

EFFECTS OF LIVELIHOOD DIVERSITY ON RURAL HOUSEHOLD'S INCOME IN VIETNAM: A COMPOSITIONAL DATA ANALYSIS APPROACH

Nguyen Thanh Nga^{1*}, Dam Thi Thu Trang²

Nguyen Thi Tuyet Mai², Trinh Thi Huong²

¹Banking Academy, ²Thuongmai University

ARTICLE INFO		ABSTRACT
Received:	02/8/2022	This research analyzes the current status and the impact of livelihood diversity on the Vietnamese rural households' income based on the 2020 Vietnam Household Living Standard Survey. The compositional data analysis (Coda) approach was used. Research results show that non-farm income has the highest proportion in all six socio-economic regions. Besides, the compositional data analysis regression model reveals that a 1% increase in non-farm wage income (resp., non-farm-services income) increased total household income by 0.011% (resp., 0.023%). In addition, household income was affected by household head characteristics, such as: age, marital status, ethnicity, and education level, and the number of household members, the rate of dependent members. Environmental factors, such as using clean water and standard toilets, agricultural land, and living regions, significantly impact household income. In the current context of urbanization and agricultural transformation, the increasing diversification of livelihoods is an essential goal in developing household income, contributing to the realization of macroeconomic objectives.
Revised:	16/9/2022	
Published:	16/9/2022	
KEYWORDS		
Livelihood diversity		
Income		
Rural households		
Compositional Data Analysis		
Vietnam Household Living Standard Survey		

ẢNH HƯỞNG CỦA ĐA DẠNG SINH KẾ TỚI THU NHẬP CỦA NÔNG HỘ TẠI VIỆT NAM: KẾT QUẢ TỪ PHƯƠNG PHÁP PHÂN TÍCH ĐA HỢP

Nguyễn Thanh Nga^{1*}, Đàm Thị Thu Trang²

Nguyễn Thị Tuyết Mai², Trịnh Thị Hương²

¹Học viện Ngân hàng, ²Trường Đại học Thương mại

THÔNG TIN BÀI BÁO		TÓM TẮT
Ngày nhận bài:	02/8/2022	Nghiên cứu này thực hiện phân tích thực trạng và đánh giá tác động của đa dạng sinh kế đến thu nhập hộ gia đình nông thôn Việt Nam dựa trên bộ số liệu điều tra mức sống dân cư 2020. Nghiên cứu sử dụng phương pháp phân tích số liệu đa hợp (Compositional data analysis, CoDA). Kết quả nghiên cứu cho thấy thu nhập phi nông nghiệp chiếm tỉ trọng cao nhất trong tất cả các vùng miền. Bên cạnh đó, thông qua mô hình hồi qui đa hợp khi tăng 1% thu nhập phi nông nghiệp từ tiền lương (tương ứng, thu nhập phi nông nghiệp từ dịch vụ) có tác động tăng tổng thu nhập hộ gia đình thêm 0,011% (tương ứng, 0,023%). Ngoài ra, các yếu tố ảnh hưởng đến thu nhập hộ gia đình là: tuổi chủ hộ, tình trạng hôn nhân của chủ hộ, dân tộc, tổng số thành viên hộ, tỷ lệ người phụ thuộc, bằng cấp chủ hộ, gia đình có sử dụng nước sạch, gia đình có nhà vệ sinh hợp tiêu chuẩn, diện tích đất nông nghiệp, khu vực sống của hộ gia đình. Trong bối cảnh đô thị hóa và chuyển đổi nông nghiệp như hiện nay, tăng cường đa dạng hóa sinh kế là mục tiêu quan trọng để phát triển kinh tế hộ gia đình, góp phần thực hiện các mục tiêu kinh tế vĩ mô.
Ngày hoàn thiện:	16/9/2022	
Ngày đăng:	16/9/2022	
TỪ KHÓA		
Đa dạng sinh kế		
Thu nhập		
Nông hộ		
Phân tích số liệu đa hợp		
Điều tra mức sống dân cư		

