

Tác động của đa dạng sinh kế đến thu nhập hộ gia đình ở Việt Nam: Kết quả từ mô hình hồi quy với biến công cụ

ĐÀM THỊ THU TRANG*

Tóm tắt

Nghiên cứu này đánh giá tác động của đa dạng sinh kế tới thu nhập hộ gia đình ở Việt Nam trên bộ điều tra mức sống dân cư năm 2020. Thông qua phương pháp hồi quy với biến công cụ, kết quả nghiên cứu cho thấy, các chỉ số đo lường sự đa dạng sinh kế, như: Số lượng nguồn thu nhập; Chỉ số Simpson và Tỷ lệ thu nhập phi nông nghiệp đều có tác động tích cực đến Thu nhập hộ gia đình. Ngoài ra, các yếu tố, như: Tuổi chủ hộ; Tình trạng hôn nhân; Dân tộc, Khu vực sinh sống... đều có ý nghĩa thống kê và ảnh hưởng đến Thu nhập. Kết quả nghiên cứu cũng cho thấy, tăng cường đa dạng hóa sinh kế là một trong những giải pháp quan trọng nhằm tăng thu nhập và phát triển kinh tế hộ gia đình, góp phần phát triển kinh tế vĩ mô.

Từ khóa: đa dạng sinh kế, chỉ số Simpson, thu nhập phi nông nghiệp, hồi quy biến công cụ, mức sống dân cư

Summary

This research evaluates the effects of livelihood diversification on household income in Vietnam. The study uses the 2020 Vietnam Household Living Standards Survey. By instrumental variables regression, the research shows that the indicators to measure livelihood diversification including the number of income sources, the Simpson index and the non-farm income have a positive impact on household's income. In addition, some other factors such as age of household head, marital status, ethnicity, living area have statistical significance and effect on the income. Also, boosting livelihood diversification is one of the important goals to increase household income and develop household economy, contributing to macroeconomic development.

Keywords: livelihood diversification, Simpson index, non-farm income, instrumental variables regression, household living standards

GIỚI THIỆU

Đa dạng sinh kế là một trong những chiến lược quan trọng giúp giảm thiểu biến động thu nhập của các hộ gia đình nông thôn trong bối cảnh hiện nay. Vì thế, đa dạng sinh kế đã và đang là vấn đề nghiên cứu thu hút sự quan tâm của nhiều nhà khoa học trong nước và trên thế giới trong các năm qua. Các nghiên cứu dù được thực hiện ở các quốc gia khác nhau, với những đặc điểm kinh tế - xã hội có nhiều điểm khác biệt nhưng kết luận đều cho thấy tác động tích cực của đa dạng sinh kế đến cuộc sống của người dân.

Nghiên cứu này nhằm mục đích đánh giá tác động của các yếu tố đến đa dạng sinh kế và ảnh hưởng của nó đến thu nhập hộ gia đình dựa trên bộ điều tra Khảo sát mức sống dân cư Việt Nam năm 2020,

đây là bộ số liệu mới nhất tính đến thời điểm hiện tại.

MÔ HÌNH VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Mô hình hồi quy với biến công cụ

Để tiện theo dõi, chúng tôi sẽ tóm lược mô hình đối với trường hợp 1 biến độc lập. Cụ thể, xét mô hình hồi quy đơn:

$$Y = \beta_0 + \beta_1 X + u \quad (1)$$

Như chúng ta đã biết, một trong các giả thuyết của mô hình là biến giải thích X và sai số ngẫu nhiên u không có tương quan, tức là:

$$Cov(X, u) = 0$$

* ThS., Khoa Toán Kinh tế, Trường Đại học Thương mại | Email: damtrang.vcu@gmail.com

Ngày nhận bài: 03/01/2023; Ngày phản biện: 13/01/2023; Ngày duyệt đăng: 16/01/2023

Tuy nhiên, khi giả thuyết trên không thỏa mãn, tức là $Cov(X,u) \neq 0$, khi đó các ước lượng OLS bị chệch và không nhất quán. Trong trường hợp này, biến giải thích X gọi là có vấn đề nội sinh (endogeneity). Câu hỏi đặt ra là liệu chúng ta có thể tìm ra biến khác để “thay thế” hoặc “đại diện” cho biến giải thích X ? Các biến như thế được gọi là các biến công cụ (instrument variables). Thông thường, biến công cụ được ký hiệu là Z trong mô hình hồi quy. Ta có thể dùng một hoặc nhiều biến công cụ cho biến giải thích X . Biến công cụ Z cần phải thỏa mãn được cả 2 điều kiện:

