tra chính**

### **I. Mô tả công việc**

Giám sát viên điều tra trong nước hoặc người điều tra chính sẽ chịu trách nhiệm đối với công việc điều tra gốc ở trong Việt Nam. Trách nhiệm bao gồm, xác định mức độ sẵn có của thông tin ở cấp xã; giúp kiểm tra và hoàn tất các bảng câu hỏi cấp huyện, cấp xã và cấp hộ gia đình; kiểm tra bảng câu hỏi ở thực địa; kết hợp các điều chỉnh vào trong bảng câu hỏi; sắp xếp để dịch bảng câu hỏi; thuê mướn và huấn luyện những người giúp việc; lập kế hoạch hậu cần cho công việc hiện trường; chuẩn bị thực hiện điều tra và tài liệu cho bảng câu hỏi; giám sát việc thực hiện điều tra và tổ chức công việc làm sạch và nhập dữ liệu. Người này cũng đóng vai trò liên lạc giữa Đơn vị quản lý dự án của Bộ Giao thông PMU18, đại diện thường trú của Ngân hàng Thế giới, đại diện của Cơ quan Viện trợ Quốc tế Canada ở Hà Nội, và người quản lý công việc dự án của Ngân hàng Thế giới tại Washington. Người này sẽ báo cáo trực tiếp cho người quản lý công việc. Người này sẽ bắt đầu công việc vào tháng Một năm 1997; hợp đồng có thể được xử lý trong kỳ hạn chín tháng, với mức lương 1.000 đô la mỗi tháng.

### **II. Công việc cụ thể**

Các công việc cụ thể gồm có:

1. Đảm nhận trách nhiệm thuê mướn, soạn thảo điều khoản tham chiếu chi tiết, huấn luyện và giám sát sáu người trợ lý chính, những người này sẽ làm việc với các trợ lý ở địa phương (có thể từ cơ quan giao thông ở địa phương), ở thực địa, và sẽ chịu trách nhiệm thu thập các dữ liệu cấp huyện, cấp xã, và cấp hộ gia đình;

2. Khai thác sự sẵn có của dữ liệu ở cấp xã và làm việc chặt chẽ với người quản lý công việc của Ngân hàng Thế giới để thiết kế bản cuối cùng của bảng câu hỏi;
3. Tiến hành kiểm tra bảng câu hỏi trên thực địa ở các xã miền Nam và miền Bắc; báo cáo lại các khó khăn tiềm tàng và những sửa đổi cần thiết; sửa chữa lại bảng câu hỏi khi cần thiết;
4. Tổ chức dịch, in ấn và photocopy bảng câu hỏi (bản cuối cùng của bảng câu hỏi cần bằng cả tiếng Anh và tiếng Việt);
5. Chọn sáu tỉnh được điều tra sao cho mỗi tỉnh đại diện cho một khu vực địa lý- khi có nhiều hơn một tỉnh như vậy, tỉnh được chọn mẫu sẽ được lựa chọn ngẫu nhiên; chuẩn bị một mẫu ngẫu nhiên gồm khoảng 200 xã nông thôn tại sáu tỉnh này, bao gồm khoảng một nửa số xã có dự án và nửa còn lại không có dự án;
- 6.. Lập kế hoạch tất cả các công việc hậu cần của nghiên cứu thực địa, bao gồm sắp xếp phương tiện đi lại, lái xe, phụ cấp đi lại, lịch trình điều tra ở các xã, và thông báo cho chính quyền xã về các chuyến thăm và mục đích của nhóm điều tra;
7. Tham dự vào việc thực hiện điều tra, thay thế giữa các nhóm trong vai trò giám sát; đảm bảo kiểm soát chất lượng; xác định các khó khăn tác động tới việc thực hiện điều tra, kiểm tra chất lượng và sự hoàn chỉnh của dữ liệu thu thập được, đề xuất các phương án giải quyết khó khăn và thực hiện chúng sau khi đã tham vấn người phụ trách chính việc nghiên cứu;
8. Đảm bảo các kỳ điều tra trong tương lai có thể lặp lại điều tra gốc, điều này đòi hỏi (a) chuẩn bị tài liệu chi tiết về tất cả thiết kế và hậu cần cho việc thực hiện điều tra (việc chọn mẫu các tỉnh, các xã, và hộ gia đình được thực hiện như thế nào; nhóm điều tra được huấn luyện và tổ chức như thế nào; công việc thực địa được tổ chức như thế nào; thủ tục nào được tiếp nối khi một khu vực được chọn mẫu không thể tiếp cận được, hay khi không tìm thấy hộ gia đình được

chọn mẫu; các khó khăn, vấn đề phát sinh và giải pháp); và (b) chuẩn bị một tài liệu hướng dẫn chi tiết về định nghĩa các thuật ngữ (ví dụ, thất nghiệp, thu nhập, nghề nghiệp chính, trẻ em hay người lớn, khoảng cách), đơn vị, số lượng tiền tệ, các mã số được sử dụng trong bảng câu hỏi; bảng câu hỏi được tiến hành như thế nào và đối tượng của bảng câu hỏi; giá cả được thu thập như thế nào, và v.v. Công việc trước là để đảm bảo cho các đợt điều tra trong tương lai có thể lại cho ta những thông tin chi tiết về tổ chức và hậu cần như đợt điều tra gốc, và công việc sau, sẽ cần thiết trong huấn luyện những người điều tra và cho công việc của những người này, cũng như hỗ trợ cho những người sử dụng dữ liệu này trong tương lai (sẽ có cả bản tiếng Anh và tiếng Việt);

9. Mua sắm trang thiết bị cần thiết như được xác định trong ngân sách nghiên cứu;
10. Thiết lập mối quan hệ tốt và bảo đảm sự hợp tác chặt chẽ với PMU18. Cung cấp thông tin cập nhật cho đơn vị này về nghiên cứu và theo dõi bước phát triển của dự án; giám sát việc thành lập một cơ sở dữ liệu cụ thể của dự án (người quản lý công việc của Ngân hàng Thế giới sẽ xác định các dữ liệu nào cần đưa vào trong cơ sở dữ liệu này);
11. Tổ chức và giám sát công việc làm sạch và nhập dữ liệu bằng cách sử dụng chương trình nhập dữ liệu được cung cấp; và
12. Đóng vai trò liên lạc và truyền tin một cách thường xuyên với người quản lý công việc.

