



DOI:10.22144/ctu.jvn.2022.185

## PHÂN TÍCH YẾU TỐ ẢNH HƯỞNG ĐẾN CƠ CẤU CHI TIÊU GIÁO DỤC PHỔ THÔNG CỦA HỘ GIA ĐÌNH VIỆT NAM NĂM 2020

Tô Thị Vân Anh<sup>1\*</sup>, Phạm Ngọc Anh<sup>2</sup> và Trịnh Thị Hương<sup>3</sup>

<sup>1</sup>Phòng Thống kê, Cục Công nghệ Thông tin, Bộ Giáo dục và Đào tạo

<sup>2</sup>Bộ môn Toán, Khoa Khoa học Cơ bản, Trường Đại học Mô Địa Chất

<sup>3</sup>Bộ môn Toán, Trường Đại học Thương mại

\*Người chịu trách nhiệm về bài viết: Tô Thị Vân Anh (email: ttvanh.vpb@moet.gov.vn)

### Thông tin chung:

Ngày nhận bài: 23/02/2022

Ngày nhận bài sửa: 19/04/2022

Ngày duyệt đăng: 22/04/2022

### Title:

Analysis of factors affecting the general education expenditure structure of Vietnamese households in 2020

### Từ khóa:

Chi tiêu giáo dục, học phí, đóng góp, điều tra mức sống dân cư, biểu đồ tam giác, phân tích đa hợp

### Keywords:

Compositional data analysis, donations, educational expenditure, ternary diagram, tuition, Vietnam Household Living Standard Survey

### ABSTRACT

Educational expenditure and its components are essential indicators of the household and society. The 2020 Vietnam Household Living Standard Survey data set, including 5491 pupils studying at all levels of education and the demographic characteristics of their households was used in this study. The cost for a school year and the proportion of each share increases as the education level increases. Tuition is a relatively small expense compared to many other expenses such as extra classes, donations, and textbooks. Using compositional data analysis (CODA), the result showed that household income, type of school, living area, and regions statistically significant impact on the expenditure structure. In particular, the gender of learners does not affect education spending. The research results provide evidence of investment in education for the government and households and suggest policy implications for increasing investment efficiency in the general education system.

### TÓM TẮT

Chi tiêu giáo dục và các khoản mục là chỉ tiêu quan trọng của hộ gia đình và cả xã hội. Bộ số liệu điều tra Mức sống dân cư 2020 gồm 5491 học sinh đang theo học tại các cấp học phổ thông và đặc điểm nhân khẩu học của hộ gia đình tương ứng được sử dụng trong nghiên cứu. Chi phí cho 1 năm học và tỉ trọng từng khoản mục đều tăng theo cấp học, trong đó, học phí là một khoản chi nhỏ so với nhiều khoản chi khác như học thêm, đóng góp, sách giáo khoa. Thông qua phương pháp phân tích đa hợp, kết quả phân tích cho thấy thu nhập hộ gia đình, loại hình trường học, và khu vực sinh sống tác động có ý nghĩa thống kê đến cơ cấu chi tiêu theo các khoản mục thiết yếu bắt buộc, đóng góp và học thêm. Đặc biệt, giới tính của người học không ảnh hưởng đến chi tiêu giáo dục. Kết quả nghiên cứu cung cấp thêm bằng chứng về đầu tư cho giáo dục của cả Chính phủ và hộ gia đình, đồng thời gợi ý các hàm ý chính sách để tăng hiệu quả đầu tư cho giáo dục ở các cấp học phổ thông.

## 1. GIỚI THIỆU

Việt Nam luôn coi giáo dục và đào tạo là trọng tâm phát triển bền vững của đất nước và đã được

Chính phủ Việt Nam cam kết thực hiện, thể hiện ở mục tiêu phát triển bền vững số 4 về “Giáo dục có chất lượng” trong 17 mục tiêu. Đầu tư cho giáo dục

