

Xác định các chỉ báo đo lường nghèo đa chiều cho hộ gia đình nông thôn Việt Nam

Trần Tiên Khai, Nguyễn Ngọc Danh

Khoa Kinh Tế, Đại Học Kinh Tế TP. Hồ Chí Minh

Tóm lược

Khái niệm nghèo đa chiều hiện nay đang được áp dụng rộng rãi trên thế giới. Mặc dù vậy, cơ sở khoa học cho chọn lựa các chỉ báo phù hợp để đo lường nghèo đa chiều vẫn còn chưa rõ ràng. Nghiên cứu này nhằm đến việc xác định các chỉ báo đo lường nghèo đa chiều cho hộ gia đình ở nông thôn Việt Nam dựa trên tiếp cận sinh kế bền vững và áp dụng các phương pháp phân tích đa biến như Principle Component Analysis, Multiple Correspondence Analysis để khám phá các vấn đề nêu trên. Bộ dữ liệu phân tích bao gồm 6.750 hộ nông thôn trích từ dữ liệu điều tra Mức sống Hộ gia đình Việt Nam năm 2010 (VHLSS 2010). Kết quả nghiên cứu xác nhận rằng có 9 chiều đo lường cho tình trạng nghèo đa chiều và đại diện cho bốn nhóm tài sản sinh kế. Một số biến có thể được sử dụng như là các chỉ báo phù hợp cho đo lường nghèo đa chiều.

Từ khóa: nghèo đa chiều, tài sản sinh kế, Principle Component Analysis, Multiple Correspondence Analysis, Cluster Analysis

1. Giới thiệu

Xác định bản chất của nghèo và cách thức đo lường nghèo là vấn đề thuộc lĩnh vực kinh tế phát triển được quan tâm ở tầm mức thế giới vì tính phức tạp của chúng. Việc xác định và đo lường nghèo một cách phù hợp có thể giúp chúng ta có được nhận thức tốt hơn về xã hội và chính phủ có những đáp ứng hiệu quả hơn để xóa đói giảm nghèo. Cách tiếp cận phổ biến nhất về đo lường nghèo dựa trên quan điểm so sánh thu nhập hay tiêu dùng của cá nhân hay hộ gia đình với một ngưỡng mà xã hội coi như là một mức chuẩn về mức sống tại một thời điểm. Cách tiếp cận này dẫn đến hai phương thức phân loại nghèo điển hình là nghèo tuyệt đối và nghèo tương đối.

Hiện nay đã có một số nghiên cứu áp dụng tiếp cận nghèo đa chiều ở Việt Nam. Tuy nhiên, các nghiên cứu này hoàn toàn không chỉ ra cơ sở khoa học cho chọn lựa các chiều đo lường và các chỉ báo. Từ đó đặt ra yêu cầu cần phải chọn lựa các chiều đo và chỉ báo hợp lý cả về phương diện lý thuyết lẫn thống kê và thực tiễn. Chính vì vậy, bài viết nhằm tìm kiếm các chỉ báo kinh tế - xã hội phù hợp để đo lường nghèo đa chiều cho hộ gia đình nông thôn Việt Nam.

Mục tiêu cụ thể của bài viết này là: 1) xác định các chỉ báo kinh tế - xã hội tiềm năng cho tình trạng nghèo đa chiều; 2) tìm hiểu được các quan hệ giữa các chỉ báo này và tình trạng nghèo về tiền; và 3) áp dụng các công cụ thống kê đa biến để xác định các chỉ báo phù hợp đo lường nghèo đa chiều.

Nội dung bài viết bao gồm các phần 1) Giới thiệu; 2) Cơ sở lý thuyết và Phương pháp nghiên cứu, 3) Kết quả nghiên cứu và Thảo luận; và 4) Kết luận và gợi ý chính sách.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1 Cơ sở lý thuyết và khung phân tích

Ngoài khái niệm nghèo về tiền, nghèo có thể được giải thích bởi các chỉ báo đa chiều (Anand & Sen, 1977). Nghèo không chỉ được đo lường bằng thu nhập, chi tiêu mà còn bởi khả năng tiếp cận một cách đồng thời đến lương thực, nhà ở, giáo dục, chăm sóc sức khỏe và các mức sống xã hội khác, ngay cả các chỉ báo phi vật chất. Tổng hòa các chỉ báo này phản ánh chất lượng cuộc sống. Hiện nay, các tổ chức quốc tế đã áp dụng khái niệm nghèo đa chiều và xây dựng các chỉ số đo lường nghèo đa chiều. Các chỉ số đa chiều phổ biến nhất là Chỉ số nghèo con người (Human Poverty Index - HPI) do Anand và Sen đề xuất (1997), Chỉ số phát triển con người (Human Development Index - HDI) được Liên Hiệp Quốc sử dụng, và Chỉ số nghèo đa chiều

(Multidimensional Poverty Index - MPI) do Đại học Oxford và UNDP áp dụng dựa trên phương pháp luận của Alkire và Foster (2007).