### Phụ lục 3

## Một mẫu ngân sách từ một đánh giá tác động của một chương trình cho trẻ ăn tại trường

Giai đoạn I: 7-1999 - 12-2000\*

Dự án nghiên cứu cho trẻ ăn tại trường- đánh giá nghiên cứu cơ sở và nghiên cứu chéo  
(7-1999- 12-2000)

Phác thảo ngân sách – 7-14-1999 - đơn vị: đô la Mỹ

	Tuần làm việc của nhân viên hoạt động		Nguồn tài chính/chi phí		
	Năm tài khoa 2000	Năm tài khoa 2001	BB	RPO	Nguồn khác
					Tổng cộng
<b>Cán bộ Ngân hàng</b>					
<b>Thế giới</b>					
Nhà kinh tế	4	2	17.640		
Chuyên gia đánh giá	5	3	23.520		
Chuyên gia dinh dưỡng	5	3	23.520		
Đồng nghiệp kiểm tra lại	0,2	0,2	1.948		
Đồng nghiệp kiểm tra lại	0,2	0,2	1.948		
					68.577
<b>Cán bộ FES</b>					
Điều phối viên của nghiên cứu	4	4		12.000	12.000
<b>Nhà tư vấn quốc tế</b>					
Đánh giá tình huống (kể cả chi phí đi lại)				7.000	
Xây dựng kiểm tra nhận thức (kể cả chi phí đi lại)			6.000		

(Xem tiếp trang sau)

**ĐÁNH GIÁ TÁC ĐỘNG CỦA CÁC DỰ ÁN PHÁT TRIỂN...**

<i>Tuần làm việc của nhân viên hoạt động</i>	<i>Nguồn tài chính/chi phí</i>				
	<i>Năm tài khoa 2000</i>	<i>Năm tài khoa 2001</i>	<i>BB</i>	<i>RPO</i>	<i>Nguồn khác</i>
Chuyên gia chọn mẫu				2.000	
Nghiên cứu chi phí - hiệu quả				25.000	40.000
<b>Hãng tư vấn trong khu vực<sup>b</sup></b>					
Thiết kế, chọn mẫu, thực hiện diều tra				42.000	
Công việc tại thực địa				25.000	
Xử lý dữ liệu				3.500	
Phân tích				30.000	
					100.500
Đi lại tới quốc gia				12.000	
Các chuyến đi	4	2			12.000
Chi phí bất thường				1.000	
Liên lạc				2.000	
Phần mềm				2.000	
Phiên dịch					5.000
<b>TỔNG CỘNG</b>			68.577	150.500	19.000 238.077

Tổng số yêu cầu từ RAD: 150.500 đô la

Tổng số yêu cầu từ ngân sách của Ngân hàng Thế giới: 68.577  
đô la

Tổng số cung cấp bởi các nguồn bên ngoài: 19.000 đô la

a. Ước lượng ngân sách cho giai đoạn II của đánh giá không được tính  
trong dự án này.

b. Sự phân chia các chi phí này được nêu ở trang sau.

**Dự trù Ngân sách - Thu thập dữ liệu ở địa phương và phân tích dữ liệu cho Giai đoạn 1**

Đánh giá tác động của chương trình cho trẻ ăn tại trường  
Chi phí tính theo đô la Mỹ

	Số người #	Số tuần làm việc của cán bộ #	Chi phí theo tuần	Tổng công
<b>Các nhà chuyên môn</b>				
Giám đốc	1	12	2.000	24.000
Chuyên gia giáo dục	1	8	1.500	12.000
Chuyên gia dinh dưỡng	1	8	1.500	12.000
Nhà thống kê/chọn mẫu	1	12	750	9.000
Nhà quản lý thực địa	1	8	750	6.000
Người lập trình	1	10	300	3.000
Giám sát viên xử lý dữ liệu	1	8	300	2.400
Trợ lý điều tra	1	10	100	1.000
Trợ lý nhân trắc	1	10	100	1.000
Trợ lý kiểm tra nhận thức	1	10	100	1.000
Kiểm soát chất lượng dữ liệu	1	8	100	800
Tổng bộ phận: cán bộ chuyên môn				72.200
<b>Cán bộ thực địa</b>				
Giám sát viên	4	6	200	4.800
Người kiểm tra nhận thức	4	6	120	2.880
Người do nhân trắc	4	6	120	2.880
Người phỏng vấn	4	6	120	2.880
Lái xe	4	5	100	2.000
Trang thiết bị cho công việc tại thực địa		Số người/dơn vị	Chi phí theo tuần hay theo đơn vị	
Ô tô (4 xe cho 5 tuần)	4	5	350	7.000
Xăng (4 xe cho 5 tuần)	4	5	80	1.600
Cân; thước đo (5 bộ)	5		20	100
Thiết bị kiểm tra nhận thức (cho 4 người kiểm tra)	4		20	80
Thiết bị điều tra (cho 4 người phỏng vấn)	4		20	80
Tổng bộ phận - công tác thực địa				24.300
<b>Xử lý dữ liệu</b>				
Mã hóa dữ liệu	3	7	75	1.575
Nhập dữ liệu	4	7	75	2.100
Tổng công bộ phận xử lý dữ liệu				3.675
<b>Tổng cộng</b>				100.175

Phụ lục 4

**Các chỉ số tác động -  
Đánh giá Quỹ đầu tư xã hội  
Bolivia**

Được xây dựng vào tháng 11/1997

**I. Giáo dục chính thức- các trường loại "A" và "B" (nhiều  
cấp lớp và thông thường)**

**1. Các chỉ số tác động cuối cùng**

Kết quả kiểm tra môn Toán và Tiếng Anh

Tỷ lệ lưu ban

Tỷ lệ bỏ học

Tỷ lệ nhập học

Trình độ giảng dạy

Nhu cầu về giáo dục (phần trăm các học sinh bị đuổi  
học).

**2. Các chỉ số tác động trung gian**

Mức độ đi học thường xuyên của học sinh

Mức độ đi làm thường xuyên của giáo viên

Phân bổ thời gian của học sinh/số giờ học tập

Phương pháp giảng dạy trong lớp học<sup>a</sup>

Mức độ luân chuyển giáo viên

**3. Các chỉ số của chương trình**

Cơ sở hạ tầng

Tỷ lệ học sinh/lớp học

Số lớp học có điều kiện vật chất tốt

Số lớp học bị thiếu

Sự sẵn có của khu vực đa chức năng

Sự sẵn có của các dịch vụ cơ bản

- Điện
- Nguồn cung cấp nước chủ yếu
- Loại hình dịch vụ vệ sinh; điều kiện vệ sinh