được đảm bảo cả từ nguồn ngân sách Nhà nước, các tầng lớp kinh tế khác nhau và cả người dân. Cùng với sự phát triển của kinh tế - xã hội, người dân ngày càng quan tâm đến tương lai của thế hệ trẻ và sẵn sàng đầu tư cho giáo dục với mức chi ngày một tăng cao. Kết quả khảo sát mức sống dân cư 2020 cho thấy đầu tư cho giáo dục ngày càng được chú trọng hơn qua các năm. Năm 2020, trung bình các hộ dân cư phải chi hơn 7,0 triệu đồng cho một thành viên đang đi học, tăng khoảng 7,0% so với năm 2018 (Tổng cục Thống kê, 2021). Tuy nhiên, tỷ lệ chi tiêu cho giáo dục chiếm tỷ trọng cao trong tổng chi tiêu của hộ gia đình cũng là một gánh nặng tương đối lớn đối với phần đông các hộ gia đình ở Việt Nam hiện nay. Trong khi đó, mức độ thiếu hụt dịch vụ xã hội cơ bản về khía cạnh giáo dục, cụ thể là trình độ giáo dục của người lớn và tình trạng đi học của trẻ em là các chỉ số được dùng để đo lường, đánh giá mức độ nghèo đa chiều ở Việt Nam. Khi gánh nặng chi trả cho các khoản trong giáo dục của hộ gia đình quá lớn, các vấn đề về công bằng và khả năng tiếp cận giáo dục có thể phát sinh. Vì vậy, thông tin về chi tiêu giáo dục rất quan trọng để đánh giá chi phí trực tiếp cho giáo dục trong các hộ gia đình và ảnh hưởng của nó đối với tổng thể phúc lợi chung. Từ đó, các nhà hoạch định chính sách có bằng chứng thực nghiệm để nâng cao hơn nữa việc thực thi các chính sách giáo dục và đào tạo để san sẻ gánh nặng với các hộ gia đình, đảm bảo công bằng xã hội và phát triển bền vững giáo dục.

Chi tiêu giáo dục của hộ gia đình biểu thị tất cả các chi phí mà một hộ gia đình phải chi trả cho việc giáo dục của một hoặc nhiều thành viên, đặc trưng không chỉ bởi độ lớn của từng khoản mục, mà còn đặc trưng bởi cơ cấu chi cho giáo dục, trong đó cho tổng tỷ trọng các khoản chi là 100% và tỉ trọng các khoản mục là phụ thuộc lẫn nhau. Cơ cấu chi tiêu giáo dục của các hộ gia đình phụ thuộc vào loại hình đào tạo, các đặc điểm cá nhân như giới tính, độ tuổi, cấp học, đặc điểm kinh tế xã hội của từng địa phương và đất nước... Hiện nay, các phân tích về chi tiêu cho giáo dục của các hộ gia đình tại Việt Nam thường chú trọng vào mức chi tiêu, tức giá trị tuyệt đối của mức chi tiêu và sử dụng các thống kê mô tả cơ bản khi phân tích cơ cấu chi tiêu. Những hạn chế trên xuất phát từ thực tế là tỷ trọng trong cơ cấu chi tiêu giáo dục phụ thuộc lẫn nhau, mang tính chất tương đối giữa các thành phần. Một số nghiên cứu trước đây như Ngoan và ctv. (2021), Tiên và Lê (2014) sử dụng mô hình hồi quy Tobit để đánh giá các ảnh hưởng, đến chi tiêu cho giáo dục của hộ gia đình. Các nghiên cứu về chi tiêu giáo dục của từng cá nhân và ảnh hưởng của gia đình, giới tính người

học được thực hiện ở nhiều quốc gia. Nghiên cứu của Aslam và Kingdon (2008) sử dụng mô hình Tobit, nhóm tác giả tìm thấy sự khác biệt giữa về giới tính của người học trong chi tiêu giáo dục. Nghiên cứu của Kuvat và Kizilgöl (2020) tìm thấy ảnh hưởng của các yếu tố đặc điểm kinh tế xã hội của hộ gia đình đến chi tiêu giáo dục của hộ gia đình năm 2017. Tuy mô hình Tobit phù hợp với khoản chi tiêu giáo dục không âm nhưng nhược điểm của mô hình này là chỉ đánh giá được tổng chi tiêu giáo dục cho từng cấp của một cá nhân chứ không đánh giá được cơ cấu từng khoản mục. Trong khi đó, cơ cấu chi tiêu giáo dục được xem xét như một vectơ đa hợp và được phân tích bằng phương pháp phân tích số liệu đa hợp (Compositional Data Analysis, CODA) (Pawlowsky-Glahn et al., 2015). Mô hình phân tích số liệu đa hợp CODA gần đây được áp dụng rộng rãi trong các nghiên cứu về kinh tế xã hội khi nghiên cứu về cơ cấu chi tiêu (Tuấn và ctv., 2020).