DOI: <https://doi.org/10.34238/tnu-jst.6296>

* Corresponding author. Email: ngant@hvn.edu.vn

1. Giới thiệu

Việt Nam là đất nước nông nghiệp với 2/3 dân số sống ở khu vực nông thôn; nông, lâm nghiệp và thủy sản chiếm tỉ trọng 10,94% trong cơ cấu nền kinh tế năm 2020. Nông nghiệp đóng vai trò quan trọng trong kinh tế hộ gia đình ở khu vực nông thôn, cùng với sự phát triển ổn định tiền lương của lao động làm công ăn lương, giúp phát triển kinh tế và giảm bất bình đẳng [1]. Trải qua gần 40 năm từ thời kỳ Đổi mới, nhiều chủ trương, chính sách về phát triển nông nghiệp và nông thôn dưới sự chỉ đạo của Đảng, Chính phủ và các ban ngành đã được thực hiện, nhằm mục tiêu xây dựng nông thôn mới theo hướng nông nghiệp sinh thái, nông thôn hiện đại và nông dân văn minh giúp nâng cao đời sống người dân. Trong đó, phương pháp tiếp cận sinh kế bền vững đã được áp dụng tại Việt Nam [2]. Đa dạng hóa sinh kế là một quá trình lâu dài và gồm nhiều thành phần như điều kiện kinh tế xã hội, tích lũy vốn con người và nhân khẩu học... để các gia đình nông thôn xây dựng danh mục hoạt động đa dạng trong khả năng nguồn vốn của họ và các hỗ trợ xã hội để tồn tại và cải thiện mức sống [3]. Nghiên cứu về đa dạng hóa sinh kế hay ảnh hưởng của đa dạng hóa sinh kế đến cuộc sống của người nông dân thu hút sự quan tâm của nhiều nhà nghiên cứu trên thế giới và cả Việt Nam. Một số nghiên cứu tại các nước đang phát triển như Nigeria, Ethiopia chỉ ra rằng, đa dạng sinh kế đóng vai trò tích cực đối với giảm nghèo thông qua đa dạng hoá các nguồn thu nhập, tạo việc làm, từ đó, góp phần quan trọng trong giảm nghèo bền vững (ví dụ: xem [4], [5]). Tại Việt Nam, các nghiên cứu về chủ đề này được tiến hành từ lâu, thường xuyên và cập nhật, ở cấp độ quốc gia [6] hoặc từng khu vực riêng [7]. Kết quả của đa dạng hóa sinh kế được đo lường qua nhiều chỉ tiêu, trong đó có chỉ tiêu về tăng thu nhập, số lượng các nguồn thu nhập hay tỉ lệ thu nhập phi nông nghiệp trong tổng chi tiêu [8]. Cụ thể, nghiên cứu của Benjamin và McCaig [1] sử dụng tỉ lệ đóng góp của 5 nguồn thu nhập chính là thu nhập từ sản xuất nông nghiệp, kinh doanh hộ gia đình, tiền lương, thu nhập từ nghề phụ, ngoại hối, thu nhập khác trên tổng chi tiêu. Tác giả Hồ và cộng sự sử dụng tỉ trọng đóng góp của các nguồn thu nhập phi nông nghiệp trong tổng thu nhập (non-farm income share) [8]. Các thang đo trên chưa bao hàm sự thay đổi của cơ cấu các nguồn thu nhập trong tổng thể hoặc các tỉ trọng của từng nguồn thu phi nông nghiệp. Ví dụ, một hộ nông nghiệp có tỉ trọng thu nhập phi nông nghiệp từ tiền lương và kinh doanh dịch vụ là 60%, trong đó tỉ trọng tiền lương là 40% và tỉ trọng của khoản kinh doanh dịch vụ là 20%. Tương tự, một hộ gia đình có tỉ trọng phi nông nghiệp cũng là 60% nhưng có tỉ lệ mỗi thành phần tương ứng là 30% và 30%. Do đó, hai hộ gia đình này có chiến lược đa dạng hóa sinh kế khác nhau. Nghiên cứu này sử dụng phương pháp phân tích số liệu đa hợp (Compositional data analysis, CoDA) – một phương pháp thống kê hiện đại và chuyên biệt để nghiên cứu về số liệu dạng tỷ lệ [9] – [11]. Phương pháp phân tích đa hợp được sử dụng gần đây trong các nghiên cứu về chủ đề kinh tế xã hội [12], [13]. Thông qua dữ liệu hộ gia đình sống ở nông thôn của bộ điều tra Khảo sát mức sống dân cư Việt Nam năm 2020 [14], chúng tôi đánh giá ảnh hưởng của đa dạng hóa sinh kế hộ gia đình nông thôn đến thu nhập hộ gia đình, thông qua hai mục tiêu cụ thể: 1) Thực trạng đa dạng sinh kế, thông qua tỷ trọng các nguồn thu nhập khác nhau; và 2) Tác động của cơ cấu các nguồn thu phi nông nghiệp đến thu nhập hộ gia đình.

2. Phương pháp nghiên cứu

2.1. Dữ liệu nghiên cứu

Bộ dữ liệu sử dụng để phân tích là Khảo sát mức sống dân cư Việt Nam (MSDC) 2020, được tiến hành theo Quyết định số 1261/QĐ-TCTK ngày 19 tháng 8 năm 2019 của Tổng cục trưởng Tổng cục Thống kê [14]. Đây là cuộc khảo sát định kỳ, được tiến hành 02 năm một lần trên phạm vi cả nước bằng phương pháp phỏng vấn trực tiếp chủ hộ đối với phiếu khảo sát hộ gia đình. Nghiên cứu hạn chế trên các hộ gia đình sống ở khu vực nông thôn từ mẫu gồm 9399 hộ. Đây là nghiên cứu khảo sát nhiều chỉ tiêu quan trọng của đời sống người dân để phục vụ các đánh giá mức sống của người dân Việt Nam của Chính phủ để hoạch định chính sách và lập kế hoạch phát

triển kinh tế - xã hội. Mẫu nghiên cứu đại diện cấp tỉnh, nông thôn thành thị và vùng miền. Ở phiếu khảo sát hộ gia đình, phiếu điều tra thu thập thông tin nhân khẩu học, lao động và việc làm, giáo dục, y tế... của từng thành viên hộ. Trong nghiên cứu này, thu nhập hộ gia đình được thu thập từ 12 nguồn thu nhập khác nhau của tất cả các thành viên trên 15 tuổi trong 12 tháng, trong đó chúng tôi chia ra:

- Thu nhập nông nghiệp là thu nhập từ các nguồn thu từ trồng trọt; thu từ chăn nuôi; thu từ săn bắt, thuần dưỡng chim thú; thu từ lâm nghiệp; thu từ thủy sản.

- Thu nhập từ lương là thu nhập từ lao động làm công ăn lương và phúc lợi.

- Thu nhập từ kinh doanh dịch vụ là các nguồn thu từ cho thuê đất nông, lâm nghiệp và mặt nước nuôi trồng thủy sản; thu từ dịch vụ nông nghiệp; thu từ ngành nghề sản xuất kinh doanh, dịch vụ phi nông nghiệp, lâm nghiệp, thủy sản, chế biến...; thu từ cho thuê nhà đất và nhà ở.

- Thu nhập khác là các nguồn thu từ các khoản trợ cấp, học bổng, thưởng nhận từ giáo dục; khoản thu trợ giúp nhận được từ y tế; thu khác tính vào thu nhập.

Nghiên cứu tập trung trên các hộ gia đình nông thôn, sau khi xử lý sạch số liệu, chúng tôi thu được 5804 hộ gia đình.

2.2. Phương pháp nghiên cứu: Mô hình hồi quy đa hợp

Các kiến thức nền tảng về phương pháp phân tích số liệu đa hợp đã được giới thiệu trong một số công trình đã công bố trước đó [12], [13].

2.2.1. Đơn hình và các phép toán

Trong phương pháp phân tích số liệu đa hợp (CoDA), dữ liệu được biểu diễn dưới dạng một véc tơ đa hợp gồm D thành phần. Đơn hình (simplex) S^D là không gian bao gồm các véc tơ đa hợp gồm D thành phần, cụ thể

$$S^D = \left\{ x = (x_1, x_2, \dots, x_D) : x_j > 0, j = 1, 2, \dots, D; \sum_{j=1}^D x_j = K \right\}. \quad (1)$$

Trong đó, tổng các thành phần của một véc tơ x bằng K là một hằng số.