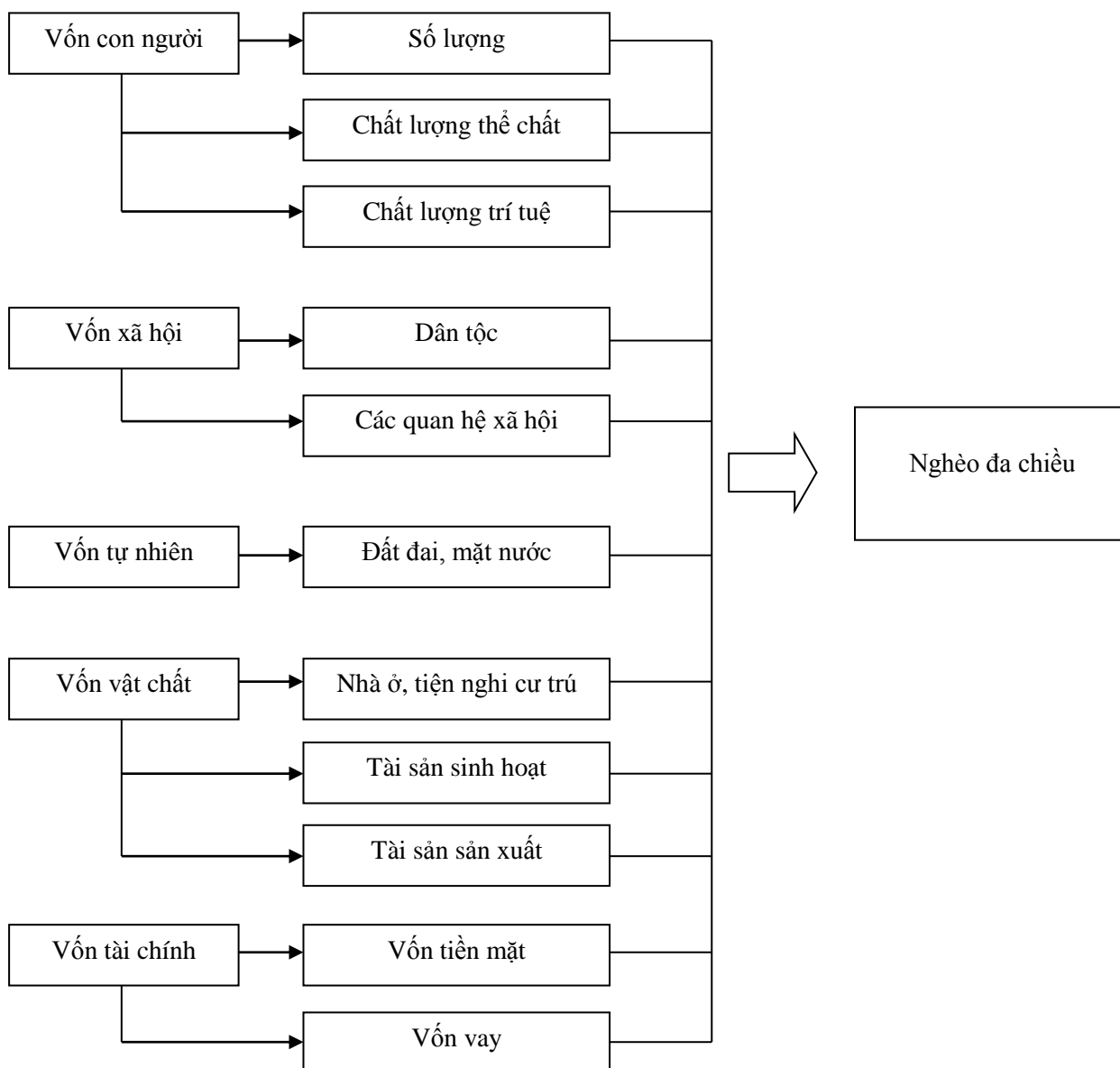
Cho đến nay hầu hết các nghiên cứu về nghèo ở Việt Nam vẫn sử dụng tiếp cận nghèo đơn chiều mặc dù Ngân hàng thế giới (2003) đã chỉ ra rằng Việt Nam đã áp dụng sáu phương pháp đo lường nghèo khác nhau, trong đó có bốn phương pháp áp dụng tiếp cận nghèo đa chiều. Gần đây, nghiên cứu đánh giá nghèo đô thị ở Hà Nội và TP. Hồ Chí Minh năm 2010 (UBND TP. Hà Nội, UBND TP. Hồ Chí Minh & UNDP, 2010) áp dụng chỉ số nghèo đa chiều MPI bao gồm tám chiều đo lường và 21 chỉ báo với trọng số ngang bằng nhau. Báo cáo Nghèo của Tổng cục Thống kê năm 2010 cũng có áp dụng chỉ số nghèo đa chiều cho trẻ em bao gồm các khía cạnh giáo dục, y tế, dinh dưỡng, nhà ở, nước sạch, vệ sinh, không làm việc trước tuổi lao động, vui chơi giải trí, hòa nhập xã hội và được xã hội bảo vệ. UNDP (2011) đã công bố Báo cáo quốc gia về phát triển con người năm 2011 cho Việt Nam trong đó áp dụng so sánh ba phương pháp đo lường là nghèo tiền tệ, HPI và MPI. Chỉ số nghèo đa chiều MPI được UNDP xây dựng dựa trên ba thước đo (chiều) là y tế, giáo dục và mức sống, được đại diện bằng chín chỉ tiêu 1) hộ phải bán tài sản, vay nợ để trả phí chăm sóc y tế hoặc ngưng chữa trị; 2) thành viên hộ chưa hoàn thành bậc tiểu học; 3) trẻ em trong độ tuổi đi học không đến trường; 4) sử dụng điện thấp sáng; 5) tiếp cận nước uống sạch; 6) tiếp cận vệ sinh; 7) tiếp cận nhà vệ sinh tiêu chuẩn; 8) sống ở nhà cố định; và 9) có sở hữu tài sản lâu bền. Quan hệ giữa nghèo và các chỉ báo kinh tế - xã hội khác cũng được áp dụng ở một số nghiên cứu thực nghiệm trên thế giới (Asselin, 2009; Ki, Faye & Faye, 2009, trích bởi Asselin, 2009; Crooks, 1995). Asselin và Vu đã áp dụng 5 chiều đo lường cho Việt Nam bao gồm giáo dục, sức khỏe, nước sạch/vệ sinh, việc làm và nhà ở (Asselin, 2009). Có thể thấy, kết quả đo lường nghèo đa chiều tùy thuộc rất nhiều vào sự tin cậy của các chiều đo và các chỉ tiêu đại diện cho từng chiều đo.

Để có thể đo lường chính xác nghèo đa chiều, cần phải xác lập hệ thống các chiều đo và các chỉ tiêu phù hợp với bối cảnh kinh tế - xã hội của từng quốc gia, vùng miền và nhóm người cần đo lường về các khía cạnh lý thuyết, thống kê và thực tiễn. Hiện nay, các nghiên cứu sử dụng MPI chủ yếu chọn lựa các chiều đo và các chỉ tiêu dựa trên lý thuyết nghèo và kinh nghiệm thực tiễn là chính. Vì vậy, nghiên cứu này hướng đến việc tìm kiếm cách thức chọn lựa các chiều đo và các chỉ tiêu hợp lý về phương diện thống kê, có nghĩa là chúng phải có quan hệ thực sự về mặt thống kê với tình trạng nghèo của hộ gia đình hoặc cá nhân cần đo lường.