#### Trang thiết bị nội thất

- Tỷ lệ học sinh/ bàn
- Tỷ lệ bàn của giáo viên/lớp học
- Tỷ lệ ghế của giáo viên/lớp học
- Tỷ lệ bảng đen/lớp học
- Tỷ lệ kệ đựng/lớp học
- Sách giáo khoa và giáo cụ
- Tỷ lệ sách giáo khoa/học sinh
- Chất lượng của các sách giáo khoa môn toán, ngôn ngữ, khoa học xã hội và khoa học tự nhiên
- Sự sẵn có của sách cho giáo viên
- Sự sẵn có và tình trạng các bản đồ và biểu đồ
- Các trò chơi sự phạm theo chu kỳ của trường (tiền cơ bản, cơ bản và trung gian)
- Sự sẵn có của bàn tính
- Chỉ số cải cách giáo dục<sup>b</sup>

#### 4. Các nhân tố tác động tới các kết quả không liên quan tới dự án SIF (nhân tố ngoại sinh)

- Dinh dưỡng
- Sự sẵn có của chương trình ăn sáng tại trường
- Chi phí của trường học
- Đặc điểm của giáo viên
- Nền tảng giáo dục
- Số năm phục vụ
- Hình thức đào tạo đã được học
- Các phương pháp áp dụng trong giảng dạy (trong một thời kỳ lớp học)
- Hình thức đào tạo đã được học, theo chủ đề và khoá học
- Thực tế đánh giá học sinh (mức độ thường xuyên cho bài tập về nhà và chừa bài)

Đánh giá giáo viên của học sinh

Lý do bỏ học

Học sinh bị nhà trường đuổi học

Khoảng cách từ nhà đến trường

Tỷ lệ học sinh/giáo viên

#### 5. Các chỉ số nhận dạng

Trường có được chương trình cải cách giáo dục ưu tiên không

Chi phí chương trình tính theo thành phần của dự án

Chi tiêu thực tế tính theo thành phần của dự án

## II. Sức khoẻ

### 1. Các chỉ số tác động cuối cùng

Tỷ lệ tử vong trẻ sơ sinh

Tỷ lệ tử vong trẻ em

Tỷ lệ nhiễm các bệnh dịch chính và mức độ phổ biến của chúng

Tình trạng phổ biến của suy dinh dưỡng (suy dinh dưỡng chung, nhẹ, trung bình, và nghiêm trọng)

### 2. Các chỉ số tác động trung hạn

Việc sử dụng các trung tâm y tế của chính phủ

Sự phổ biến của tiêm vắc-xin phòng uốn ván

Địa điểm tiêm vắc-xin

Sự phổ biến của kiểm soát sinh đẻ

Số biện pháp kiểm soát sinh đẻ

Chất lượng của biện pháp kiểm soát

Sự phổ biến của việc sinh con tại các trung tâm y tế

Chất lượng của ca sinh

Sự phổ biến của sinh con tại nhà với sự hỗ trợ của cán bộ y tế

Chiều cao khi sinh

Cân nặng khi sinh

**Đánh giá nhân trắc học****Địa điểm thực hiện đánh giá****Tuổi khi tiến hành đánh giá****Tỷ lệ nhiễm bệnh và sự phổ biến của việc tiêm chủng miễn dịch****Bệnh đậu mùa****Bệnh bạch hầu- uốn ván- pertussis (DPT)****Bệnh sởi****Bệnh lao (TB)****Kiến thức về địa điểm cần đến để tiêm chủng****Tỷ lệ nhiễm và điều trị bệnh ho****Tỷ lệ nhiễm và điều trị bệnh tiêu chảy****Sự phổ biến về kiến thức và cách sử dụng gói muối uống rehydration****Kiến thức của trạm y tế về sự phổ biến của tình trạng có thai****Sự chăm sóc các sản phụ có rủi ro cao****Sự phổ biến của thói quen vệ sinh tốt và sử dụng nước****Thời kỳ cho con bú sữa mẹ****3. Các chỉ số của chương trình<sup>d</sup>****Chất lượng cơ sở hạ tầng theo loại hình trung tâm y tế****Sự sẵn có của các dịch vụ cơ bản trong trung tâm y tế (nước uống, hệ thống nước thải, và điện)****Sự đầy đủ các cơ sở hạ tầng dựa trên các tiêu chuẩn đã được xác lập theo loại hình trung tâm y tế****Sự đầy đủ của trang thiết bị dựa trên các tiêu chuẩn đã được xác lập theo loại hình trung tâm y tế****Số giường bệnh trong trung tâm y tế****Sự sẵn có của các loại thuốc thiết yếu theo loại hình của trung tâm y tế****Sự sẵn có của các công cụ y tế thiết yếu theo loại hình của trung tâm y tế****Sự sẵn có của các trang bị nội thất thiết yếu theo loại hình của trung tâm y tế**

**4. Các nhân tố tác động đến kết quả không có liên hệ với dự án SIF (nhân tố ngoại sinh)**

Đặc điểm của hộ gia đình

Chất lượng hộ gia đình

Loại hình hộ gia đình

Các dịch vụ cơ bản trong hộ gia đình

Điện

Nguồn nước

Loại hình dịch vụ vệ sinh

Khả năng tiếp cận dễ dàng tới các dịch vụ cơ bản

Khoảng cách giữa hộ gia đình và trung tâm y tế gần nhất

Khoảng cách giữa dịch vụ vệ sinh và nguồn nước

Khoảng cách giữa hộ gia đình và nguồn nước chính

Số giờ có sẵn nước trong ngày

Mức độ đầy đủ của nước trong ngày

Chi phí của việc khám bệnh tại trung tâm y tế

Nhận thức của chủ hộ về chất lượng của :

“Dịch vụ” trong trung tâm y tế mà hộ gia đình từng đến

“Cơ sở hạ tầng” của trung tâm y tế mà hộ gia đình từng đến

“Mức độ sẵn có thuốc men” trong trung tâm y tế mà hộ gia đình từng đến

Chi tiêu của hộ gia đình

Các đặc điểm cá nhân của các thành viên trong hộ gia đình

Tuổi tác

Ngôn ngữ

Trình độ giáo dục

Nghề nghiệp

Các đặc điểm về địa lý

- Tình hình y tế trong huyện
- Tình hình y tế trong vùng
- Khu vực y tế
- Tỉnh
- Địa phương
- Nguồn nhân lực trong trung tâm y tế (bác sĩ, chuyên gia dinh dưỡng, y tá, trợ lý y tá, kỹ thuật viên, cán bộ hành chính)
- Số dân nằm trong khu vực tác động của trung tâm y tế theo nhóm tuổi
- Chi phí khám bệnh ở trung tâm y tế
- Các chương trình y tế không được SIF tài trợ

#### **5. Các chỉ số nhận dạng**

- Chi phí chương trình theo thành phần dự án
- Chi tiêu thực tế theo thành phần dự án