Mỗi véc tơ $x \in S^D$ có thể được chuẩn hóa, kí hiệu $C(x)$, bằng cách đưa về véc tơ có tổng các thành phần bằng 1 (hoặc 100%) qua phép chuẩn hóa sau:

$$C(x) = C(x_1, x_2, \dots, x_D) = \left(\frac{k.x_1}{\sum_{j=1}^D x_j}, \frac{k.x_2}{\sum_{j=1}^D x_j}, \dots, \frac{k.x_D}{\sum_{j=1}^D x_j} \right). \quad (2)$$

Định nghĩa (Véc tơ đa hợp con): Xét một véc tơ đa hợp $x = (x_1, x_2, \dots, x_D) \in S^D$ và một tập chỉ số được chọn $I = \{i_1, i_2, \dots, i_s\}$. Khi đó, véc tơ đa hợp con của x ứng với tập chỉ số I , kí hiệu $x_I = (x_{i_1}, x_{i_2}, \dots, x_{i_s})$.

Định nghĩa (Véc tơ con nhất quán - Subcompositional coherence): Giả sử véc tơ đa hợp $x = (x_1, x_2, \dots, x_D) \in S^D$ và véc tơ chuẩn hóa của x là $C(x)$. Xét véc tơ đa hợp con của x là x_I và véc tơ chuẩn hóa của x_I là $C(x_I)$. Khi đó, nguyên tắc véc tơ con nhất quán đảm bảo cho các suy luận thống kê về mối quan hệ giữa các thành phần tương ứng trong véc tơ đa hợp x và trong véc tơ đa hợp con x_I sẽ cho cùng một kết quả.

Nhận thấy, do các thành phần của các véc tơ phụ thuộc lẫn nhau, nên các tính toán thông thường không phù hợp (ví dụ khoảng cách Euclide). Do đó, các phép toán được định nghĩa lại trên đơn hình S^D . Trong đó, các phép toán phải đảm bảo các nguyên tắc của độ đo Aichison (1986) [10], cụ thể là Tỷ lệ bất biến (scale invariance), Hoán vị bất biến (permutation invariance) và Véc tơ con nhất quán.

- *Phép cộng hai véc tơ (Perturbation)*

Với $x, y \in S^D$ thì tổng của x và y cũng là một véc tơ đa hợp được kí hiệu $x \oplus y$, và

$$x \oplus y = C(x_1 y_1, x_2 y_2, \dots, x_D y_D). \quad (3)$$

- *Phép nhân với một số thực (Power transformation)*

Với $x \in S^D$, $\alpha \in \mathbb{R}$ thì tích giữa x và α được kí hiệu $\alpha \odot x$, và

$$\alpha \odot x = C(x_1^\alpha, x_2^\alpha, \dots, x_D^\alpha). \quad (4)$$

Đơn hình S^D cùng với hai phép toán trên, tức $(S^D, \oplus, \odot)$ lập thành một không gian véc-tơ.

Bên cạnh đó, véc-tơ đa hợp trong đơn hình S^D thường được chuyển đổi về không gian các số thực $\mathbb{R}$, tức không gian Euclide, để thuận lợi tính toán và giải thích thông qua một số phép biến đổi dựa trên phép toán lôgarít. Trong nghiên cứu này, chúng tôi sử dụng phép biến đổi lôgarít đẳng cự (isometric log-ratio, ILR) chuyển đổi D tỉ trọng trong đơn hình về véc-tơ gồm D-1 tọa độ trong không gian Euclide thông thường [11]. ILR được thực hiện dựa trên sự phân nhóm không giao giữa các khoản mục trong véc-tơ đa hợp. Cụ thể như sau: để tạo tọa độ cân bằng đầu tiên, D tỉ trọng được chia thành 2 nhóm: một nhóm cho tử số và một nhóm cho mẫu số. Ở bước tiếp theo, một trong hai nhóm lại được tách thành 2 nhóm con để tạo ra tọa độ cân bằng thứ hai. Để tạo bảng phân tổ, tại mỗi bước, nhóm ở tử đánh số là 1, nhóm ở mẫu đánh số là -1, còn lại đánh số là 0. Ở bước thứ k, ILR_k (tọa độ cân bằng thứ k) được tạo ra từ nhóm có r_k+t_k phần tử trong đó: r_k phần tử $S_{n1} \dots S_{nr}$ ở tử số (các tỉ trọng thuộc phân tổ $\{1\}$), t_k phần tử $S_{d1} \dots S_{dt}$ ở mẫu số (các tỉ trọng thuộc phân tổ $\{-1\}$). Công thức tính ILR_k là:

$$ILR_k = \sqrt{\frac{r_k \cdot t_k}{r_k + t_k}} \ln \frac{r_k \sqrt{S_{n1} \dots S_{nr}}}{t_k \sqrt{S_{d1} \dots S_{dt}}} \text{ với } k = 1, 2, \dots, D-1. \quad (5)$$

Với phép chuyển đổi ILR ở trên, các mô hình hồi quy đa hợp được chuyển đổi về không gian Euclide thông thường và có thể được ước lượng thông qua phương pháp bình phương nhỏ nhất. Có nhiều loại mô hình hồi quy đa hợp và mô hình có thể cho phép cả véc-tơ đa hợp và tổng của các thành phần, tuy nhiên, hệ số hồi quy khi đó không được phân tích trực quan mà phải dựa trên các biến đổi và tính toán thêm. Trong nghiên cứu này, chúng tôi quan tâm mô hình có véc-tơ đa hợp đóng vai trò biến độc lập.

2.2.2. Mô hình thực nghiệm

Chúng tôi chia cơ cấu thu nhập của hộ gia đình gồm 3 thành phần: thu nhập từ nông nghiệp, thu nhập từ phi nông nghiệp và thu nhập từ các nguồn thu khác (Other, OT). Trong đó, thu nhập từ phi nông nghiệp bao gồm thu nhập từ làm công ăn lương (Non-farm-wage, NW) và thu nhập từ kinh doanh, dịch vụ (Non-farm-service, NS). Để đánh giá tác động của đa dạng sinh kế lên thu nhập của nông hộ, chúng tôi xem xét ảnh hưởng của các nguồn thu ngoài nông nghiệp lên tổng thu nhập của hộ.