Tiếp cận sinh kế ngày nay đã được áp dụng rộng rãi khi nghiên cứu về đặc điểm kinh tế - xã hội của hộ gia đình ở các nước đang phát triển. Khung phân tích sinh kế bền vững (DFID, 1999) xác định năm nhóm tài sản, hay còn được hiểu là “vốn” mà sinh kế dựa vào. Các tài sản này bao gồm tài sản (hoặc là vốn) con người, tự nhiên, vật chất, tài chính và xã hội. Gia tăng khả năng tiếp cận đến các tài sản sinh kế này bằng cách sở hữu hay sử dụng được hiểu là hỗ trợ cho sinh kế và giảm nghèo. Khái niệm tài sản sinh kế cho phép hiểu nghèo đa chiều thông qua các chỉ báo về tài sản sinh kế. Như vậy, có thể tồn tại các quan hệ chặt chẽ giữa các chỉ báo nghèo về tiền và các chỉ báo về tài sản sinh kế. Mỗi tài sản sinh kế cũng có thể được coi là một chiều đo lường của nghèo đa chiều, và được biểu thị bằng nhiều chỉ báo khác nhau. Nghiên cứu này dựa vào lý thuyết sinh kế bền vững vì tính chất toàn diện của lý thuyết cho phép tạo ra nền tảng hình thành các chiều đo lường khác nhau cho nghèo đa chiều. Sự giàu có hay nghèo nàn về các tài sản sinh kế cũng đồng nghĩa với sự giàu có hay nghèo theo quan niệm đa chiều.

Dựa trên tiếp cận sinh kế bền vững, Trần Tiến Khai và Nguyễn Ngọc Danh (2012) đã thử sử dụng bộ số liệu VHLSS năm 2008 và đã xác định mười chiều đo đại diện cho bốn nhóm tài sản sinh kế là vốn con người, vốn tự nhiên, vốn vật chất, và vốn tài chính của hộ gia đình nông thôn Việt Nam dựa trên các phương pháp thống kê đa biến là phân tích thành phần chính PCA và Multiple Correspondence Analysis (MCA). Kế thừa kết quả trên, nghiên cứu này tiếp tục dựa vào khung phân tích sinh kế bền vững (DFID, 1999) để xác định các chỉ báo đo lường nghèo đa chiều cho hộ

gia đình nông thôn Việt Nam và sử dụng bộ dữ liệu VHLSS 2010 để kiểm tra lại tính nhất quán của các chiều đo và các chỉ báo nghèo đa chiều.



Hình 1. Khung phân tích nghèo đa chiều (rút ra từ Khung sinh kế bền vững, DFID, 1999)

2.2 Phương pháp nghiên cứu

Nguồn dữ liệu

Bài viết sử dụng bộ dữ liệu VHLSS 2010 với 6.750 hộ gia đình nông thôn ở sáu vùng kinh tế xã hội. Các chỉ tiêu điều tra được chia tách thành tám nhóm thông tin là: 1) cấu trúc hộ gia đình; 2) giáo dục; 3) chăm sóc sức khỏe; 4) thu nhập và việc làm; 5) tiêu dùng; 6) tài sản cố định và đồ dùng lâu bền; 7) nhà ở; và 8) tham gia xóa đói giảm nghèo và tín dụng.

Phương pháp phân tích

Nghiên cứu của Asselin (2009) đã khai thác rất sâu các khía cạnh lý thuyết và ứng dụng của các phương pháp đo lường nghèo đa chiều khác nhau để xây dựng Chỉ số nghèo tổng hợp (Composite Indicator of Poverty - CIP). Asselin đã xem xét nhiều phương pháp khác nhau như CIP dựa trên Chỉ

số bất bình đẳng (Inequality Indices), CIP dựa trên Phân tích cấu trúc nghèo (Poverty Structure Analysis), và phương pháp tiếp cận “tập mờ” (Fuzzy Subset Approach). Phương pháp Phân tích thành phần chính (Principal Component Analysis - PCA) chỉ phù hợp với các chỉ báo định lượng trong khi các biến phân loại (định tính) lại rất quan trọng ở các bộ dữ liệu điều tra mức sống. Vì vậy phương pháp phân tích Multiple Correspondence Analysis (MCA) được đề nghị sử dụng để xử lý các chỉ báo định tính và dĩ nhiên là phải mã hóa dữ liệu theo phương cách phù hợp.

Các tác giả Ki, Faye và Faye (2009, trích bởi Asselin, 2009) và Asselin và Vu (2009, trích bởi Asselin, 2009) đã áp dụng phương pháp Phân tích cấu trúc nghèo kết hợp với MCA. Trần Tiến Khai và Nguyễn Ngọc Danh (2012) đã thử nghiệm kết hợp PCA, MCA cho bộ dữ liệu VHLSS 2008 để tìm kiếm các chỉ báo nghèo đa chiều cho hộ gia đình nông thôn Việt Nam.

Thừa kế kinh nghiệm về áp dụng các công cụ thống kê đa biến trên, cũng như dựa trên khung phân tích, bài viết đã trích ra hơn 30 biến kinh tế - xã hội bao gồm cả hai dạng định lượng và định tính và chia thành năm nhóm theo tài sản sinh kế (Bảng 1).

Dữ liệu được phân tích theo các bước sau:

Bước 1. Phân tích tương quan được áp dụng để khám phá các quan hệ giữa các biến dùng làm chỉ báo cho nghèo đa chiều. Xác định tương quan giữa tình trạng nghèo đơn chiều dựa trên chi tiêu và thu nhập bình quân đầu người và các chỉ báo về tài sản sinh kế.

Bước 2. Xác định các biến phù hợp đại diện cho năm nhóm tài sản sinh kế để sử dụng như các chỉ báo tổng hợp của nghèo đa chiều bằng hai phương pháp phân tích thống kê đa biến PCA và MCA.

Phân tích PCA được áp dụng để xác định các chỉ báo đo lường nghèo đa chiều có tính chất định lượng. Phương pháp xoay áp dụng là Varimax với Kaiser Normalization. Các biến có giá trị tuyệt đối của hệ số tải nhân tố (factor loadings) nhỏ hơn 0,4 bị loại khỏi phép xoay. Có 15 biến được đưa vào phân tích cho toàn bộ nông thôn Việt Nam và 6 vùng kinh tế - xã hội.

Phân tích MCA được dùng để chọn ra các biến đại diện phân loại và thứ bậc cho các tài sản sinh kế của hộ gia đình. Giá trị của tất cả các biến đều được mã hóa theo trật tự thứ bậc để tính toán. Phương pháp chuẩn hóa số liệu được chọn là Variable Principal để tối ưu hóa sự kết hợp giữa các biến. Có 19 biến được đưa vào phân tích cho toàn bộ nông thôn Việt Nam và khu vực nông thôn ở 6 vùng kinh tế - xã hội.