(i) Biến Z không tương quan với u , tức là: $Cov(X,u) = 0$

(ii) Biến Z tương quan với X , tức là: $Cov(X,u) \neq 0$

Khi đó, từ phương trình (1), hiệp phương sai giữa z và y là:

$Cov(Z,Y) = \beta_1 Cov(Z,X) + Cov(Z,u)$
Do điều kiện của biến công cụ, ta có:

$$\beta_1 = \frac{Cov(Z,Y)}{Cov(Z,X)}$$

Dựa vào mẫu dữ liệu quan sát kích thước n , ta có được ước lượng của β_1 là:

$$\hat{\beta}_1^{IV} = \frac{\sum_{i=1}^n z_i y_i}{\sum_{i=1}^n z_i x_i}$$

trong đó, $x_i = X_i - \bar{X}$, $y_i = Y_i - \bar{Y}$, $z_i = Z_i - \bar{Z}$

Khi đó, hệ số chặn được ước lượng bởi:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1^{IV} \bar{X} \quad (2)$$

Có thể chỉ ra rằng, ước lượng là có tính nhất quán. Trong trường hợp mẫu có kích thước lớn thì:

$$\hat{\beta}_1^{IV} \sim N \left(\beta_1, \left(\frac{\sigma_u^2}{\sum_{i=1}^n x_i^2} \cdot \frac{1}{\rho_{xz}^2} \right) \right) \quad (3)$$

với ρ_{xz}^2 là bình phương của hệ số tương quan tổng giữa X và Z .

Trong thực hành, việc ước lượng mô hình (1) với biến công cụ Z được tiến hành qua 2 bước, cụ thể:

Bước 1: Hồi quy biến X qua biến Z để thu được giá trị ước lượng, theo mô hình:

$$\hat{X} = \alpha_0 + \alpha_1 Z + r \quad (4)$$

Bước 2: Hồi quy biến Y theo giá trị ước lượng thu được từ bước 1, cụ thể:

$$Y = \beta_0 + \beta_1 \hat{X} + u \quad (5)$$

Các hệ số của mô hình (4) và (5) được ước lượng qua phương pháp bình phương nhỏ nhất (OLS), nên phương pháp ước lượng ở đây được gọi là phương pháp bình phương nhỏ nhất hai bước - 2SLS (two-stage least squares).

Trong trường hợp hồi quy bội, mô hình hồi quy với biến công cụ được tiến hành tương tự và tổng quát hơn, tức là bước 1 ước lượng biến nội sinh qua các

BẢNG 1: THỐNG KÊ MÔ TẢ THEO VÙNG SINH THÁI

Biến quan sát		Toàn mẫu
Số quan sát		9.389
Thu nhập hộ gia đình (Nghìn đồng/năm)		179.445,98 (163.580,25)
Số lượng nguồn thu nhập (NIS)		4,54 (1,44)
Chỉ số Simpson (SID)		0,59 (0,06)
Tỷ lệ thu nhập phi nông nghiệp (NFS)		65,15 (35,49)
Bằng cấp chủ hộ (%)	Không bằng cấp	20,8
	Tiểu học	24,3
	Trung học cơ sở và trung học phổ thông	46,6
	Đại học và cao hơn	8,3
Tổng số thành viên hộ		3,69 (1,61)
Giới tính (%)		
Mặc định: Nam		Nữ 26,2
Tuổi chủ hộ		50,97 (14,04)
Tình trạng hôn nhân		
Mặc định: Khác		Đã kết hôn 79,8
Dân tộc (%)		
Mặc định: dân tộc Kinh		Dân tộc khác 17
Nghề nghiệp chủ hộ (%)	Kinh doanh dịch vụ	28
	Làm công ăn lương	43,5
	Làm nông thủy sản	28,5
Tỷ lệ người phụ thuộc		60,57 (62,10)
Tỷ lệ hộ gia đình không tiếp cận nước sạch (%)		10,2
Tỷ lệ hộ gia đình có nhà vệ không hợp tiêu chuẩn (%)		16,5
Diện tích đất nông nghiệp		4.486,86 (12.078,94)
Số lượng nguồn vay tín dụng chính thức		0,24 (0,49)
Nơi sống (%)	Bắc Trung Bộ và Duyên hải miền Trung	22
	Đồng bằng sông Hồng	21,2
	Đồng bằng sông Cửu Long	20,2
	Đông Nam Bộ	12
	Tây Nguyên	6,9
	Trung du và miền núi phía Bắc	17,7