### **III. Nước**

#### **1. Các chỉ số tác động cuối cùng**

- Tỷ lệ tử vong trẻ sơ sinh
- Tỷ lệ tử vong trẻ em
- Tỷ lệ nhiễm bệnh tiêu chảy trong các hộ gia đình
- Sự phổ biến của tình trạng suy dinh dưỡng (chung, nhẹ, trung bình, và nghiêm trọng)

#### **2. Các chỉ số tác động trung gian**

- Tỷ lệ nhiễm và điều trị bệnh tiêu chảy tại các trung tâm y tế
- Sự phổ biến của việc sử dụng và hiểu biết về sử dụng viên muối rehydration
- Sự phổ biến của thói quen vệ sinh và sử dụng nước tốt

**3. Các chỉ số của chương trình (của đầu vào)**

Sự phổ biến của đào tạo về các chủ đề sức khoẻ  
Khả năng tiếp cận các dịch vụ cơ bản

Nguồn nước chính

Sự tồn tại và loại hình dịch vụ vệ sinh

Khoảng cách giữa dịch vụ vệ sinh và nguồn nước

Khoảng cách giữa hộ gia đình và nguồn nước  
chính

Số giờ có sẵn nước trong ngày

Mức độ đầy đủ về nước trong ngày

Sự sẵn có về nước trong năm

Lượng nước tiêu thụ của hộ gia đình<sup>a</sup>

Chất lượng nước<sup>a</sup>

**4. Các nhân tố tác động tới các kết quả không liên quan  
tới dự án SIF (nhân tố ngoại sinh)**

Việc sử dụng các trung tâm y tế của chính phủ (MSSP)

Kích cỡ trẻ khi sinh

Cân nặng trẻ khi sinh

Thời kỳ bú sữa mẹ

Đặc điểm của hộ gia đình

Chất lượng của hộ gia đình

Loại hình hộ gia đình

Khả năng tiếp cận các dịch vụ cơ bản

Khoảng cách giữa hộ gia đình và trung tâm y  
tế gần nhất

Chi phí của việc khám bệnh tại trung tâm y tế

Chi tiêu của hộ gia đình

Các đặc điểm cá nhân của các thành viên trong hộ gia  
đình

Tuổi tác

Ngôn ngữ

Trình độ giáo dục

Nghề nghiệp

### 5. Các chỉ số nhận dạng

Chi phí chương trình theo thành phần dự án  
Chi tiêu thực tế theo thành phần dự án

- a. Không được xem xét trong điều tra gốc
- b. Sẽ được xây dựng trên cơ sở phối hợp với cán bộ của chương trình Cải cách giáo dục; sẽ được coi là nhân tố ngoại sinh đối với chương trình trừ khi được xem xét kết hợp cùng với chương trình Cải cách giáo dục SIF.
- c. Loại trừ tỷ lệ tử vong chung, tỷ lệ sinh, tỷ lệ sinh tự nhiên toàn cầu, tỷ lệ tử vong người lớn, và tuổi thọ bình quân tính tại thời điểm sinh.
- d. Loại trừ đào tạo về các chủ đề sức khoẻ
- e. Loại trừ tỷ lệ tử vong chung, tỷ lệ sinh, tỷ lệ sinh tự nhiên toàn cầu, tỷ lệ tử vong người lớn (nam giới và phụ nữ), tuổi thọ bình quân tính tại thời điểm sinh, sự phổ biến của các bệnh hô hấp nghiêm trọng và điều trị bệnh ho.

## Phụ lục 5

# Mẫu về khung logarit cho tóm tắt thiết kế dự án phục vụ cho tài liệu hoàn thành dự án hay tài liệu đánh giá dự án (CAS)

Thú bắc mục tiêu

Các chỉ số kết quả

Theo dõi và đánh giá

Giả định quan trọng  
(Từ mục tiêu tới sứ  
mệnh của Ngân hàng  
Thế giới)

**Mục tiêu chiến lược hỗ  
trợ Quốc gia theo lĩnh  
vực:**

Đưa ra phát biểu một  
câu về mục tiêu chiến  
lực dài hạn (như phản  
 ánh trong CAS) cho dự  
 án. Phát biểu này cần  
 phải mô tả được sự thay  
 đổi đáng kể trong (các)  
 lĩnh vực được quan tâm

Chỉ số của lĩnh vực:  
1. Các chỉ số kèm theo  
mục tiêu CAS cho lĩnh  
vực có liên quan bao  
gồm các do lượng  
thường không phải do  
dự án tài trợ mà do  
Ngân hàng Thế giới tài  
trợ như là một phần  
trong các công việc  
khác.

2. Thông thường, người  
đi vay sẽ theo dõi các  
chỉ số này như một

quốc gia:  
1. Cột này xác định sẽ  
tìm thấy thông tin để  
kiểm chứng các chỉ số ở  
dầu và quá trình liên  
quan

2. Các chỉ số đi kèm với  
mục tiêu CAS theo lĩnh  
vực thường được theo  
dõi và đánh giá thông  
qua các báo cáo lĩnh vực  
hay quốc gia được viết ra  
ở bên ngoài dự án

Báo cáo lĩnh vực và

Giả định là mục tiêu  
CAS liên quan tới lĩnh  
vực (nêu ra ở hộp bên  
trái ngoài cùng) sẽ đặt  
dược trong dài hạn, liệt  
kê bất kỳ giả định bổ  
sung nào cần thiết để  
nối kết mục tiêu này  
với sứ mệnh của Ngân  
hàng Thế giới (tức là  
giảm đói nghèo).

Các giả định này  
thường liên quan đến

(Xem tiếp trang sau)

## Phụ lục 5 (Tiếp)

phần trong công việc quản lý lĩnh vực tốt

### Mục tiêu phát triển dự án:

1. Dưa ra phát biểu một câu về thay đổi hành vi dự kiến từ một nhóm hay thế chế thu hưởng mục tiêu vào cuối thời kỳ thực hiện dự án. Việc đặt được mục tiêu này đóng vai trò như là sự kiểm định đơn giản đối với nhu cầu đầu ra của dự án. Mục tiêu cần phải thể hiện một mục đích phát triển duy nhất có tính thực tế, cụ thể, có thể đo lường và được dẫn dắt bởi nhu cầu. Tham khảo hướng dẫn xác định mục tiêu dự án
2. Các chỉ số kết quả (cấp độ PDO) không phải là một phát biểu lặp lại những gì ở cấp độ đầu ra.
3. Việc thu thập dữ liệu để do luồng các chỉ số

### Các chỉ số kết quả/lúc động:

1. Các chỉ số kết quả liên quan tới kết quả sẽ đạt được vào cuối thời kỳ thực hiện dự án, trong khi có thể không hoàn toàn đạt được tác động của dự án cho tới khi đã hoàn thành giai đoạn thực hiện dự án được 5 năm hoặc lâu hơn nữa.
2. Các chỉ số kết quả (cấp độ PDO) không phải là một phát biểu lặp lại những gì ở cấp độ đầu ra.