Kí hiệu V_{NW}, V_{NS}, V_{OT} lần lượt là độ lớn của thu nhập hộ gia đình từ làm công ăn lương, kinh doanh dịch vụ và các nguồn thu khác. Tổng thu nhập ngoài nông nghiệp là $T = V_{NW} + V_{NS} + V_{OT}$ và $S = (S_{NW}, S_{NS}, S_{OT})$ là véc-tơ thuộc không gian các véc-tơ đa hợp S^3 :

$$S_{NW} = \frac{V_{NW}}{T}; S_{NS} = \frac{V_{NS}}{T}; S_{OT} = \frac{V_{OT}}{T}. \quad (6)$$

Phép chuyển đổi $ILR: S^3 \rightarrow \mathbb{R}^2$ được thực hiện dựa trên việc phân tổ trong đơn hình S^3 , chúng tôi tạo bảng phân tổ như trong Bảng 1.

Bảng 1. Bảng phân tổ trong đơn hình S^3

s_1	s_2	s_3	r	s
1	-1	-1	1	2
0	1	-1	1	1

Khi đó, phép biến đổi ILR tương ứng, gồm 2 tọa độ dựa trên Bảng 1 là:

$$ILR_1 = \sqrt{\frac{2}{3}} \ln \frac{s_1}{\sqrt{s_2 s_3}}; ILR_2 = \sqrt{\frac{1}{2}} \ln \frac{s_2}{s_3}. \quad (7)$$

Tọa độ ILR_1 chứa thông tin so sánh sự thay đổi giữa s_1 và $\{s_2, s_3\}$. Tọa độ ILR_2 chứa thông tin so sánh tỷ trọng s_2 và s_3 .

Để đánh giá tác động của thu nhập ngoài nông nghiệp đến thu nhập của nông hộ, chúng tôi sử dụng mô hình hồi quy đa hợp trong đơn hình S^3 có dạng:

$$\ln(\text{Income}_i) = b_0 + \bigoplus_{k=1}^3 b_k \odot S_{ki} + \sum_l \beta_l Z_{il} + \epsilon_i. \quad (8)$$

Với *Income* là thu nhập của nông hộ, *i* là chỉ số hộ gia đình, $S_i = (S_{NW_i}, S_{NS_i}, S_{OT_i})$ là vectơ cơ cấu thu nhập phi nông nghiệp, Z_i là vectơ về đặc điểm hộ gia đình, và ϵ là sai số của mô hình.

Mô hình (8) được biểu diễn qua các chuyển đổi ILR như sau:

$$\ln(\text{Income}_i) = a_0 + \sum_{j=1}^2 \alpha_j ILR_{ji} + \sum_l \gamma_l Z_{il} + \epsilon_i. \quad (9)$$

Các hệ số hồi quy của mô hình được ước lượng bằng phương pháp bình phương bé nhất. Các biến được lựa chọn vào mô hình dựa trên các nghiên cứu trước đó đã tiến hành về đánh giá tác động của đa dạng sinh kế đến thu nhập hộ gia đình [8], [15].

Việc giải thích các kết quả ước lượng từ mô hình hồi quy (2) được thực hiện tương tự như mô hình hồi quy tuyến tính bội thông thường. Cụ thể, hệ số hồi quy dương thể hiện tác động cùng chiều và hệ số hồi quy âm thể hiện tác động ngược chiều. Hơn nữa, tác động biên của 1 nhân tố đến chuyển đổi ILR được giải thích tương tự mô hình hồi quy tuyến tính thông thường và khi giữ cố định các nhân tố khác. Chẳng hạn, nếu mô hình (8) sử dụng biến đổi ILR dạng

$$ILR_1 = \sqrt{\frac{2}{3}} \ln \frac{S_{NW}}{\sqrt{S_{NS}S_{OT}}} = \sqrt{\frac{2}{3}} \ln \frac{V_{NW}}{\sqrt{V_{NS}V_{OT}}} ; \quad ILR_2 = \sqrt{\frac{1}{2}} \ln \frac{S_{NS}}{S_{OT}} = \sqrt{\frac{1}{2}} \ln \frac{V_{NS}}{V_{OT}}. \quad (10)$$

Khi đó, nếu thu nhập phi nông nghiệp từ tiền lương tăng lên $\delta\%$, ta có $V'_{NM} = V_{NM}(1 + \delta)$, trong khi các thu nhập phi nông nghiệp khác không thay đổi. Kí hiệu, ILR'_1, ILR'_2 là giá trị của ILR_1, ILR_2 sau khi thay đổi. Khi đó, theo (10) thì

$$ILR'_1 = \sqrt{\frac{2}{3}} \ln \frac{V'_{NW}}{\sqrt{V_{NS}V_{OT}}} = \sqrt{\frac{2}{3}} \ln \frac{V_{NM}(1+\delta)}{\sqrt{V_{NS}V_{OT}}} = \sqrt{\frac{2}{3}} \ln(1 + \delta) ILR_1; \quad ILR'_2 = ILR_2. \quad (11)$$

Do đó, theo kết quả ước lượng từ mô hình (9), thu nhập của hộ gia đình tăng lên $\alpha_1 \cdot \sqrt{\frac{2}{3}} \ln(1 + \delta)\%$.

Tương tự như vậy, nếu mô hình (8) sử dụng biến đổi ILR dạng

$$ILR_1 = \sqrt{\frac{2}{3}} \ln \frac{S_{NS}}{\sqrt{S_{NW}S_{OT}}} = \sqrt{\frac{2}{3}} \ln \frac{V_{NS}}{\sqrt{V_{NW}V_{OT}}} ; \quad ILR_2 = \sqrt{\frac{1}{2}} \ln \frac{S_{NW}}{S_{OT}} = \sqrt{\frac{1}{2}} \ln \frac{V_{NW}}{V_{OT}}. \quad (12)$$

Khi đó, nếu thu nhập phi nông nghiệp từ dịch vụ tăng $\delta\%$ trong khi các thu nhập phi nông nghiệp khác không thay đổi thì thu nhập hộ gia đình tăng lên bình quân khoảng $\alpha_2 \cdot \sqrt{\frac{2}{3}} \ln(1 + \delta)\%$.