Nguồn: Điều tra mức sống dân cư 2020

biến công cụ và kết quả của bước 1 được sử dụng trong bước 2.

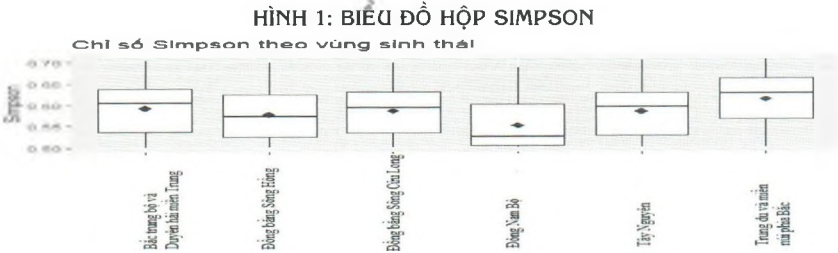
Dữ liệu điều tra mức sống dân cư

Chúng tôi sử dụng mẫu điều tra với 9.399 hộ, bao gồm: điều tra thu nhập, chi tiêu và các chủ đề khác, như: giáo dục, nông nghiệp. Nghiên cứu tác động của đa dạng sinh kế đến thu nhập của các hộ gia đình Việt Nam, sau khi xử lý số liệu mẫu nghiên cứu bao gồm có 9.389 hộ. Nghiên cứu dựa trên 14 nguồn thu được liệt kê của Tổng cục Thống kê chia ra thành: Thu nhập nông nghiệp; Thu nhập từ lương; Thu nhập từ kinh doanh dịch vụ và Thu nhập khác.

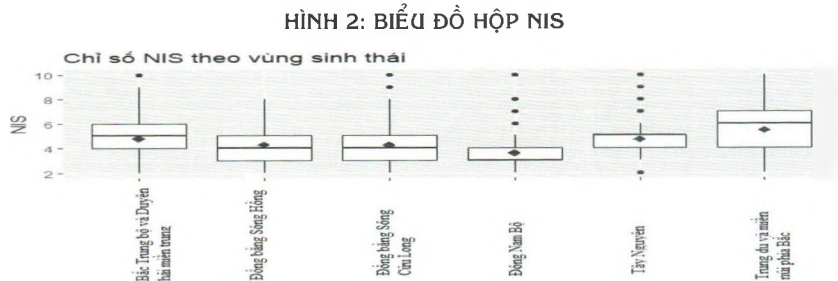
Đo lường đa dạng sinh kế

Để đo lường đa dạng sinh kế của mỗi hộ gia đình, nghiên cứu của chúng tôi dựa trên một số thước đo bao gồm:

- Chỉ số *NIS* là số lượng các nguồn thu của hộ (tính theo danh sách 14 nguồn thu được liệt kê của Tổng cục Thống kê).



Chú thích: Biểu đồ hộp (boxplot) thể hiện 4 giá trị tứ phân vị (đường nằm ngang là trung vị (mức phân vị 50%); 2 cạnh chiều rộng là mức phân vị thứ nhất (25%) và mức phân vị thứ 3 (75%). Chấm đỏ là giá trị trung bình.



Chú thích: Biểu đồ hộp (boxplot) thể hiện 4 giá trị tứ phân vị (đường nằm ngang là trung vị (mức phân vị 50%); 2 cạnh chiều rộng là mức phân vị thứ nhất (25%) và mức phân vị thứ 3 (75%). Chấm đỏ là giá trị trung bình.