Bộ dữ liệu Điều tra mức sống Dân cư Việt Nam (VHLSS) 2020 mới nhất đại diện cho toàn quốc, các thống kê mô tả, biểu đồ tam giác và phương pháp phân tích số liệu đa hợp được sử dụng. Hai vấn đề cụ thể được tập trung nghiên cứu là: 1) thực trạng chi tiêu cho mỗi cá nhân theo hệ thống giáo dục phổ thông của các hộ gia đình, 2) phân tích các nhân tố tác động đến cơ cấu chi tiêu cho một số khoản mục chính của các cấp học trong chương trình giáo dục phổ thông Việt Nam.

## 2. PHƯƠNG PHÁP NGHIÊN CỨU

### 2.1. Dữ liệu

Bộ dữ liệu sử dụng trong nghiên cứu được lấy từ cuộc khảo sát Điều tra mức sống Dân cư Việt Nam năm 2020, được thực hiện bởi Tổng cục Thống kê Việt Nam. Đây là cuộc khảo sát được tiến hành 02 năm một lần trên phạm vi toàn quốc bằng phương pháp phỏng vấn trực tiếp chủ hộ. Kết quả điều tra nhằm thu thập thông tin để tổng hợp, biên soạn các chi tiêu thống kê quốc gia về mức sống hộ dân cư, đáp ứng nhu cầu thông tin thống kê cho các cấp, các ngành để đánh giá mức sống, tình trạng nghèo đói và phân hoá giàu nghèo của dân cư theo hướng tiếp cận đa chiều. Thông tin thống kê từ Khảo sát mức sống 2020 bảo đảm mức độ đại diện cho cả nước, khu vực thành thị - nông thôn, 6 vùng kinh tế - xã hội và tỉnh, thành phố trực thuộc Trung ương (Tổng cục Thống kê, 2021). Cấp hộ gia đình gồm 9.388 hộ được sử dụng, hạn chế trên các hộ gia đình có các thành viên đang theo học trên ba cấp học của giáo dục phổ thông là tiểu học, trung học cơ sở (THCS) và trung học phổ thông (THPT). Sau quá trình xử lý

sạch số liệu, dữ liệu được thu gồm 5.491 thành viên đang theo học trong chương trình giáo dục phổ thông và các đặc điểm hộ gia đình tương ứng. Cuộc khảo sát gồm các mục chính như sau:

Mục về một số đặc điểm nhân khẩu học cơ bản liên quan đến mức sống: Mục này được sử dụng để tạo các biến điều tra về đặc điểm chủ hộ như trình độ giáo dục, giới tính, tuổi, số thành viên hộ, nơi sống và vùng kinh tế - xã hội. Ngoài ra, thông tin giáo dục của các thành viên trong hộ cũng được điều tra: Thông tin về số người đang học trong 12 tháng qua, thông tin về cấp học đang theo học.

Mục chỉ tiêu giáo dục: Mục này thu thập thông tin về nhiều loại chỉ tiêu, trong đó có chỉ tiêu cho các thành viên trong hộ đi học trong 12 tháng qua cho các môn học theo quy định của nhà trường. Từ đó, tổng chỉ tiêu giáo dục của hộ gia đình được tính toán. Trong đó, các khoản chi bao gồm: học phí; đóng góp cho trường, lớp (quỹ xây dựng...); trái tuyển; quần áo đồng phục và trang phục theo quy định; sách giáo khoa, sách tham khảo; dụng cụ học tập (giấy, bút, cặp, vở...); học thêm cho môn học thuộc chương trình quy định; chi phí giáo dục khác (lệ phí thi, đi lại, trọ, bảo hiểm thân thể,...). Bên cạnh đó, các thông tin về loại hình trường học (công lập, tư thục/dân lập) có hưởng trợ cấp giáo dục hay không cũng được thu thập cho từng cá nhân đang theo học. Nghiên cứu này hạn chế trên các cấp học của giáo dục phổ thông, ngoài ra nghiên cứu cũng có thành viên đang học mẫu giáo, trường nghề, đại học và cao hơn.