Phần mềm PASW Statistics 18.0 được dùng cho các phân tích thống kê.

Bảng 1. Mô tả các biến trích ra từ bộ dữ liệu VHLSS 2010

Nhóm tài sản sinh kế	Biến	Thang đo, mô tả
Vốn con người	Quy mô nhân khẩu	Tỷ số; tổng số thành viên của hộ gia đình
	Số năm đi học trung bình của một thành viên của hộ	Tỷ số; tổng số năm đi học của tất cả các thành viên gia đình chia cho số thành viên gia đình
	Bằng cấp cao nhất	Thứ bậc; mã của loại bằng cấp cao nhất mà thành viên trong gia đình có được
	Số người đau ốm trong năm	Tỷ số; tổng số thành viên trong hộ có đau ốm trong vòng 12 tháng qua
	Số ngày đau ốm trung bình trong năm	Tỷ số; tổng số ngày đau ốm trong năm chia cho tổng số thành viên bộ đau ốm
	Số ngày khám chữa bệnh trung bình trong năm	Tỷ số; số ngày khám chữa bệnh trung bình trong năm tại các cơ sở y tế của các thành viên hộ gia đình trong vòng 12 tháng qua
	Chi phí khám chữa bệnh trung bình trong năm	Tỷ số; chi phí khám chữa bệnh trung bình trong năm tại các cơ sở y tế của các thành viên hộ gia đình trong vòng 12 tháng qua
	Số lao động	Tỷ số; số thành viên gia đình có việc làm chính từ (1) làm công làm thuê, (2) hoạt động nông nghiệp; (3) hoạt động phi nông nghiệp

Nhóm tài sản sinh kế	Biến	Thang đo, mô tả
Vốn tự nhiên	Đất nông nghiệp	Tỷ số; diện tích đất nông nghiệp của hộ gia đình
	Tưới tiêu	Danh nghĩa; đất được tưới tiêu hay không
	Đất nông nghiệp chia theo loại	Tỷ số; diện tích đất nông nghiệp sử dụng cho (1) cây hàng năm; (2) cây lâu năm; (3) lâm nghiệp; (4) diện tích mặt nước nuôi trồng thủy sản; (5) đồng cỏ; (6) vườn cây ăn quả; (7) du canh; (8) khác
Vốn vật chất	Tài sản sản xuất	Danh nghĩa; có hay không có các tài sản như đất vườn cây lâu năm; gia súc cày kéo; chuồng trại; máy kéo; thuyền máy; thuyền; bơm nước;
	Tài sản tiêu dùng	Danh nghĩa; có hay không có các tài sản như xe máy; điện thoại di động; máy thu hình màu; dàn nhạc; máy vi tính; tủ lạnh; máy điều hòa không khí;
	Diện tích nhà ở	Tỷ số; diện tích của nhà ở
	Giá trị nhà ở	Tỷ số; giá trị của nhà ở
	Nguồn nước sinh hoạt	Thứ bậc; nước máy; giếng khoan; giếng đào; suối; nước mưa; nước sông
	Nước sinh hoạt	Danh nghĩa; nước máy; không phải nước máy
	Nhà vệ sinh	Thứ bậc; nhà vệ sinh tự hoại – bán tự hoại; dội thối nước; hai ngăn; cầu cá; các loại khác; không có nhà vệ sinh
Điện thắp sáng	Thứ bậc; điện lưới; ắc quy/máy nổ; đèn dầu; khác	
Vốn tài chính	Có vay tín dụng	Danh nghĩa; có; không từ các nguồn khác nhau
	Giá trị món vay tín dụng	Tỷ số; giá trị món vay tín dụng tại thời điểm vay
Vốn xã hội	Dân tộc	Danh nghĩa; dân tộc Kinh hoặc các nhóm dân tộc khác
	Quan hệ	Danh nghĩa; có thành viên làm công chức viên chức
Thu nhập và chi tiêu	Tổng thu nhập	Tổng thu nhập trong năm của hộ gia đình
	Tổng chi tiêu	Tổng chi tiêu trong năm của hộ gia đình
	Chi tiêu bình quân đầu người	Tỷ số; tổng chi tiêu của hộ chia cho số thành viên của hộ
	Nhóm ngũ phân vị theo chi tiêu bình quân đầu người	Thứ bậc; ngũ phân vị theo chi tiêu
	Thu nhập bình quân đầu người	Tỷ số; tổng thu nhập của hộ chia cho số thành viên của hộ
	Nhóm ngũ phân vị theo chi tiêu bình quân đầu người	Thứ bậc; ngũ phân vị theo thu nhập

3. Kết quả và thảo luận

3.1 Kết quả nghiên cứu

Quan hệ giữa các chỉ báo về tài sản sinh kế và tình trạng nghèo về tiền

Tình trạng nghèo về tiền của hộ gia đình được đo lường bằng thu nhập và chi tiêu bình quân đầu người của hộ cũng như ngũ phân vị của hai biến này. Các chỉ báo về tài sản sinh kế của hộ gia đình được thể hiện bằng các biến liên tục lẫn thứ bậc. Vì vậy, hệ số tương quan Pearson được dùng để xem xét quan hệ giữa các biến liên tục, và các chỉ số thống kê khác như Kendall's Tau-b và hệ số tương quan hạng Spearman được áp dụng để đo lường mức độ quan hệ giữa các biến thứ bậc. Kết quả cho thấy hầu hết các chỉ báo về tài sản sinh kế của hộ gia đình nông thôn có tương quan chặt và hợp lý với tình trạng nghèo về tiền¹. Điều này có nghĩa là tình trạng nghèo của hộ gia đình có thể được phản ánh thông qua chất lượng và số lượng của tài sản sinh kế mà hộ gia đình có được.

¹ Chỉ có các biến diện tích đất lâm nghiệp, mặt nước nuôi trồng thủy sản, diện tích đồng cỏ và vườn cây ăn quả là không tương quan với thu nhập và chi tiêu bình quân đầu người. Kết quả tính cho toàn bộ 6.750 quan sát.

Áp dụng phân tích thành phần chính (PCA) để tìm kiếm các chỉ báo cho nghèo đa chiều

Kết quả phân tích PCA cho toàn bộ vùng nông thôn và khu vực nông thôn ở 6 vùng kinh tế - xã hội cho phép rút ra sáu nhân tố (Bảng 2 và 3).