3. Khi việc thu thập dữ liệu là bắt buộc, cần phải lưu ý tới các phương pháp và trách

các điều kiện, hành động hay phản ứng ở bên ngoài dự án và bên ngoài lĩnh vực

(Từ mục tiêu phát triển dự án tới mục tiêu CAS liên quan tới lĩnh vực)

Giả định là đạt được mục tiêu phát triển dự án, liệt kê bất kỳ giả định bổ sung nào cần thiết để minh chứng sự đóng góp của dự án cho mục tiêu đã xác định. Các giả định này thường liên quan đến (các) đóng góp của những dự án bổ sung, đều vào bổ sung hay phản ứng bổ sung từ các nhóm, và thể chế thu hưởng có vai trò thiết yếu tới việc đạt mục tiêu đã xác định

(Xem tiếp trang sau)

## Phụ lục 5 (Tiếp)

trong "Những điều nên làm và không nên làm khi xác định mục tiêu của một dự án" (gọi là "một dự án")

này thường được tài trợ bởi dự án

nhiệm, mà có thể bao gồm những yêu cầu từ phía những người thụ hưởng

bản sao)

Dầu ra từ mỗi thành phần:

1. Nếu rõ ở đây (dưới hình thức thời gian) giá trị tăng bổ sung của việc hoàn thành từng thành phần.

2. Phát biểu chính xác về giá trị dầu ra bổ sung sẽ dễ do lường được (như được phản ánh trong các chỉ số bên phải).

3. Để lập luận đơn giản và rõ ràng, cần phải có

Các chỉ số dầu ra:

1. Các chỉ số dầu ra có các đặc điểm về mặt số lượng, chất lượng, và thời gian. Nếu thời gian không được nêu ra, thì giả định là tới khi kết thúc dự án.

2. Các chỉ số dầu ra nhìn chung, bao gồm các thước đo hiệu quả về mặt chi phí.

3. Việc thu thập dữ liệu để do lường các chỉ số

Báo cáo dự án:

1. Các chỉ số dầu ra nhìn chung được theo dõi và đánh giá thông qua các báo cáo dự án khác nhau, báo cáo công việc giám sát và báo cáo đánh giá (giữa kỳ và cuối cùng).

2. Nguồn dữ liệu để theo dõi và đánh giá các chỉ số này thường bao gồm các hệ thống ghi chép hành chính và

(Từ các đầu ra đến mục tiêu phát triển dự án).

- Giả định là đạt được các đầu ra được liệt kê ở hộp phía ngoài bên trái vào cuối thời kỳ dự án, nếu ra bất kỳ giả định bổ sung nào cần thiết để đạt được mục tiêu dự án.
- Những giả định này có thể bao gồm các điều kiện, thay đổi chính sách, hay hành vi dự kiến của các nhóm hoặc thể chế thụ hưởng cần thiết cho thành công của

(Xem tiếp trang sau)

## Phụ lục 5 (Tiếp)

phát biểu đầu ra đối với từng thành phần dự án tương ứng.

4. Mỗi đầu ra cần phải tương ứng về giá trị số với một thành phần dự án.

5. Nhóm dự án chịu trách nhiệm chung trong việc đảm bảo cung ứng các đầu ra như một phân trong thiết kế dự án tốt và thực hiện, lên kế hoạch, cung cấp dự án tốt.

đầu ra được tài trợ của dự án.

4. Đối với các dự án phức tạp, một bảng riêng biệt (có thể là bảng bổ sung cho phụ lục 1) có thể được sử dụng để cung cấp danh sách các chỉ số đầu ra chi tiết hơn.

5. Có một số ít các chỉ số đầu ra có ý nghĩa và dễ đo lường, tốt hơn là có vô số chỉ số nhưng lại khó thu thập dữ liệu để đo lường.

6. Các chỉ số đầu ra được thoả thuận với người đi vay trong giai đoạn PCD (về mức độ sẵn có của dữ liệu và sự thu thập dữ liệu dễ dàng), và điều tra gốc cần được tiến hành trước khi đánh giá.

quản lý và các báo cáo tóm tắt của dự án.

5. Nhóm

dự án

- Những giả định này là quan trọng đối với việc đặt được mục tiêu đã xác định của dự án, nhưng nằm ngoài phạm vi kiểm soát trực tiếp của dự án.

dự án.

- Những giả định này là quan trọng đối với việc đặt được mục tiêu đã xác định của dự án, nhưng nằm ngoài phạm vi kiểm soát trực tiếp của dự án.

## Phụ lục 5 (Tiếp)

### Đầu vào dự án (ngân sách cho mỗi thành phần):

**Phân phong của dự án:**  
1. Một thành phần là một nhóm các thành phần phụ hay các hoạt động được thiết kế nhằm tạo ra một đầu ra dự án duy nhất.  
2. Liệt kê mỗi thành phần dự án như là các phân dự án tiêu đề chính, tiếp theo là thành phần phụ dự án, nếu có, được tài trợ như là một phần trong thành phần dự án.

### Đầu vào dự án (ngân sách cho mỗi thành phần):

1. Liệt kê các đầu vào thành phần dưới hình thức tổng chi phí cho mỗi thành phần bao gồm cả các khoản chi phí bất thường (ví dụ — đó là Mỹ).  
2. Đối với các dự án lớn hay phức tạp, chi phí cho thành phần Phụ cũng có thể được nêu ra (viết lui vào đầu hàng, nhằm tách bạch nó với chi phí thành phần).

### (Từ thành phần dự án đến đầu ra dự án).

- Giả định là các thành phần và hoạt động được liệt kê ở hộp bên ngoài, phía trái, được thực hiện thành công, liệt kê bất kỳ giả định cần thiết nào để đạt được đầu ra này.
- Những giả định này là các điều kiện ở bên ngoài phạm vi kiểm soát trực tiếp của dự án và là bắt buộc để đạt được các đầu ra đã xác định của dự án.
- Dự án không nên chỉ tiền để đạt được bất kỳ điều kiện nào trong số này (do các giả định này đã được đưa vào trong các thành phần của dự án).

### Báo cáo dự án:

1. Các đầu vào thường được theo dõi thông qua báo cáo tiền độ và báo cáo giải ngân (cả hai đều định kỳ hàng quý).
2. Các đầu vào thường được đánh giá thông qua báo cáo công việc giám sát (định kỳ nửa năm) và báo cáo kiểm toán (hàng năm).