**2.2. Phương pháp phân tích đa hợp**

Khái niệm về phân tích số liệu đa hợp CODA dựa trên các nghiên cứu của Van den Boogaart và Tolosana-Delgado (2013), Pawlowsky-Glahn và ctv. (2015) trên nền tảng về độ đo Aitchison (Aitchison, 1986). CODA là một lĩnh vực thống kê lâu đời với các lĩnh vực ứng dụng đa dạng, chẳng hạn như địa chất hoặc kinh tế (Pawlowsky-Glahn et al., 2015). Phương pháp này gần đây đã được áp dụng trong y tế và nghiên cứu dịch tễ học dinh dưỡng (Dumuid et al., 2018; Trinh et al., 2018).

**2.2.1. Một số định nghĩa**

Dữ liệu đa hợp được biểu diễn bởi một vectơ đa hợp S gồm D thành phần được biểu diễn trong đơn hình (simplex)  $S^D$ ,

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

trong đó, tổng các các thành phần của một véc tơ x bằng T và là một số, thường được chọn T = 1 hoặc T = 100 (%).

Mỗi vectơ  $x \in S^D$  có thể được chuẩn hóa, kí hiệu  $C(x)$ , qua phép chuẩn hóa sau:

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

Do các thành phần của các vectơ 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), cụ thể là tỷ lệ bất biến (scale invariance); hoán vị bất biến (permutation invariance) và vectơ con nhất quán (subcompositional coherence).

Cho hai vectơ đa hợp trong đơn hình  $S^D$  và  $\alpha \in R$ , các phép toán cơ bản là

$$x = (x_1, x_2, \dots, x_D),$$

$$y = (y_1, y_2, \dots, y_D)$$

- Cộng hai vectơ (Perturbation)

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

- Tích của một vectơ và một số thực (Power transformation)

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

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

Bên cạnh đó, vectơ đa hợp trong đơn hình  $S^D$  thường được chuyển đổi về không gian các số thực R, tức không gian Euclide, để thuận lợi tính toán và giải thích thông qua một số phép biến đổi dựa trên phép toán lôgarit. Trong nghiên cứu này, phép biến đổi lôgarit đẳng cự (isometric log-ratio, ILR) được sử dụng để chuyển đổi D tỉ trọng trong đơn hình về vectơ gồm D-1 tọa độ trong không gian Euclide thông thường (Egozcue et al., 2003). ILR được thực hiện dựa trên sự phân nhóm không giao giữa các khoản mục trong vectơ đa hợp. Cụ thể, để tạo tọa độ cân bằng đầu tiên, D tỉ trọng được chia thành 2 nhóm: một nhóm cho từ 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 từ 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 từ 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:

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

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 vectơ đa hợp và tổng của các thành phần, tuy nhiên hệ số hồi quy khi đó không được phân tích trực quan mà phải qua các tính toán thêm. Trong nghiên cứu này, chúng tôi sử dụng mô hình có vectơ đa hợp là biến phụ thuộc.

2.2.2. *Mô hình hồi quy đa hợp trong cơ cấu chi tiêu giáo dục*

Dữ liệu gồm 5.491 quan sát đại diện cho 06 vùng kinh tế - xã hội của Việt Nam, khoản chi cho giáo dục gồm 08 khoản chi nêu trên. Dựa vào số liệu và các ưu điểm của đơn hình trong không gian 3 chiều  $S^3$  như biểu đồ tam giác và thuận lợi trong chuyển đổi ILR. Các khoản mục được gộp thành 03 khoản mục chi. Thứ nhất là khoản chi tiêu thiết yếu bắt buộc gồm học phí; quần áo đồng phục và trang phục theo quy định; sách giáo khoa, sách tham khảo. Thứ hai là các khoản đóng góp và học thêm bao gồm đóng góp cho trường, lớp; học thêm cho môn học thuộc chương trình quy định và chi giáo dục khác là nhóm các khoản chi không bắt buộc. Thứ 3 là dụng cụ học tập và trái tuyến.