3. Kết quả và bàn luận

3.1. Thông tin chung về đối tượng khảo sát

Bảng 2 thể hiện đối tượng nghiên cứu là các hộ gia đình khu vực nông thôn Việt Nam với tổng số hộ là 5804. Thu nhập bình quân khoảng 155.426 nghìn đồng/năm. Trong nghiên cứu này thu nhập hộ gia đình chia làm 3 nguồn chính: thu nhập nông nghiệp với tỷ lệ trung bình 23,13%, thu nhập phi nông nghiệp có tỷ lệ trung bình 58,59%, thu nhập khác tỷ lệ trung bình là 18,29%. Trong 58,59% đến từ thu nhập phi nông nghiệp thì có 2 nguồn thu chính là thu từ làm công ăn lương với tỷ lệ là 43,48% và thu nhập từ kinh doanh dịch vụ 15,11%.

Bảng 2. Thống kê mô tả đối tượng quan sát

Biến quan sát		Giá trị
Số quan sát		5804
Thu nhập hộ gia đình (Nghìn đồng/năm)		155426,70 (148292,55)
Tỷ lệ thu nhập từ nông nghiệp (%)		23,13 (28,93)
Tỷ lệ thu nhập phi nông nghiệp (%)		58,59 (36,27)
Tỷ lệ thu nhập khác (%)		18,29 (25,96)
Tỷ lệ thu nhập phi nông nghiệp (%)	Tỷ lệ thu nhập từ làm công ăn lương (%)	43,48 (37,31)
	Tỷ lệ thu nhập từ kinh doanh dịch vụ (%)	15,11 (28,22)

	Biến quan sát	Giá trị
Bằng cấp chủ hộ (%)	Không bằng cấp	22,50
	Tiểu học	27,30
	Trung học cơ sở và Trung học phổ thông	46,20
	Đại học và cao hơn	3,90
Tổng số thành viên hộ		3,74 (1,64)
Giới tính (%)		
Mặc định: Nam	Nữ	21,60
Tuổi chủ hộ		51,12 (14,22)
Tình trạng hôn nhân		
Mặc định: Khác	Đã kết hôn	81,20
Dân tộc (%)		
Mặc định: dân tộc Kinh	Dân tộc khác	21,10
Nghề nghiệp chủ hộ (%)	Kinh doanh dịch vụ	20,50
	Làm công ăn lương	42,30
	Làm nông thủy sản	37,20
Tỷ lệ người phụ thuộc		62,81 (63,19)
Tỉ lệ hộ gia đình không tiếp cận nước sạch (%)		13,90
Tỉ lệ hộ gia đình có nhà vệ không hợp tiêu chuẩn (%)		21,30
Diện tích đất nông nghiệp		5799,59 (12620,46)
Số lượng nguồn vay tín dụng chính thức		0,29 (0,52)
Nơi sống (%)	Bắc Trung Bộ và duyên hải miền Trung	22,30
	Đồng bằng sông Hồng	22,20
	Đồng bằng sông Cửu Long	20,90
	Đông Nam Bộ	6,80
	Tây Nguyên	7,60
	Trung du và miền núi phía Bắc	20,20

Chú thích: Biến liên tục thể hiện giá trị trung bình và độ lệch chuẩn trong ngoặc đơn, biến rời rạc thể hiện tỉ lệ phần trăm của từng mức độ.

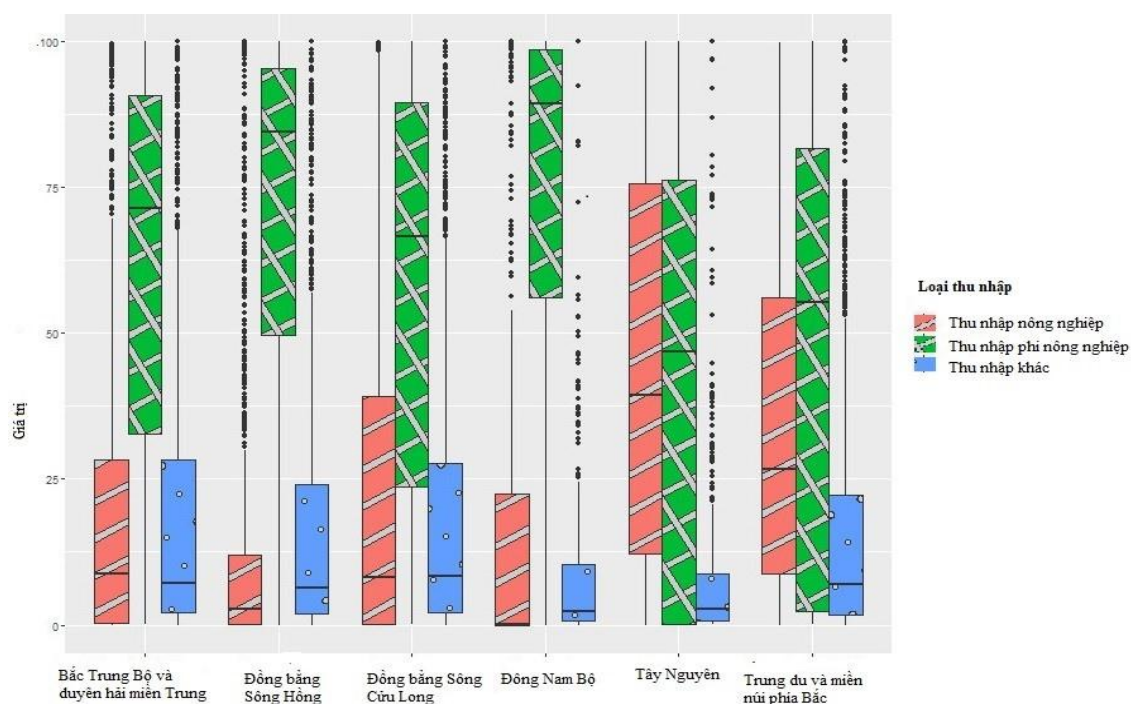
(Nguồn: Tính toán của tác giả từ Điều tra mức sống dân cư 2020)