*Nguồn: Phòng dịch vụ cơ bản hoạt động, Ngân hàng Thế giới. Để có các ví dụ hoàn chỉnh về phụ lục này, xin truy cập Bàn giao đỡ M&E trên trang Web nội bộ của Ngân hàng Thế giới tại địa chỉ <http://Lnts012/helpdesk.nsf>.*

## Phụ lục 6

### Ma trận phân tích

#### Danh giá tác động của Quỹ đầu tư xã hội Khẩn cấp của Nicaragua (SF) năm 1998

##### A. Định hướng theo đối tượng

Các виды	Các chỉ số chung	Các phương pháp	Nhân xét	Nguồn dữ liệu
Mức độ đói nghèo của xã/huyện SF	<ul style="list-style-type: none"> <li>Tỷ lệ phần trăm các hộ gia đình trong cộng đồng ở dưới ngưỡng nghèo hay mức tiêu dùng của những người rất nghèo</li> </ul>	Đòi hỏi điều tra thu nhập và tiêu dùng của hộ gia đình và xác định các hoạt động của SF theo xã và theo huyện	Để so sánh giữa các nước, cần phải có những định nghĩa giống nhau về ngưỡng nghèo	Chọn mẫu nhiều hơn từ điều tra hộ gia đình toàn quốc (LSMS) tại các xã có SF- chỉ cho các dự án giáo dục, sức khoẻ, nước, và vệ sinh
	<ul style="list-style-type: none"> <li>Mức tiêu dùng trung bình của các xã tham dự so với mức tiêu dùng trong nước</li> </ul>	Đòi hỏi điều tra thu nhập và tiêu dùng của hộ gia đình và xác định các hoạt động của SF theo xã và theo huyện	Đòi hỏi điều tra thu nhập và tiêu dùng của hộ gia đình và xác định các hoạt động của SF theo xã và theo huyện	Chọn mẫu nhiều hơn từ điều tra hộ gia đình toàn quốc (LSMS) tại các xã có SF- chỉ cho các dự án giáo dục, sức khoẻ, nước, và vệ sinh
• Chỉ số bần đói đối với nghèo (được SF sử dụng)	Các bản đồ thường sử dụng các thước đo dài	Các bản đồ thường sử dụng các thước đo dài	Nhược điểm là các chỉ số và trọng số được	SF sử dụng bản đồ đối với nghèo căn cứ vào dữ (Xem tiếp trang sau)

## Phụ lục 6 (Tiếp)

<p>diện, ví dụ chỉ số dồi nghèo tổng hợp cần cứ vào kết hợp một số biến số.</p>	<p>Lựa chọn một cách tuy ý và mỗi nước có chỉ số khác nhau. Ưu diểm là phương pháp này thường chia khu vực địa lý tối hon diều tra về thu nhập và tiêu dùng- hai điều tra này có thể đi kèm để xác định mức tiêu dùng dự kiến ở các đơn vị địa lý chia nhỏ</p>	<p>liều LMS93, sử dụng chỉ số dồi nghèo tổng hợp, sẽ cấp nhất bằng LMS98 và điều tra dân số để dự đoán tiêu dùng ở cấp địa phương</p>	<p>Có thể khác nhau đáng kể tuy theo loại dự án SF</p>
<p>Mức độ dồi nghèo của những người thụ hưởng SF (cấp độ hộ gia đình)</p>	<ul style="list-style-type: none"> <li>• Tỷ lệ phần trăm những người thụ hưởng nằm dưới ngưỡng nghèo hay ở trong tình trạng rất nghèo</li> </ul>	<p>Điều tra thu nhập và tiêu dùng chọn ra những người thụ hưởng SF hoặc bằng quy mô của SF hoặc bằng cách chọn mẫu nhiều hơn trong các xã có SF</p>	<p>Điều tra thu nhập và tiêu dùng chọn ra những người thụ hưởng SF, hoặc bằng mức trung bình</p>
	<ul style="list-style-type: none"> <li>• Mức tiêu dùng trung bình của các hộ gia đình thụ hưởng dự án với mức trung bình</li> </ul>	<p>Chọn mẫu nhiều hơn từ điều tra hộ gia đình tổ quốc (LMS9) tại các xã có SF</p>	<p>Chọn mẫu nhiều hơn từ điều tra hộ gia đình tổ quốc (LMS9) tại các xã có SF (Xem tiếp trang sau)</p>

## Phụ lục 6 (Tiếp)

<p>Phân bố nguồn lực của SF</p> <p>Các đặc điểm thiết kế thể chế có tác động tới kết quả định hướng đổi tương của SF</p>	<p>towane quốc của các hộ gia đình tương tự trong mỗi dạng dự án (ví dụ, với trẻ em tại các trường tiểu học, với khả năng tiếp cận nước đường ống, ai là người sử dụng)</p> <p>Tỷ lệ phần trăm các dự án và nguồn lực SF ở nhóm phần năm thấp nhất tại các huyện</p> <p>Các đặc điểm thiết kế thể chế có tác động tới kết quả định hướng đổi tương của SF</p>	<p>quy mô của SF, hoặc bằng cách chọn mẫu nhiều hơn trong các xã có SF; cũng có thể chọn mẫu nhiều hơn ở các xã "tương tự" nhưng không có chương trình SF</p> <p>Cần có phương pháp xếp thứ tự nhât quán giữa các nước</p> <p>Sử dụng bản đồ đối nghèo</p> <ul style="list-style-type: none"> <li>● Các khuyến khích</li> <li>● Sự tiếp cận trực tiếp các nhóm thu hưởng</li> <li>● Phản đóng góp trong dự án của cơ quan yêu cầu</li> </ul>	<p>Cần xem xét lại hệ thống xếp thứ tự và xác định lại đơn vị do luồng</p> <p>Xây dựng các biến thể chế tiêu chuẩn có thể được sử dụng để giải thích các kết quả định hướng - có thể để dàng có được các biến này từ các SF.</p> <ul style="list-style-type: none"> <li>● Các văn phòng được phân quyền</li> <li>● Phản bối nguồn lực mục tiêu</li> <li>● Danh mục dự án phụ</li> </ul>
			<p>Thông tin có được từ SF</p> <p>(Xem tiếp trang sau)</p>