Bảng 2. Ma trận nhân tố đã xoay - Rotated Component Matrix^a cho toàn vùng nông thôn

	Nhân tố					
	1	2	3	4	5	6
Tổng số lao động	0,923					
Quy mô nhân khẩu	0,797					
Lao động nông nghiệp	0,780					
Lao động làm công làm thuê	0,478			-0,704		
Diện tích nhà ở		0,729				
Giá trị nhà ở		0,696				
Số năm đi học trung bình		0,502				
Diện tích đất trồng cây lâu năm		0,499			-0,484	
Số người đau ốm			0,818			
Số ngày khám chữa bệnh trung bình			0,806			
Chi phí chữa trị bệnh bình quân đầu người			0,465			
Diện tích đất trồng cây hàng năm				0,637		
Lao động phi nông nghiệp					0,827	
Giá trị món vay tín dụng						0,855
Tổng diện tích đất nông nghiệp						0,526

Nguồn: tính toán từ VHLSS 2010

Bảng 3. Tổng hợp kết quả phân tích PCA cho nông thôn 6 vùng kinh tế - xã hội

	Vùng kinh tế - xã hội*					
	1	2	3	4	5	6
Tổng số lao động	(1)	(1)	(1)	(1)	(2)	(1)
Quy mô nhân khẩu	(1)	(1)	(1)	(1)	(2)	(1)
Số năm đi học trung bình	(5)	(4)	(3)	(2)	(7)	(3)
Lao động nông nghiệp	(1)	(1)	(1)	(1)	(3)	(2)
Lao động phi nông nghiệp	(4)	(4)	(6)	(6)	(5)	(3)
Lao động làm công làm thuê	(4)	(5)	(1)	(1)	(2)	(1)
Số người đau ốm	(2)	(3)	(4)	(4)	(4)	(4)
Số ngày khám chữa bệnh trung bình	(2)	(3)	(4)	(4)	(4)	(4)
Chi phí chữa trị bệnh bình quân đầu người	(2)	(3)	(4)	(4)	(4)	(4)
Diện tích nhà ở	(3)	(4)	(3)	(2)	(1)	(3)
Giá trị nhà ở	(3)	(4)	(3)	(2)	(1)	(3)
Tổng diện tích đất nông nghiệp	x	(2)	(2)	(3)	(1)	(2)
Diện tích đất trồng cây lâu năm	(6)	(6)	(2)	(3)	(1)	(7)
Diện tích đất trồng cây hàng năm	x	-	(2)	(3)	(3)	(2)
Diện tích mặt nước nuôi trồng thủy sản	(6)	(7)	(7)	(5)	(6)	(5)
Diện tích đất lâm nghiệp	x	(2)	(5)	(6)	(7)	(5)
Diện tích đất đồng cỏ	-	(6)	(5)	(6)	(7)	(6)
Giá trị món vay tín dụng	x	(7)	x	x	x	(6)

Nguồn: tính toán từ VHLSS 2010

Ghi chú:

- Vùng kinh tế - xã hội: 1: Đồng bằng sông Hồng; 2: Miền núi phía Bắc; 3: Duyên hải Trung bộ; 4: Tây Nguyên; 5: Đông Nam Bộ; 6: Đồng bằng Sông Cửu Long
- Các số trong dấu ngoặc đơn biểu thị số thứ tự của chiều đo lường mà các biến có tương quan
- x: biểu thị biến được đưa vào tính toán nhưng bị loại vì giá trị hệ số tải nhân tố nhỏ hơn 0,4
- -: biểu thị biến không được đưa vào tính toán vì có các giá trị quan sát đồng nhất

Áp dụng phân tích Multiple Correspondence Analysis (MCA) để tìm kiếm các chỉ báo cho nghèo đa chiều

Phương pháp MCA được dùng để phân tích các biến thứ bậc. Kết quả phân tích MCA cho phép chọn được bốn chiều giải thích được 100% tổng phương sai. Độ tin cậy của các biến thành phần của các chiều này được xác nhận với các giá trị rất cao của các hệ số Cronbach's Alpha (từ 0,395 đến 0,878). Kết quả tính các hệ số discrimination measures được trình bày ở Bảng 4 (cho toàn bộ quan sát) và Bảng 5 (cho 6 vùng kinh tế - xã hội).

Bảng 4. Hệ số Discrimination Measures của các biến định tính ở các chiều đo

	Chiều			
	1	2	3	4
Bảng cấp cao nhất				0,152
Có thành viên là công chức viên chức				0,068
Gia súc cày kéo	0,639			
Chuồng trại gia súc	0,712			
Máy kéo		0,562		
Thuyền máy	0,639			
Máy bơm nước	0,712			
Phương tiện vận chuyển khác	0,721			
Xe máy	0,322			
Điện thoại di động	0,250			
Máy thu hình màu		0,486		
Dàn nhạc Hi Fi		0,513		
Máy vi tính		0,375		
Tủ lạnh		0,219		
Nguồn nước sinh hoạt	0,317		0,345	
Kiểu nhà vệ sinh	0,375		0,353	0,249
Nguồn điện thấp sáng	0,276		0,102	
Tín dụng				
Dân tộc	0,480		0,495	0,229
Tổng	5,938	2,907	1,744	1,598

Nguồn: tính toán từ VHLSS 2010

Ghi chú: những giá trị được thể hiện biểu thị hệ số Discrimination Measures của các biến ở các chiều đo phù hợp nhất

Bảng 5. Tổng hợp kết quả phân tích MCA cho nông thôn 6 vùng kinh tế - xã hội

	Vùng kinh tế - xã hội*					
	1	2	3	4	5	6
Bảng cấp cao nhất	(2)	(1)	(1)	(1)	(1)	(3)
Có thành viên là công chức viên chức	x	x	x	(4)	(4)	(3)
Dân tộc	(3)	(1)	(1)	(4)	(3)	x
Tín dụng	x	x	x	x	x	x
Gia súc cày kéo	(1)	(1)	(1)	(1)	(1)	(1)
Chuồng trại gia súc	(1)	(1)	(1)	(1)	(1)	(1)
Máy kéo	(2)	(2)	(2)	(2)	(2)	(2)
Thuyền máy	(1)	(1)	(1)	(1)	(1)	(1)
Máy bơm nước	(1)	(1)	(1)	(1)	(1)	(1)
Phương tiện vận chuyển khác	(1)	(1)	(1)	(1)	(1)	(1)
Xe máy	(1)	(1)	(1)	(1)	(1)	(1)
Điện thoại di động	(1)	(1)	(1)	(1)	(1)	(1)
Máy thu hình màu	(2)	(2)	(2)	(2)	(2)	(2)
Dàn nhạc Hi Fi	(2)	(2)	(2)	(2)	(2)	(2)
Máy vi tính	(2)	(2)	(2)	x	(2)	(2)
Tủ lạnh	x	(2)	x	x	(2)	(2)
Nguồn nước sinh hoạt	(3)	(1)	(1)	(3)	(3)	(4)
Kiểu nhà vệ sinh	(1)	(1)	(1)	(3)	(3)	x
Nguồn điện thấp sáng	(3)	(1)	(1)	(3)	(3)	(4)