Ngoài ra, nghiên cứu cũng quan tâm đến các yếu tố nhân khẩu học. Bằng cấp của chủ hộ được tính là bằng cấp cao nhất. Chủ hộ có trình độ Trung học cơ sở và Trung học phổ thông chiếm tỷ lệ cao nhất 46,2%, trong khi đó, 5 chủ hộ có trình độ Đại học và cao hơn có tỷ lệ thấp nhất, chỉ chiếm 3,9%, chủ hộ không có bằng cấp thì vẫn chiếm tỷ lệ lên đến 22,5%; nguyên nhân có thể do nghiên cứu tập trung vào khu vực nông thôn là nơi trình độ của người dân chưa cao và có sự chênh lệch lớn. Chủ hộ chủ yếu là nam giới, tỷ lệ nữ giới chỉ là 21,6%; tuổi trung bình là 51,12 và đa số là dân tộc Kinh, đã kết hôn. Bảng 2 cũng cho biết nghề nghiệp chủ hộ làm công ăn lương chiếm tỷ lệ cao nhất 42,3%, tiếp theo, làm nông thủy sản 37,2% và kinh doanh dịch vụ chiếm tỷ lệ thấp nhất 20,5%. Một đặc điểm nữa mà nghiên cứu quan tâm đến là tỷ lệ người phụ thuộc, đó là những người ngoài độ tuổi lao động, trong nghiên cứu xét đến là những người dưới 15 tuổi và trên 60 tuổi, tỷ lệ này chiếm đến 62,81%.

Nghiên cứu tập trung vào khu vực nông thôn Việt Nam và chia làm 6 vùng sinh thái: Bắc Trung Bộ và duyên hải miền Trung (22,3%), Đồng bằng sông Hồng (22,2%), Đồng bằng sông Cửu Long (20,9%), Đông Nam Bộ (6,8%), Tây Nguyên (7,6%), Trung du và Miền núi phía Bắc (20,2%).

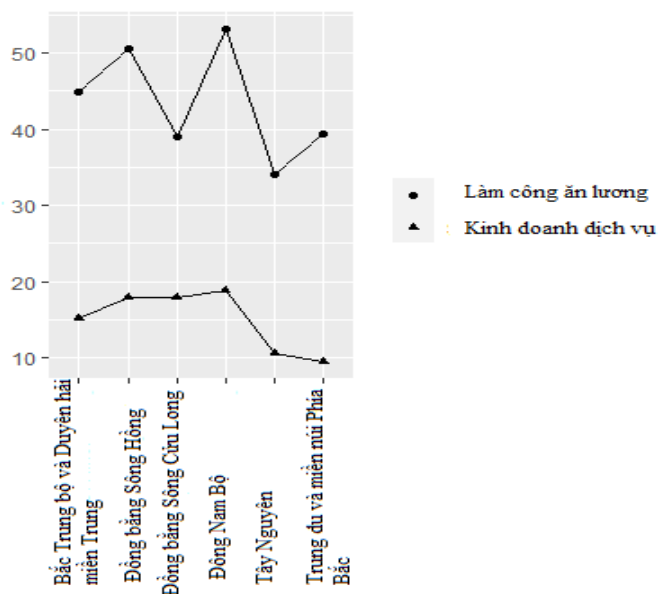
3.2. Thực trạng đa dạng nguồn thu nhập của nông hộ

Hình 1 cho thấy, nguồn thu phi nông nghiệp chiếm tỷ trọng lớn nhất ở hầu hết các vùng miền với mức trung vị đa số từ 50% trở lên (trừ khu vực Tây Nguyên) và cũng có sự dao động rất lớn. Tiếp theo là tỷ trọng thu nhập từ nông nghiệp, mức trung vị cao nhất là khu vực Tây Nguyên khoảng 40%, mức trung vị thấp nhất là Đông Nam Bộ xấp xỉ 0%. Nguồn thu nhập khác chiếm tỉ trọng ít nhất.



Hình 1. Biểu đồ hộp về tỷ lệ các nguồn thu nhập
(Nguồn: Tính toán của tác giả từ Điều tra mức sống dân cư 2020)

Hình 2 là biểu đồ đường về tỷ lệ thu nhập phi nông nghiệp trong thu nhập hộ gia đình nông thôn. Hình 2 chia tổng thu nhập phi nông nghiệp thành hai thành phần: Đường có kí hiệu chấm tròn là thu nhập phi nông nghiệp từ tiền công, tiền lương; đường có kí hiệu hình tam giác là thu nhập phi nông nghiệp từ các ngành nghề sản xuất kinh doanh và dịch vụ.



Hình 2. Biểu đồ đường về tỷ lệ thu nhập phi nông nghiệp
(Nguồn: Tính toán của tác giả từ Điều tra mức sống dân cư 2020)

Hình 2 cho thấy ở tất cả các vùng miền thì thu nhập từ tiền lương chiếm tỷ lệ cao hơn nhiều so với thu từ ngành nghề sản xuất kinh doanh và dịch vụ. Vùng Đông Nam Bộ có thu nhập từ tiền lương cao nhất chiếm khoảng 53%, nguồn thu từ sản xuất kinh doanh và dịch vụ khoảng 19%; như vậy, tổng là khoảng 72% thu nhập từ hoạt động phi nông nghiệp. Vùng Tây Nguyên có thu nhập từ hoạt động phi nông nghiệp thấp nhất, khoảng 34% thu nhập từ làm công ăn lương và 10% từ kinh doanh dịch vụ, tổng thu nhập từ hoạt động phi nông nghiệp ở Tây Nguyên khoảng 44%. Các vùng còn lại có tổng thu nhập từ các hoạt động phi nông nghiệp là: Trung du và miền núi phía Bắc (49%), Đồng bằng sông Cửu Long (57%), Bắc Trung Bộ và duyên hải miền Trung (60%), Đồng bằng sông Hồng (68%). Từ đây cho thấy đối với các vùng phát triển, đông dân cư như Đông Nam Bộ và Đồng bằng sông Hồng thì thu nhập từ hoạt động phi nông nghiệp chiếm tỷ lệ cao, còn đối với các vùng kém phát triển hơn, dân cư thưa thớt hơn như Tây Nguyên và Trung du miền núi phía Bắc chiếm tỷ lệ thấp hơn nhiều.