## Phụ lục 6 (Tiếp)

Các nhân tố khác tác động tới kết quả định hướng của dự án	<ul style="list-style-type: none"><li>Số năm hoạt động của SF</li><li>"Vốn xã hội" của công đồng</li><li>Khoảng cách tới văn phòng chính của SF</li><li>Trình độ giáo dục cao nhất của những người thụ hưởng</li><li>Có các chương trình khác của chính phủ hay các tổ chức phi chính phủ</li><li>Mức độ bất bình đẳng về thu nhập trong nước</li></ul>	Cùng căn các định nghĩa tiêu chuẩn về các biến - các biến có được từ SF, từ các điều tra hộ gia đình (xác định được những người thụ hưởng SF), và các điều tra toàn quốc	Chỉ số duy nhất đáng ngại ngờ là "vốn xã hội của công đồng"
So sánh các phương án khác nhau	<ul style="list-style-type: none"><li>Tỷ lệ phần trăm các dự án và nguồn lực của SF trong nhóm phân năm các quận thấp nhất so với các chương trình và cơ chế tương tự khác</li></ul>	So sánh kết quả định hướng đối tượng căn cứ vào vị trí địa lý hay mức độ đối nghèo của những người thụ hưởng, tuy thuộc vào thiết kế điều tra, quy mô của SF và các chương trình khác	Lập kế hoạch cho phân tích hiệu quả về chi phí
		Khó tìm được các chương trình để so sánh phù hợp; cần các thông tin riêng biệt được thu thập từ các chương trình so sánh	(Xem tiếp trang sau)

## Phụ lục 6 (Tiếp)

B. Lợi ích	Các vấn đề	Các chỉ số chung	Nguồn dữ liệu/ phương pháp	Nhận xét	Nghiên cứu tình huống Nicaragua
Vốn vật chất	<ul style="list-style-type: none"> <li>Mức độ các dự án phụ đáp ứng các ưu tiên của các xã</li> <li>Nhận thức của người thụ hưởng về mức độ lợi ích và cải thiện phúc lợi</li> </ul>	<p>Điều tra cấp xã, đánh giá đổi tượng thụ hưởng hoặc điều tra hộ gia đình trong đó lấy mẫu nhiều hơn tại các khu vực có SF</p> <p>Điều tra hộ gia đình hoặc đánh giá người thụ hưởng tại các xã có SF</p> <p>Cải thiện khả năng tiếp cận cơ sở hạ tầng kinh tế xã hội (trước và sau chương trình)</p>	<p>Điều tra cấp xã, đánh giá đổi tượng thụ hưởng hoặc điều tra hộ gia đình trong đó lấy mẫu nhiều hơn tại các khu vực có SF</p> <p>Điều tra hộ gia đình hoặc đánh giá người thụ hưởng tại các xã có SF</p> <p>Cải thiện khả năng tiếp cận cơ sở hạ tầng kinh tế xã hội (trước và sau chương trình)</p>	<p>Được thực hiện trong Điều tra Danh giá đối tượng thụ hưởng và cơ sở vật chất do IDB tài trợ</p> <p>Được thực hiện trong điều tra hộ gia đình và Đánh giá đổi tượng thụ hưởng do IDB tài trợ</p> <p>Cần phải có điều tra gốc hay các câu hỏi gởi nhỡ, cần xây dựng các chủ số riêng biệt cho từng kiểu dự án SF</p>	<p>Được thực hiện trong điều tra hộ gia đình và Đánh giá đổi tượng thụ hưởng do IDB tài trợ</p> <p>Một số câu hỏi gởi nhỡ trong điều tra hộ gia đình và một số câu hỏi trong LSMS trước đó, có thể so sánh người thụ hưởng SF với những đặc điểm của dân số toàn quốc và với các xã tương tự</p>
					(Xem tiếp trang sau)

## Phụ lục 6 (Tiếp)

	<ul style="list-style-type: none"> <li>Cải thiện khả năng tiếp cận cơ sở hạ tầng kinh tế xã hội so với các dự án so sánh</li> </ul>	Điều tra hộ gia đình về người thụ hưởng SF và người thụ hưởng SF và các dự án so sánh	Cần phải có điều tra gốc hay các câu hỏi gợi nhớ; cần xây dựng các chỉ số riêng biệt cho từng kiểu dự án SF	Có thể so sánh người thụ hưởng SF với những đặc điểm của dân số toàn quốc và với các xã tương tự
	<ul style="list-style-type: none"> <li>Cải thiện chất lượng cơ sở hạ tầng và dịch vụ (trước và sau)</li> </ul>	Điều tra trang bị và điều tra hộ gia đình, bao gồm một phần trong đánh giá người thụ hưởng	Cần phải có điều tra gốc hay các câu hỏi gợi nhớ; cần xây dựng các chỉ số riêng biệt cho từng kiểu dự án SF	Đối với lĩnh vực giáo dục, y tế, nước và vệ sinh, có các câu hỏi ghi nhớ và thông tin quá khứ từ điều tra gốc vật chất và đánh giá trước đối với SF
Vốn con người	<ul style="list-style-type: none"> <li>Cải thiện chất lượng cơ sở hạ tầng và dịch vụ so với các dự án so sánh</li> </ul>	Điều tra cơ sở vật chất và điều tra hộ gia đình, bao gồm một phần đánh giá người thụ hưởng (trong SF và các dự án được so sánh)	Cần phải có điều tra gốc hay các câu hỏi gợi nhớ; cần xây dựng các chỉ số riêng biệt cho từng kiểu dự án SF	Đối với lĩnh vực giáo dục, y tế, nước và vệ sinh, thông tin về cơ sở vật chất của SF và các dự án không phải SF có được từ điều tra hộ gia đình và điều tra gốc vật chất
	<ul style="list-style-type: none"> <li>Cải thiện tình hình giáo dục: tỷ lệ nhập học, số năm hoàn</li> </ul>	Điều tra hộ gia đình và thông tin từ nhà trường	Điều tra hộ gia đình và thông tin từ nhà trường đối với các	(Xem tiếp trang sau)

## Phụ lục 6 (Tiếp)

thành, tỷ lệ bỏ học  
và lưu ban (trước và  
sau khi có dự án và  
so với dự án so  
sánh)

- Cải thiện tình hình sức khỏe: ví dụ, tỷ lệ nhiễm bệnh, tỷ lệ tử vong trẻ sơ sinh, tình trạng suy dinh dưỡng, tăng cường việc nuôi con bằng sữa mẹ (trước và sau khi có dự án và so với dự án so  
sánh)

Điều tra hộ gia đình  
và thông tin từ các  
trung tâm y tế đối với  
các trung tâm y tế và  
các hộ gia đình tham  
gia và không tham gia  
SF

trường học và hộ gia  
đình tham gia SF và  
không tham gia

- Cải thiện tình hình kinh tế: tăng cường thu nhập, giảm thời gian đi lấy nước, giảm chi phí dịch vụ, tăng việc làm (trước và sau và so  
với dự án so sánh)

Điều tra hộ gia đình

đối với người thu  
nhưởng SF và các xã  
tương tự không thu  
nhưởng SF

Điều tra hộ gia đình

(Xem tiếp trang sau)