Vùng kinh tế - xã hội: 1: Đồng bằng sông Hồng; 2: Miền núi phía Bắc; 3: Duyên hải Trung bộ; 4: Tây Nguyên; 5: Đông Nam Bộ; 6: Đồng bằng Sông Cửu Long
Các số trong dấu ngoặc đơn biểu thị số thứ tự của chiều đo lường mà các biến có tương quan
x: biểu thị biến được đưa vào tính toán nhưng bị loại vì giá trị hệ số discrimination nhỏ hơn 0,1

3.2 Thảo luận

Kết quả phân tích nhân tố PCA

Nhân tố thứ nhất bao gồm bốn biến định lượng là quy mô nhân khẩu, tổng số lao động, lao động nông nghiệp và lao động làm công làm thuê. Ba biến đầu tiên chỉ có tương quan với nhân tố thứ nhất. Với cấu trúc các biến có liên quan, nhân tố này ám chỉ đến nguồn nhân lực của hộ gia đình nông thôn về mặt số lượng và liên quan đến hoạt động nông nghiệp. Vì vậy, nhân tố này có thể được đặt tên là “nguồn nhân lực cho nông nghiệp”.

Nhân tố thứ hai chứa đựng hai biến là diện tích nhà ở và giá trị nhà. Nhân tố này chỉ ra tài sản vật chất của hộ gia đình dựa trên tài sản quan trọng nhất chính là nhà ở. Vì vậy, nhân tố này có thể được đặt tên là “điều kiện nhà ở”.

Nhân tố thứ ba bao gồm tất cả ba biến chỉ thị cho tình trạng sức khỏe của các thành viên hộ gia đình. Vì vậy, nhân tố này có thể đại diện cho một chiều đo lường của vốn con người về khía cạnh “tình trạng sức khỏe” của của hộ gia đình.

Biến diện tích đất trồng cây hàng năm đại diện cho nhân tố thứ tư. Biến này rõ ràng là phù hợp với hoạt động sinh kế chủ yếu của hộ gia đình nông thôn. Nhân tố này có thể được đặt tên là “nguồn lực đất đai”.

Nhân tố thứ năm bao gồm biến lao động phi nông nghiệp và biến diện tích đất trồng cây lâu năm nhưng với dấu âm. Có thể thấy biến này đại diện cho tính đa dạng hóa của lao động gia đình nông thôn theo hướng ngành nghề phi nông nghiệp, và tương quan nghịch với diện tích đất trồng cây lâu năm. Vì vậy, nhân tố thứ năm được đặt tên là “khả năng đa dạng hóa việc làm” của hộ gia đình nông thôn.

Nhân tố thứ sáu chỉ chứa đựng biến đại diện là giá trị món vay tín dụng của hộ gia đình trong năm. Có thể đặt tên nhân tố này là “thu nhập phụ” và coi như là một nhân tố đại diện cho vốn tài chính.

Kết quả phân tích riêng rẽ cho khu vực nông thôn thuộc 6 vùng kinh tế - xã hội cho thấy một số nhân tố có tính ổn định và nhất quán cao giữa các vùng kinh tế - xã hội cũng như cho toàn bộ khu vực nông thôn, thể hiện qua sự hội tụ của các biến thành phần của chúng.

Tuy nhiên, các biến đại diện cho nhân tố “nguồn lực đất đai” khác biệt khá cao giữa các vùng kinh tế - xã hội.

Rõ ràng là kết quả phân tích thành phần chính sử dụng dữ liệu VHLSS 2010 phản ánh khá rõ tính đặc thù của các vùng miền kinh tế - xã hội khác nhau ở nông thôn Việt Nam. Qua đó, có thể thấy tính tin cậy của bộ dữ liệu và tính phù hợp của phương pháp phân tích thành phần chính.

Kết quả phân tích MCA cho thấy có 4 nhóm chiều đo. Chiều thứ nhất phản ánh cả tài sản sản xuất như gia súc cày kéo; chuồng trại chăn nuôi; xuống máy, máy bơm nước và tài sản tiêu dùng như xe máy, phương tiện vận chuyên khác và điện thoại di động.

Chiều đo thứ hai đại diện cho tình trạng sở hữu các tài sản tiêu dùng sang trọng của hộ gia đình nông thôn. Máy thu hình màu, dàn nhạc Hi Fi, máy vi tính, tủ lạnh là những tài sản hiếm khi được trang bị ở nông thôn, ngoại trừ những hộ gia đình giàu có. Ngoài ra, việc sở hữu máy vi tính còn có liên quan với các hộ gia đình có trình độ học vấn cao.

Chiều đo thứ ba bao gồm hầu hết các biến chỉ thị cho điều kiện nhà ở, nước sạch và điều kiện vệ sinh. Nguồn điện thấp sáng, nguồn nước sinh hoạt và kiểu nhà vệ sinh là các chỉ báo đại diện tin

cây. Mặc dù biến nguồn điện thấp sáng có tương quan mạnh hơn với chiều đo thứ nhất, nhưng xét về tính hợp lý, biến này được chọn vào chiều đo thứ ba.

Chiều đo thứ tư chỉ ra vốn con người (mức bằng cấp cao nhất) và cả vốn xã hội (công chức viên chức và dân tộc). Tuy nhiên, chiều đo này chưa đủ mạnh để thể hiện sự tập trung mang tính nhất quán với quan hệ chặt chẽ giữa các biến và chiều đo này. Vì sự không chắc chắn này, chiều đo lường thứ tư không được xác định.