3.3. Ảnh hưởng của đa dạng sinh kế tới thu nhập của nông hộ

Hệ số hồi quy của các mô hình được thể hiện trong Bảng 3. Theo đó, các biến đổi tỉ lệ logarit đẳng cự ILR đều có ý nghĩa thống kê trong cả 2 mô hình. Các hệ số hồi quy ILR giải thích ảnh hưởng của cơ cấu các nguồn thu phi nông nghiệp đến thu nhập của nông hộ. Cụ thể, kết quả được giải thích như sau: nếu thu nhập phi nông nghiệp từ tiền lương tăng lên $\delta\% = 1\%$ trong khi các thu nhập phi nông nghiệp khác không thay đổi thì thu nhập hộ gia đình tăng lên bình quân khoảng $0,02 \times \sqrt{\frac{2}{3}} \ln(1 + \delta)\% = 0,011\%$. Tương tự như vậy, theo hệ số mô hình (2), nếu thu nhập phi nông nghiệp từ dịch vụ tăng lên $\delta\% = 1\%$ trong khi các thu nhập phi nông nghiệp khác không thay đổi thì thu nhập hộ gia đình tăng lên bình quân khoảng $0,04 \times \sqrt{\frac{2}{3}} \ln(1 + \delta)\% = 0,023\%$. Với các hệ số đã thu được từ mô hình hồi quy đa hợp, ta cũng hoàn toàn có thể tính toán được sự thay đổi của thu nhập hộ gia đình khi giá trị từ số trong phép biến đổi ILR có thay đổi và chỉ cần giữ nguyên giá trị trung bình nhân của các yếu tố khác. Tuy nhiên, kết quả cụ thể phụ thuộc vào giá trị độ thay đổi cụ thể của các yếu tố còn lại.

Kết quả hồi quy cũng cho thấy, có sự gia tăng về thu nhập trong hộ nếu chủ hộ đã kết hôn, hoặc chủ hộ là người Kinh so với các dân tộc thiểu số khác. Khi chủ hộ có bằng cấp càng cao thì thu nhập của hộ càng có xu hướng cao hơn. Điều này là phù hợp với thực tiễn cuộc sống. Đối với các nông hộ ở nước ta, có một số chỉ tiêu cũng phản ánh mức sống của hộ gia đình như chỉ tiêu về việc sử dụng nước sạch và nhà vệ sinh đạt tiêu chuẩn. Kết quả hồi quy mang dấu âm và có ý nghĩa thống kê trên mô hình cũng cho thấy việc hộ gia đình có sử dụng nước sạch và có nhà vệ sinh hợp tiêu chuẩn có thu nhập bình quân cao hơn so với các hộ không có các yếu tố trên. Điều này cũng đã được đề cập đến trong một số nghiên cứu về đa dạng sinh kế với thu nhập nông hộ [8] và chương trình nông thôn mới [16]. Nếu phân chia theo khu vực sinh thái, các nông hộ ở Đồng bằng sông Hồng, Đồng bằng sông Cửu Long và Đông Nam Bộ có thu nhập bình quân cao hơn so với khu vực Bắc Trung bộ và duyên hải miền Trung.

Bảng 3. Hệ số hồi quy

Biến quan sát	Logarit thu nhập hộ gia đình	
	Mô hình (1)	Mô hình (2)
Hệ số chặn	11,01 ***(0,06)	
ILR1	0,02 ***(0,00)	0,04 ***(0,00)
ILR2	0,03 ***(0,00)	0,03 ***(0,00)
Giới tính chủ hộ (mặc định: Nam)	Nữ	0,00 (0,02)
Tuổi chủ hộ		0,00 ***(0,00)
Hôn nhân chủ hộ (mặc định: Khác)	Kết hôn	0,20 ***(0,03)
Dân tộc chủ hộ (mặc định: Kinh)	Dân tộc thiểu số	-0,37 ***(0,02)
Tổng số thành viên hộ		0,20 ***(0,01)

Biến quan sát		Logarit thu nhập hộ gia đình	
		Mô hình (1)	Mô hình (2)
Nghề nghiệp chủ hộ (mặc định: Kinh doanh dịch vụ)	Làm công ăn lương	-0,08 ** (0,03)	
	Nông lâm thủy sản	-0,05 * (0,02)	
Bằng cấp chủ hộ (Mặc định: Không bằng cấp)	Tiểu học	0,10 *** (0,02)	
	Trung học cơ sở, trung học phổ thông	0,23 *** (0,02)	
	Đại học	0,60 *** (0,04)	
Tỉ lệ người phụ thuộc		0,00 *** (0,00)	
Gia đình có sử dụng nước sạch (Mặc định: Có)	Không	-0,15 *** (0,02)	
Gia đình có nhà vệ sinh đạt tiêu chuẩn (Mặc định: Có)	Không	-0,27 *** (0,02)	
Diện tích đất nông nghiệp		0,00 *** (0,00)	
	Đồng bằng sông Hồng	0,24 *** (0,02)	
	Đồng bằng sông Cửu Long	0,25 *** (0,02)	
	Đông Nam Bộ	0,41 *** (0,03)	
	Tây Nguyên	-0,06 (0,03)	
Khu vực sinh thái (Mặc định: Bắc Trung bộ và Duyên hải miền Trung)	Trung du và miền núi phía Bắc	-0,04 (0,03)	
Hệ số R2 hiệu chỉnh		0,51	0,51

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

$$\text{Mô hình (1) sử dụng } ILR_1 = \sqrt{\frac{2}{3}} \ln \frac{S_{NW}}{\sqrt{S_{NS}S_{OT}}} = \sqrt{\frac{2}{3}} \ln \frac{V_{NW}}{\sqrt{V_{NS}V_{OT}}}; \quad ILR_2 = \sqrt{\frac{1}{2}} \ln \frac{S_{NS}}{S_{OT}} = \sqrt{\frac{1}{2}} \ln \frac{V_{NS}}{V_{OT}}.$$

$$\text{Mô hình (2) sử dụng } ILR_1 = \sqrt{\frac{2}{3}} \ln \frac{S_{NS}}{\sqrt{S_{NW}S_{OT}}} = \sqrt{\frac{2}{3}} \ln \frac{V_{NS}}{\sqrt{V_{NW}V_{OT}}}; \quad ILR_2 = \sqrt{\frac{1}{2}} \ln \frac{S_{NW}}{S_{OT}} = \sqrt{\frac{1}{2}} \ln \frac{V_{NW}}{V_{OT}}.$$