Nguồn: Điều tra mức sống dân cư 2020

BẢNG 2: HỆ SỐ HỒI QUY BƯỚC 1 THEO PHƯƠNG PHÁP BÌNH PHƯƠNG NHỎ NHẤT HAI BƯỚC

	NIS	SID	Tỷ lệ thu nhập phi nông nghiệp
Hệ số chặn	3,51 ***(0,04)	0,57 ***(0)	38,72 ***(1,14)
Bằng cấp chủ hộ (Tiểu học)	0,04 (0,04)	0 (0)	5,11 ***(1,08)
Bằng cấp chủ hộ (trung học cơ sở và phổ thông)	0 (0,04)	-0,01 ***(0)	15,36 ***(0,95)
Bằng cấp chủ hộ (đại học)	-0,59 ***(0,06)	-0,03 ***(0)	31,97 ***(1,49)
Số lượng nguồn vay tín dụng chính thức	0,5 ***(0,03)	0,01 ***(0)	-4,03 ***(0,75)
Tổng số người	0,26 ***(0,01)	0 ***(0)	4,41 ***(0,23)
R^2	0,14	0,04	0,1

Chú thích: *, ** và *** lần lượt biểu diễn các mức ý nghĩa 10%, 5% và 1%.

Nguồn: Tác giả tính toán từ bộ số liệu Điều tra mức sống dân cư 2020.

- Chỉ số Simpson (Simpson’s Diversity Index - SID) dùng để đo lường sự đa dạng, được xác định bằng công thức:

$$SID = 1 - \sum_{i=1}^I p_i^2$$

trong đó, p_i là tỷ lệ thu nhập từ nguồn thu thứ i trên tổng thu nhập của hộ.

- NFS (non - farm income share) là tỷ lệ thu nhập phi nông nghiệp của hộ, tỷ lệ này được tính dựa trên tổng tỷ lệ thu nhập từ lương, tiền công và thu nhập từ ngành nghề sản xuất, kinh doanh, dịch vụ phi nông, lâm, nghiệp, thủy sản, chế biến...

Để đánh giá tác động của đa dạng sinh kế đến thu nhập của nông hộ, chúng tôi sử dụng 3 mô hình sau:

$$Y_1 = f(NIS, X_1)$$
$$Y_2 = f(SID, X_1)$$
$$Y_3 = f(NFS, X_1)$$

Trong đó, Y là biến Thu nhập của hộ gia đình. NIS , SID , NFS là các thước đo về đa dạng sinh kế. X là các biến thể hiện đặc điểm hộ gia đình, vùng miền.

Do NIS , SID , NFS là các thước đo được tính toán dựa trên thu nhập, nên trong 3 mô hình trên, NIS , SID , NFS là các biến nội sinh. Chúng tôi tham khảo các nghiên cứu trước đây, như Ho và Ha (2017) và nhận thấy, biến công cụ được sử dụng, cụ thể gồm các biến: Bằng cấp chủ hộ; Tổng số thành viên của hộ; Số lượng nguồn vay tín dụng chính thức.

Áp dụng mô hình hồi quy với 3 biến công cụ đã chọn với 3 mô hình trên, khi đó phương pháp ước lượng bình phương nhỏ nhất 2 bước được tiến hành như sau:

Bước 1: Hồi quy từng biến NIS , SID , NFS theo 3 biến: Bằng cấp chủ hộ (X_1), Tổng số thành viên của hộ (X_2), Số lượng nguồn vay tín dụng chính thức (X_3) để thu được các ước lượng tương ứng $\widehat{NIS}$, $\widehat{SID}$, $\widehat{NFS}$

$$\widehat{NIS} = \alpha_{01} + \alpha_{11}X_1 + \alpha_{21}X_2 + \alpha_{31}X_3 + r_1$$
$$\widehat{SID} = \alpha_{02} + \alpha_{12}X_1 + \alpha_{22}X_2 + \alpha_{32}X_3 + r_2$$
$$\widehat{NFS} = \alpha_{03} + \alpha_{13}X_1 + \alpha_{23}X_2 + \alpha_{33}X_3 + r_3$$

Bước 2: Hồi quy biến thu nhập (Y) theo các biến còn lại và các giá trị ước lượng được $\widehat{NIS}$, $\widehat{SID}$, $\widehat{NFS}$ từ bước 1

$$Y = \beta_{01} + \beta_{11}\widehat{NIS} + \sum_{i=4}^n \beta_{i1}X_i + u_1$$
$$Y = \beta_{02} + \beta_{12}\widehat{SID} + \sum_{i=4}^n \beta_{i2}X_i + u_2$$
$$Y = \beta_{03} + \beta_{13}\widehat{NFS} + \sum_{i=4}^n \beta_{i3}X_i + u_3.$$