Kí hiệu  $V_1, V_2, V_3$  lần lượt là độ lớn của ba khoản mục chi cho giáo dục tương ứng như trên, tổng chi tiêu là  $T = V_1 + V_2 + V_3$  và  $S = (s_1, s_2, s_3) \in S^3$ ,  $s_i = \frac{V_i}{T}$ ,  $i = 1, 2, 3$  là đơn hình của không gian các vectơ đa hợp.

Với trường hợp  $D = 3$ , bảng phân tổ được trình bày như 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:

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

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  $s_3$ .

Tác động của các đặc điểm hộ gia đình Việt Nam đến cơ cấu chi tiêu cho giáo dục được quan tâm, mô hình hồi quy đa hợp đơn hình có dạng:

$$s_i = a \oplus_{k=1}^K X_{ki} \ominus b_k \oplus \epsilon_i, \quad i = 1, 2, \dots, n \quad (1)$$

Với  $s_i = (s_{i1}, s_{i2}, s_{i3})'$  và ký hiệu  $i$  là hộ gia đình thứ  $i$ , vectơ đa hợp  $s \in S^D$  và  $X_k$  là biến giải thích, gồm nhân khẩu học của cá nhân đang theo học hệ phổ thông (giới tính, loại hình trường học và có nhận hỗ trợ), đặc điểm của chủ hộ (trình độ giáo dục, nghề nghiệp, giới tính) và đặc điểm hộ gia đình (số thành viên hộ, số người đang đi học, nơi sống),  $a$  và  $b_k$  là hệ số hồi quy trong đơn hình,  $\epsilon$  là sai số của mô hình.

Mô hình (1) được biểu diễn qua các chuyển đổi ILR như sau với  $j = 1, 2$  và  $i = 1, 2, \dots, n$ :

$$ILR_{j,i} = a_j^* + \sum_{k=1}^K b_{j,k}^* X_{ki} + \epsilon_{j,i}^* \quad (2)$$

Các hệ số hồi quy của mô hình được ước lượng bằng phương pháp bình phương bé nhất (OLS). Các biến được lựa chọn vào mô hình dựa trên các nghiên cứu trước đó đã tiến hành về yếu tố ảnh hưởng đến chi tiêu giáo dục hộ gia đình (Tiến và Lê, 2014; Trang, 2020; Ngoan và ctv., 2021). Sau đó, thủ tục lùi từng bước (stepwise backward elimination) được thực hiện để lựa chọn các biến đưa vào mô hình. Với từng mô hình được thử nghiệm, kiểm định về đa cộng tuyến được thực hiện thông qua hệ số phóng đại phương sai (Variance inflation factor, VIF), kiểm định phương sai sai số thay đổi bằng kiểm định Breusch Pagan và kiểm định phân phối chuẩn chuỗi phần dư thông qua biểu đồ xác suất (q-q plot) và kiểm định Shapiro-Wilk.

Giải thích các hệ số ước lượng của mô hình hồi quy (2) tương tự như mô hình hồi quy tuyến tính 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 khi giữ cố định các nhân tố khác. Một yếu tố có tác động dương đến  $ILR_1$  và có tác động âm đến  $ILR_2$ , có nghĩa là khi tăng 1 đơn vị (hoặc so với phạm trù mặc định) sẽ làm tăng tỉ lệ chi tiêu cho thành phần  $s_1$  so với trung bình (hình học) hai thành phần  $s_2$  và  $s_3$ , đồng thời làm giảm tỉ lệ chi tiêu cho  $s_2$  so với  $s_3$ . Ngoài ra, hệ số mô hình hồi quy (2) có thể được biến đổi ngược về đơn hình và giải thích trực tiếp trên mô hình (1) nhưng tính toán phức tạp hơn (Morais & Thomas-Agnan, 2021). Trong trường hợp này, hệ số hồi quy của mô hình (1) là bất biến đối với các phân tổ khác nhau ở Bảng 1.