(Nguồn: Tính toán của tác giả từ Điều tra mức sống dân cư 2020)

4. Kết luận

Nghiên cứu của nhóm tác giả đã phân tích ảnh hưởng của đa dạng sinh kế đến đời sống của người dân khu vực nông thôn Việt Nam, trong đó, đặc biệt tập trung phân tích sự thay đổi của cơ cấu các nguồn thu nhập phi nông nghiệp. Nghiên cứu sử dụng phương pháp phân tích số liệu đa hợp (Compositional data analysis, CoDa) – một phương pháp thống kê hiện đại và chuyên biệt để nghiên cứu về cơ cấu các thành phần, dựa trên bộ số liệu điều tra Khảo sát mức sống dân cư Việt Nam năm 2020. Kết quả nghiên cứu cho thấy, các yếu tố ảnh hưởng đến thu nhập hộ gia đình là: tuổi chủ hộ, tình trạng hôn nhân của chủ hộ, dân tộc, tổng số thành viên hộ, tỷ lệ người phụ thuộc, bằng cấp chủ hộ, gia đình có sử dụng nước sạch, gia đình có nhà vệ sinh hợp tiêu chuẩn, diện tích đất nông nghiệp, khu vực sống của hộ gia đình. Ngoài ra, thu nhập từ các hoạt động phi nông nghiệp ảnh hưởng tích cực đến thu nhập hộ gia đình: nếu thu nhập phi nông nghiệp từ tiền lương tăng, thu nhập phi nông nghiệp từ dịch vụ tăng trong khi các thu nhập phi nông nghiệp khác không thay đổi thì thu nhập hộ gia đình tăng lên. Để tăng thu nhập hộ gia đình thì đa dạng hóa sinh kế là vấn đề cần được Đảng và Nhà nước quan tâm, đặc biệt chú trọng phát triển các nguồn thu nhập phi nông nghiệp phù hợp với đặc điểm của từng địa phương. Hạn chế của nghiên cứu này là chỉ áp dụng trên bộ số liệu điều tra năm 2020, nhóm nghiên cứu đề xuất thực hiện các nghiên cứu cụ thể và chi tiết hơn trên bộ số liệu mảng về hộ gia đình. Đồng thời, các công cụ đo lường mới về đa dạng sinh kế hộ gia đình nên tiếp tục được nghiên cứu.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] D. Benjamin, L. Brandt, and B. McCaig, "Growth with equity: income inequality in Vietnam, 2002–14," *J. Econ. Inequal.*, vol. 15, no. 1, pp. 25–46, 2017, doi: 10.1007/s10888-016-9341-7.
- [2] P. Baumann, *Improving access to natural resources for the rural poor*, FAO LSP Paper (FAO), Rome, 2002.
- [3] F. Ellis, "Household Strategies and Rural Livelihood Diversification," *J. Dev. Stud.*, vol. 35, no. 1, pp. 1–38, 1998.

-
- [4] O. Oyino and K. T. Olaleye, "Farm Households Livelihood Diversification and Poverty Alleviation in Giwa Local Government Area of Kaduna State, Nigeria," *Consilience*, no. 15, SE-Field Notes, Feb. 2016, doi: 10.7916/consilience.v0i15.3914.
- [5] Y. Gecho, "Rural Household Livelihood Strategies: Options and Determinants in the Case of Wolaita Zone, Southern Ethiopia," *Soc. Sci.*, vol. 3, no. 3, pp. 92 - 104, 2014, doi: 10.11648/j.ss.20140303.15.
- [6] H. T. Pham, B. A. Tuan, and T. D. Le, "Is Nonfarm Diversification a Way Out of Poverty for Rural Households? Evidence from Vietnam in 1993-2006," *SSRN Electron. J.*, no. 2010-17, 2010, doi: 10.2139/ssrn.1715603.
- [7] V. Q. Khuc, Q. B. Tran, and L. S. Hoang, "Analysis of factors affecting household income diversification in the buffer zone of U Minh Ha National Park, Ca Mau (in Vietnamese)," *J. Agric. Rural Dev.*, vol. 1, no. 1, pp. 1-11, 2016.
- [8] T. N. D. Ho and V. H. Thuc, "Determinants of income diversification and its effects on rural household income in Vietnam," *Can Tho Univ. J. Sci.*, vol. 06, pp. 153-162, 2017, doi: 10.22144/ctu.jen.2017.039.
- [9] K. G. van den Boogaart and R. Tolosana-Delgado, *Analyzing Compositional Data with R*, vol. 122, Berlin: Springer, 2013.
- [10] J. Aitchison, *The statistical analysis of compositional data*. Chapman and Hall, London, 1986.
- [11] J. J. Egozcue, V. Pawłowsky-Glahn, G. Mateu-Figueras, and C. Barceló-Vidal, "Isometric logratio transformations for compositional data analysis," *Math. Geol.*, vol. 35, no. 3, pp. 279-300, 2003.
- [12] V. T. Le, T. T. T. Dam, and T. H. Trinh, "Factors effecting expenditure shares of international travel tourism to Vietnam: A compositional data analysis approach," *Can Tho Univ. J. Sci.*, vol. 56, no. 4, pp. 208-218, 2020.
- [13] T. V. A. To, N. A. Pham, and T. H. Trinh, "Determinants of Vietnamese household expenditure for school education in 2020," *Can Tho Univ. J. Sci.*, vol. 58, no. 4, pp. 37-49, 2022.
- [14] General Statistics Office, *Report of labor force survey 2020*. Department of Population and Labour statistics, 2021.
- [15] T. N. Nguyen, T. T. M. Nguyen, T. T. T. Dam, and T. H. Trinh, "Impact of livelihood diversity on household income in rural areas in Vietnam," in *Proceedings of the National Workshop on Quantitative Analysis of Economic and Social Issues in the Digital Environment*, 2022, pp. 106-118.
- [16] Q. V. Hoang, "Determinants of the result of new rural development program in Vietnam," *J. Econ. Dev.*, vol. 22, no. 1, pp. 81-90, 2020, doi: 10.1108/jed-12-2019-0076.