KẾT QUẢ NGHIÊN CỨU

Thông kê mô tả

Bảng 1 thể hiện thông tin chung về đối tượng nghiên cứu, cho biết thu nhập bình quân hộ gia đình khoảng 179.445 nghìn đồng/năm và mô tả các đặc điểm nhân khẩu học, như: Trình độ của chủ hộ; Giới tính; Dân tộc; Tình trạng hôn nhân; Nghề nghiệp; Tỷ lệ người phụ thuộc.

Hình 1 cho thấy, miền Đông Nam Bộ có SID trung bình thấp nhất là khoảng 0,55 tức sự đa dạng thu nhập cao nhất; Trung du và Miền núi phía Bắc có SID cao nhất khoảng 0,63, có nghĩa là vùng ít đa dạng thu nhập nhất.

Hình 2 thể hiện số lượng nguồn thu nhập (NIS). NIS giữa các vùng là khác nhau, Trung du và Miền núi phía Bắc có NIS trung bình cao nhất là 5,7; Đông

Nam Bộ có NIS thấp nhất vào khoảng 3,7; các vùng còn lại có chỉ số NIS trung bình từ 4 đến 5.

Kết quả ước lượng
Bảng 2 là kết quả hồi quy bước 1, hồi quy các biến: NIS; SID; Tỷ lệ thu nhập phi nông nghiệp theo các biến công cụ: Bằng cấp chủ hộ; Số lượng nguồn vay tín dụng chính thức và Tổng số người. Theo đó, các hệ số của các biến công cụ đều có ý nghĩa thống kê, tức là có ảnh hưởng đến chỉ số đa dạng sinh kế xét theo cả 3 thước đo: NIS; SID và Tỷ lệ thu nhập phi nông nghiệp. Điều này cho thấy, sự phù hợp của các biến công cụ trong mô hình phân tích ảnh hưởng của biến đa dạng sinh kế đến thu nhập của hộ gia đình.

Bảng 3 thể hiện hồi quy bước 2 theo phương pháp bình phương nhỏ nhất. Trong cả 3 trường hợp chỉ số đa dạng sinh kế lần lượt là: NIS; SID và Tỷ lệ thu nhập phi nông nghiệp, các hệ số đều dương và có ý nghĩa thống kê. Theo đó, nếu NIS tăng thêm 1, thì thu nhập tăng thêm khoảng 39%, trong khi nếu tăng tỷ lệ thu nhập phi nông nghiệp, thì thu nhập chỉ tăng khoảng 3% với điều kiện 2 hộ gia đình có cùng đặc điểm nhân khẩu học, địa điểm và các điều kiện cơ bản như nhau. Chủ hộ đã kết hôn và chủ hộ dân tộc Kinh, thì thu nhập hộ gia đình cao hơn so với chủ hộ chưa kết hôn, dân tộc khác. Sự khác nhau về các khu vực sinh sống có ý nghĩa thống kê giữa các vùng. Các biến độc lập giải thích được 36%, 30% và 44% sự thay đổi thu nhập trong 3 trường hợp.

KẾT LUẬN

Nghiên cứu đánh giá tác động của đa dạng sinh kế đến thu nhập hộ gia đình bằng phương pháp hồi quy với biến công cụ và dựa trên bộ số liệu điều tra mức sống dân cư năm 2020. Kết quả nghiên