Kết quả phân tích MCA cho nông thôn 6 vùng kinh tế - xã hội cũng thể hiện tính ổn định và nhất quán. Nhóm biến đại diện cho chiều đo về tài sản sản xuất, tài sản tiêu dùng thông thường (chiều đo 1) và tài sản tiêu dùng sang trọng (chiều đo 2) khá ổn định ở các vùng, ngoại trừ sự thay đổi ở hai biến sở hữu máy vi tính và tủ lạnh. Chiều đo về tiện nghi cư trú cũng có tính nhất quán cao, và có đại diện là ba biến nguồn nước sinh hoạt, kiểu nhà vệ sinh và nguồn điện thấp sáng. Tuy nhiên, đối với Miền núi và trung du phía Bắc và Bắc trung bộ và Duyên hải miền Trung thì ba biến này không đại diện cho một chiều riêng biệt mà cùng chung chiều đo với tài sản sản xuất và tài sản tiêu dùng thông thường.

Các biến chỉ thị vốn xã hội không có quan hệ rõ ràng, chặt chẽ và hợp lý với các chiều đo ở hầu hết các vùng kinh tế - xã hội, nhưng lại có tính đặc trưng rất cao ở vùng Tây Nguyên. Ở vùng này, hai biến có thành viên gia đình là công chức viên chức và dân tộc đi cùng với nhau và đại diện cho một chiều đo độc lập về vốn xã hội. Rõ ràng là ở một vùng đặc trưng đồng bào dân tộc ít người chiếm tỷ trọng khá cao ở khu vực nông thôn, và hạn chế về việc làm từ khu vực công là những nguyên nhân hợp lý giải thích cho sự xuất hiện của chiều đo về vốn xã hội này.

Chọn lựa các chỉ báo phù hợp cho nghèo đa chiều của hộ gia đình nông thôn Việt Nam theo tiếp cận tài sản sinh kế

Phối hợp kết quả thống kê của hai phương pháp phân tích PCA và MCA ta có thể xác định 22 chỉ báo cho tài sản sinh kế của hộ gia đình nông thôn (Bảng 6). Các chỉ báo này có thể được phân loại thành nhóm chỉ báo định lượng (gồm 10 chỉ báo) và nhóm chỉ báo định tính (gồm 12 chỉ báo). Các chỉ báo này đại diện cho chín chiều đo lường của bốn nhóm tài sản sinh kế.

Trong đó, vốn con người có ba chiều đo lường độc lập là nguồn nhân lực cho nông nghiệp, tình trạng sức khỏe, và khả năng đa dạng hóa việc làm. Vốn vật chất có năm chiều đo lường độc lập là điều kiện nhà ở, tiện nghi cư trú, tài sản sản xuất, tài sản tiêu dùng thông thường và tài sản tiêu dùng sang trọng. Vốn tự nhiên có một chiều đo lường là nguồn lực đất đai. Vốn tài chính không có biên đại diện chắc chắn nào. Vốn xã hội có một biến đại diện nhưng chỉ phù hợp cho vùng Tây Nguyên.

Đối với chiều đo lường về nguồn lực đất đai, các biến đại diện rất đa dạng và có tính đặc trưng cho các vùng kinh tế - xã hội khác nhau.

Kết quả phân tích này khá tương đồng với kết quả phân tích sử dụng bộ dữ liệu VHLSS 2008, vì thế khẳng định tính nhất quán của các kết quả.

Bảng 6. Các chỉ báo khả dụng cho đo lường nghèo đa chiều theo tiếp cận tài sản sinh kế áp dụng cho hộ gia đình nông thôn Việt Nam và ở các vùng kinh tế - xã hội

Tài sản sinh kế	Nhân tố	Các chỉ báo phù hợp
(1) Vốn con người	(1) Nguồn nhân lực cho nông nghiệp	Tổng số lao động Quy mô nhân khẩu Lao động nông nghiệp
	(2) Tình trạng sức khỏe	Số người đau ốm Số ngày khám chữa bệnh trung bình Chi phí chữa trị bệnh bình quân đầu người
	(3) Khả năng đa dạng hóa việc làm	Lao động phi nông nghiệp
(2) Vốn tự nhiên	(4) Nguồn lực đất đai	Tổng diện tích đất nông nghiệp Diện tích đất trồng cây lâu năm và mặt nước nuôi trồng thủy sản ¹ Tổng diện tích đất nông nghiệp và diện tích đất lâm nghiệp ² Diện tích đất trồng cây lâu năm và đất đồng cỏ ³ Diện tích đất trồng cây hàng năm và đất trồng cây lâu năm ⁴
(3) Vốn vật chất	(5) Điều kiện nhà ở	Giá trị nhà ở Diện tích nhà ở
	(6) Tiện nghi cư trú	Kiểu nhà vệ sinh Nguồn nước sinh hoạt Nguồn điện thấp sáng
	(7) Tài sản tiêu dùng sang trọng	Máy thu hình màu Dàn nhạc Hi Fi Máy vi tính* Tủ lạnh*
	(8) Tài sản tiêu dùng thông thường	Xe máy Điện thoại di động
	(9) Tài sản sản xuất	Gia súc cày kéo Chuồng trại gia súc Thuyền máy Bơm nước Phương tiện vận chuyển khác
(4) Vốn xã hội	(10) Quan hệ xã hội ⁵	Thành phần dân tộc

Nguồn: tổng hợp từ kết quả phân tích thành phần chính và Multiple Correspondence Analysis sử dụng bộ dữ liệu VHLSS 2010

Ghi chú: *: các biến có thể được loại ra khỏi tính toán nghèo đa chiều

¹: áp dụng cho vùng Đồng bằng sông Hồng; ²: áp dụng cho vùng Miền núi và Trung du phía Bắc; ³: áp dụng cho các vùng Bắc trung bộ và Duyên hải miền Trung, Tây Nguyên và Đông Nam Bộ; ⁴: áp dụng cho vùng Đồng bằng sông Cửu Long; ⁵: áp dụng cho vùng Tây Nguyên

4. Kết luận và Khuyến nghị

4.1 Kết luận

Kết quả phân tích cho phép rút ra các kết luận sau đây:

Có thể xác định tình trạng nghèo đa chiều của hộ gia đình nông thôn dựa trên một số chỉ báo kinh tế - xã hội chỉ thị cho các tài sản sinh kế của hộ. Tình trạng nghèo đa chiều có thể được biểu thị thông qua chín chiều đo đại diện cho ba nhóm tài sản sinh kế là vốn con người, vốn tự nhiên, vốn vật chất, của hộ gia đình. Trong đó, vốn con người được đại diện bởi ba chiều độc lập là nguồn nhân lực cho nông nghiệp, tình trạng sức khỏe và khả năng đa dạng hóa việc làm. Vốn vật chất có năm chiều đo đại diện là tình trạng nhà ở, tiện nghi cư trú, tài sản sản xuất, tài sản tiêu dùng thông thường và tài sản tiêu dùng sang trọng. Diện tích đất nông nghiệp đại diện cho vốn tự nhiên. Kết quả này nhìn chung là tương đồng với kết quả phân tích với bộ dữ liệu VHLSS 2008.