## Phụ lục 6 (Tiếp)

Vốn xã hội	<ul style="list-style-type: none"> <li>Tăng cường năng lực của cộng đồng trong giải quyết khó khăn (so với dự án so sánh)</li> </ul>	Điều tra hộ gia đình, điều tra cộng đồng và/hoặc đánh giá người thụ hưởng (so với dự án so sánh)	Điều tra hộ gia đình, điều tra cộng đồng và/hoặc đánh giá (so với dự án so sánh)	Cần xây dựng các chỉ tiêu	Thông tin trong điều tra hộ gia đình về sự tham dự	Không được giải quyết
C. Sự bền vững của các lợi ích						
Các vấn đề	Các chỉ số chung	Nguồn dữ liệu/ phương pháp	Nhận xét	Nghiên cứu tình huống Nicaragua		
Sự bền vững của hoạt động	<ul style="list-style-type: none"> <li>Các điều kiện tại đó các dự án SF đang hoạt động sau khi có chương trình SF (sự bền vững tuyệt đối)</li> </ul>	Điều tra cơ sở vật chất	Có thể có thêm thông tin từ đánh giá người thụ hưởng	Đối với các điều tra về dự án giáo dục và sức khoẻ, cả với các cộng đồng có SF và không có SF		
	<ul style="list-style-type: none"> <li>Các điều kiện tại đó các dự án SF hoạt động sau khi triển khai chương trình SF so với các dự án so sánh (sự bền vững tương đối)</li> </ul>	Điều tra cơ sở vật chất	Có thể có thêm thông tin từ đánh giá người thụ hưởng	Đối với các điều tra về dự án giáo dục và sức khoẻ, cả với các cộng đồng có SF và không có SF		

## Phụ lục 6 (Tiếp)

Sự bén vững của bảo quản	<ul style="list-style-type: none"> <li>Bảo quản cơ sở hạ tầng và dịch vụ theo thời gian (tuyệt đối)</li> </ul>	Điều tra cơ sở vật chất	Có thể có thêm thông tin từ đánh giá người thu hướng	Đối với các điều tra về dự án giáo dục và sức khoẻ, cả với các cộng đồng có SF và không có SF
Sự bén vững của tác động	<ul style="list-style-type: none"> <li>Chất lượng và số lượng cơ sở hạ tầng và dịch vụ theo thời gian</li> </ul>	Điều tra cơ sở vật chất và điều tra hộ gia đình	Có thể có thêm thông tin từ đánh giá người thu hướng	Đối với các điều tra về dự án giáo dục và sức khoẻ, cả với các cộng đồng có SF và không có SF
Sự bén vững trong tác động tới cộng đồng	<ul style="list-style-type: none"> <li>Xu hướng các cộng đồng SF để trình dự án của mình (tối SF và các chương trình khác) theo thời gian</li> </ul>	Cơ sở dữ liệu của SF, điều tra cộng đồng, hay đánh giá người thu hướng	Sẽ cần công việc bổ sung	

## Phụ lục 6 (Tiếp)

- Sự tham dự của công đồng trong các nhu cầu về cơ sở hạ tầng kinh tế xã hội theo thời gian

Bao gồm đánh giá người thu hưởng tiếp theo; đánh giá tác động

### D. Các vấn đề về chi phí- hiệu quả

#### Các vấn đề

#### Các chỉ số chung

#### Nghiên cứu tình huống

#### Nicaragua

Chi phí - hiệu quả	Cơ sở dữ liệu của SF và thông tin từ các bộ và chính quyền địa phương	Nhân xét	Nghiên cứu tình huống
Chi phí - hiệu quả của các dự án phụ	<ul style="list-style-type: none"> <li>Chi phí trung bình cho mỗi trường học, trạm y tế, hệ thống nước mới so với các phương án thay thế và so với các dự án so sánh</li> </ul>	<p>Chi phí thay đổi theo thời gian và các dự án để so sánh phải giống nhau</p>	<p>Dữ liệu về SF và không phải SF từ điều tra gốc vật chất.Các ước lượng về chi phí không phải SF có thể không đáng tin cậy</p>
Chi phí đơn vị chi phí trên mỗi mét vuông xây dựng, mỗi kilômét đường xá, v.v, so với các dự án so sánh	<ul style="list-style-type: none"> <li>Chi phí đơn vị chi phí trên mỗi mét vuông xây dựng, mỗi kilômét đường xá, v.v, so với các dự án so sánh</li> </ul>	<p>Cơ sở dữ liệu của SF và thông tin từ các bộ và chính quyền địa phương</p>	<p>Có thể tính được chi phí trung bình của SF. Số ở trong phân tích hiệu quả về chi phí</p>

(Xem tiếp trang sau)

## Phụ lục 6 (Tiếp)

Chi phí - hiệu quả của cơ chế giao nhận	<ul style="list-style-type: none"><li>Chi phí trung bình cho việc làm được tạo ra so với các dự án so sánh</li><li>Các chi phí thể chế của SF (đầu tư và hoạt động) tính theo tỷ lệ chi phí dự án của SF so với các dự án so sánh</li></ul>	Cơ sở dữ liệu của SF và thông tin từ các bộ và chính quyền địa phương	<p>Có thể tính được chi phí trung bình của SF. Sẽ ở trong phân tích hiệu quả về chi phí</p> <p>Cần phát triển các định nghĩa chuẩn xác về chi phí thể chế, nếu rõ thời kỳ</p>
	<ul style="list-style-type: none"><li>Thời gian hoàn thành trung bình so với các dự án so sánh</li></ul>	Cơ sở dữ liệu của SF và thông tin từ các bộ và chính quyền địa phương	<p>Có thể tính được chi phí trung bình của SF. Sẽ ở trong phân tích hiệu quả về chi phí</p>

**NHÀ XUẤT BẢN VĂN HÓA - THÔNG TIN**  
**43 LÒ ĐÚC - HÀ NỘI**  
**Điện thoại: 04 8.214756 - 04 8257 477**

---

Chịu trách nhiệm xuất bản  
**VŨ AN CHƯƠNG**

Biên tập: **VIỆT HÀ**  
**PHƯƠNG MAI**  
Trình bày: **PHAN BÍCH**  
Vẽ bìa: **PHAN BÍCH**

---

In 2.000 cuốn, khổ 15,5 x 23,5 cm.  
Tai Công ty in và Văn hoá phẩm - Bộ Văn hoá.  
Giấy phép xuất bản số: 589/XB-QLXB/03-VHTT,  
cấp ngày 30 tháng 5 năm 2002  
In xong và nộp lưu chiểu tháng 6 năm 2002.