BẢNG 3: HỆ SỐ HỒI QUY BƯỚC 2 THEO PHƯƠNG PHÁP BÌNH PHƯƠNG NHỎ NHẤT 2 BƯỚC

	NIS	SID	Tỷ lệ thu nhập phi nông nghiệp
Hệ số chặn	10,1 ***(0,08)	10,34 ***(0,36)	9,88 ***(0,06)
Ước lượng chỉ số đa dạng sinh kế	0,39 ***(0,01)	2,48 ***(0,62)	0,03 ***(0)
Giới tính chủ hộ - Nữ (Mặc định: Nam)	0,05 *(0,02)	0,03 (0,02)	0,02 (0,02)
Tuổi chủ hộ	-0,01 ***(0)	-0,01 ***(0)	0 ***(0)
Tình trạng hôn nhân của chủ hộ - Đã kết hôn (Mặc định: Khác)	0,41 ***(0,02)	0,5 ***(0,02)	0,29 ***(0,02)
Dân tộc -Thiểu số (Mặc định: Kinh)	-0,45 ***(0,02)	-0,4 ***(0,02)	-0,3 ***(0,02)
Nghề nghiệp chủ hộ (Làm công ăn lương)	-0,06 ***(0,02)	-0,08 ***(0,02)	-0,1 ***(0,02)
Nghề nghiệp chủ hộ (Nông lâm thủy sản)	-0,23 ***(0,02)	-0,22 ***(0,02)	-0,16 ***(0,02)
Tỷ lệ người phụ thuộc	0 ***(0)	0 ***(0)	0 ** (0)
Tỷ lệ hộ gia đình không tiếp cận nguồn nước sạch	-0,26 ***(0,02)	-0,26 ***(0,02)	-0,2 ***(0,02)
Tỷ lệ hộ gia đình không có nhà vệ sinh hợp tiêu chuẩn	-0,42 ***(0,02)	-0,42 ***(0,02)	-0,29 ***(0,02)
Diện tích đất nông nghiệp	0 ***(0)	0 ***(0)	0 ***(0)
Đồng bằng sông Hồng	0,27 ***(0,02)	0,24 ***(0,02)	0,15 ***(0,02)
Đồng bằng sông Cửu Long	0,15 ***(0,02)	0,15 ***(0,02)	0,26 ***(0,02)
Đồng Nam Bộ	0,43 ***(0,02)	0,39 ***(0,03)	0,45 ***(0,02)
Tây Nguyên	-0,05. (0,03)	-0,03 (0,03)	-0,05 (0,03)
Trung du miền núi phía Bắc	0 (0,02)	-0,01(0,03)	-0,08 ***(0,02)
R ²	0,36	0,3	0,44

Chú thích: *, ** và *** lần lượt biểu diễn các mức ý nghĩa 10%, 5% và 1%.
Nguồn: Tác giả tính toán từ bộ số liệu Điều tra mức sống dân cư 2020.

cứu đã cho thấy, các chỉ số đo lường sự đa dạng sinh kế, như: NIS, SID và Tỷ lệ thu nhập phi nông nghiệp đều có tác động tích cực đến Thu nhập hộ gia đình. Ngoài ra, các yếu tố, như: Tuổi chủ hộ; Tình trạng hôn nhân; Dân tộc, Khu vực sinh sống... đều có ý nghĩa thống kê và ảnh hưởng đến Thu nhập. Điểm mới của nghiên cứu đó là đưa vào các biến công cụ có ý nghĩa như bằng cấp chủ hộ, số lượng nguồn vay chính thức và tổng số người trong hộ.

Kết quả nghiên cứu đã đóng góp thêm bằng chứng thực nghiệm để các nhà hoạch định chính sách có thể kịp thời đưa ra các chính sách khuyến khích sản xuất, phát triển thêm nhiều ngành nghề, đa dạng hóa nguồn thu nhập nâng cao đời sống của người dân. □

TÀI LIỆU THAM KHẢO

1. Benjamin, D., Brandt, L., and McCaig, B. (2017), Growth with equity: income inequality in Vietnam, 2002-2014, *The Journal of Economic Inequality*, 15(1), 25-46.
2. Ho, T.N.D., and Ha, T.V. (2017), Determinants of income diversification and its effects on rural household income in Vietnam, *Can Tho University Journal of Science*, 6, 153-162.
3. Khai T. T., and Danh N. N. (2014), Determinants of income diversification and its effect on household income in rural Viet Nam, *Journal of economic development*, 221(7), 20-41.
4. Jeffrey M. Wooldridge (2017), *Nhập môn Kinh tế lượng cách tiếp cận hiện đại, Tập 2*, Trần Thị Tuấn Anh dịch, Nxb Kinh Tế TP. Hồ Chí Minh.
5. Tổng cục Thống kê (2020), *Kết quả Khảo sát mức sống dân cư*, Nxb Thống kê.