Có thể sử dụng 22 chỉ báo trích ra từ bộ dữ liệu VHLSS 2010 đại diện cho mười chiều đo để mô tả tình trạng nghèo đa chiều của hộ gia đình nông thôn. Nếu áp dụng cho các vùng kinh tế - xã hội khác nhau, các chỉ báo đại diện cho chiều đo nguồn lực đất đai sẽ được áp dụng khác nhau để thể hiện tính phù hợp với đặc trưng sản xuất nông nghiệp ở từng vùng. Ngoài ra, chỉ báo thành phần dân tộc nhất thiết phải được sử dụng khi đo lường nghèo đa chiều ở những vùng có nhiều đồng bào dân tộc ít người.

Các chỉ báo đo lường nghèo đa chiều có tương quan chặt chẽ với tình trạng nghèo về tiền. Vì vậy chúng có thể cung cấp sự hiểu biết sâu sắc hơn về tình hình kinh tế - xã hội của hộ gia đình nông thôn theo tiếp cận nghèo đa chiều.

Hạn chế của nghiên cứu là bộ dữ liệu VHLSS 2010 (kể cả 2008) cung cấp rất giới hạn các chỉ báo đại diện cho vốn xã hội và vốn tài chính. Vì vậy, hai chiều đo này chưa có các chỉ báo đại diện một cách chắc chắn để có thể áp dụng vào thực tế.

4.2 Khuyến nghị

Đo lường nghèo đa chiều là một vấn đề học thuật và thực tiễn quan trọng vì có thể cung cấp cho ta sự hiểu biết sâu sắc và toàn diện về tình trạng nghèo đói, nhất là ở khu vực nông thôn nơi có rất nhiều người nghèo sinh sống. Một lần nữa, nghiên cứu này cho thấy các kỹ thuật thống kê đa biến như Principle Component Analysis và Multiple Correspondence Analysis là khả thi và phù hợp để phát hiện các chỉ báo kinh tế - xã hội đại diện cho nghèo đa chiều. Ta có thể dựa vào các chỉ báo tin cậy này để xây dựng các chỉ số đo lường nghèo đa chiều như Chỉ số nghèo tổng hợp (Composite Indicator of Poverty - CIP) hay Chỉ số nghèo đa chiều (Multidimensional Poverty Index - MPI) thay vì chỉ dựa vào kinh nghiệm hay các phân tích thống kê đơn biến giản đơn khi chọn các chỉ báo phù hợp.

Cũng cần tìm kiếm các chỉ báo đặc trưng cho từng vùng kinh tế - xã hội, và bổ sung đầy đủ hơn các chỉ báo nếu có thể, nhất là các chỉ báo đại diện cho vốn xã hội và vốn tài chính của hộ gia đình nông thôn.

Tài liệu tham khảo

- Alkire, S. and Foster, J. (2008), “Counting and Multidimensional Poverty Measurement”, OPHI Working paper No.7, Working paper series, *Oxford Poverty and Human Development Initiative*.
- Anand, S. and Sen, A. (1977), “Concept of Human Development and Poverty: a Multidimensional Perspectives”, *Human Development Papers*.
- Asselin, L. M. (2009), “Analysis of multidimensional poverty. Theory and case studies”, *Economic Studies in Inequality, Social Exclusion and Well-Being*, Vol 7, IDRC, Springer.
- Chambers, R. and Conway G.R. (1991), “Sustainable rural livelihood: practical concepts for the 21st century”, Discussion paper 296, IDS, Institute of Development Studies, Brighton.
<http://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/775/Dp296.pdf?sequence=1>, accessed 10 Jun 2012.
- Crooks, D. L. (1995), “American children at risk: Poverty and its consequences for children’s health, growth and school achievement”, *Yearbook of Physical Anthropology*, (38),57-86.
- DFID (1999), Sustainable livelihoods guidance sheets, <http://www.enonline.net/pool/files/ife/dfid-sustainable-livelihoods-guidance-sheet-section1.pdf>, accessed 18 Oct 2009.
- FAO (2005), *Impacts of Policies on Poverty – The Definition of Poverty*, EASYPoL, Module 004, FAO, Italy.

- Jahan, S. (2002), Measuring human development: evolution of the human development index, UNDP, HDR, http://hdr.undp.org/en/media/Jahan_HDI.pdf, accessed 16 Jan 2012.
- OPHI (2011), Multidimensional Poverty Index 2011, www.ophi.org.uk/wp-content/uploads/OPHI-MPI-Brief-2011.pdf, accessed 12 Dec 2011.
- OPHI (2012), Multidimensional poverty. Alkire Foster Method, <http://www.ophi.org.uk/research/multidimensional-poverty/>, accessed 15 Dec 2011.
- World Bank (2003), *Vietnam Development Report 2004: Poverty*, Report No. 27130-VN, Poverty Reduction and Economic Management Unit East Asia and Pacific Region, World Bank.
- World Bank Institute (2005), *Introduction to Poverty Analysis. Poverty Manual*, All, JH Revision of August 8, 2005.
- Oxfam và ActionAid (2012), *Theo dõi nghèo theo phương pháp cùng tham gia tại một số cộng đồng dân cư nông thôn Việt Nam. Báo cáo tổng hợp 5 năm (2007-2011)*, UKAid, DFID.
- UBND Hà Nội, UBND TP. Hồ Chí Minh, UNDP (2010), *Đánh giá nghèo đô thị ở Hà Nội và TP. Hồ Chí Minh*, Nhà Xuất Bản Thanh Niên, Hà Nội.
- UNDP (2011), *Dịch vụ xã hội phục vụ phát triển con người - Báo cáo quốc gia về phát triển con người năm 2011*, UNDP, Hà Nội.
- Viện Khoa học Xã hội Việt Nam, (2011), *Giảm nghèo ở Việt Nam: Thành tựu và Thách thức*, Nhà Xuất Bản Thế Giới, Hà Nội.