Các tính toán trong nghiên cứu này thực hiện trên phần mềm R, phiên bản 4.0.2 và các gói lệnh *tidyverse*, *tableone*, *Compositions*.

### 3. KẾT QUẢ VÀ THẢO LUẬN

#### 3.1. Đặc điểm đối tượng quan sát

Bảng 2 thể hiện đặc điểm của các cá nhân đang theo học hệ phổ thông; đặc điểm chủ hộ và gia đình tương ứng. Số lượng cá nhân theo học hệ tiểu học là nhiều nhất. Tỷ lệ học sinh là nam giới và nữ giới tương đối đồng đều. Các khoản trợ cấp cho việc miễn giảm học phí hoặc các khoản đóng góp cho giáo dục lớn nhất với cấp tiểu học (98,2%) và thấp nhất với cấp THPT (11%). Tỷ lệ này là phù hợp với chính sách Phổ cập giáo dục tiểu học và hỗ trợ học phí đối với học sinh tiểu học. Đối với hệ THCS và THPT, tỷ lệ học sinh nhận trợ cấp ít hơn so với bậc tiểu học. Trung bình chỉ cho một học sinh trong 1

năm giáo dục ở cấp THPT là lớn nhất và thấp nhất là chỉ tiêu cho giáo dục cấp tiểu học và độ dao động cũng tăng theo các cấp học. Về đặc điểm gia đình, thu nhập bình quân đầu người/tháng và phân theo cấp học có sự khác biệt không nhiều. Gia đình có học sinh tiểu học có mức thu nhập bình quân/người/tháng thấp nhất và cao nhất là các gia đình có thành viên học trung học cơ sở. Các gia đình nghiên cứu phân bố ở vùng nông thôn chiếm đa số (khoảng 70%) và phân bố ở cả 6 vùng kinh tế - xã hội. Số lượng hộ gia đình ở khu vực Tây Nguyên và Đông Nam Bộ là ít hơn các vùng khác. Về đặc điểm chủ hộ, họ đa số làm công ăn lương, đã kết hôn và sở hữu bằng trung học (THCS hoặc THPT).

**Bảng 2. Thống kê mô tả chung đặc điểm cá nhân đi học và hộ gia đình**

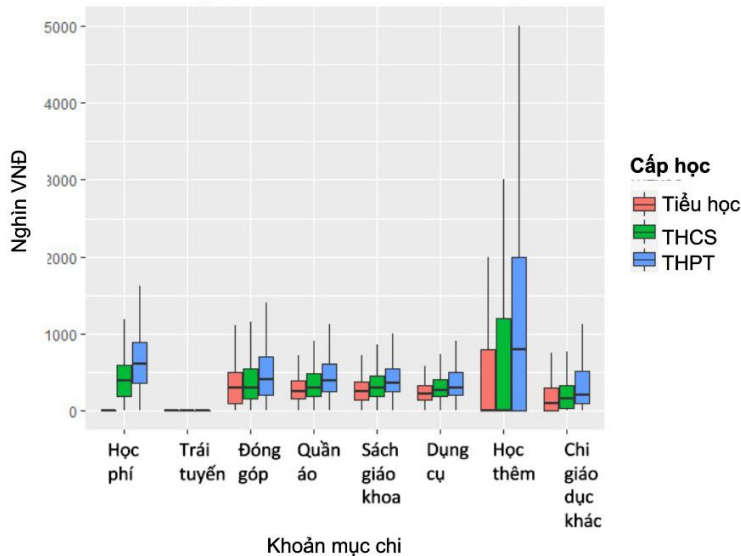
Biến quan sát		Cấp học		
		Tiểu học	Trung học cơ sở	Trung học phổ thông
Số quan sát		2749	1757	985
Giới tính người học (%)	Nữ	47,9	49,1	49,8
	Nam	52,1	50,9	50,2
Nhận được trợ cấp (%)	Không	1,8	78,7	89,0
	Có	98,2	23,1	11,0
Loại hình trường (%)	Công lập	99,3	99,5	96,8
	Dân lập - tư thục	0,7	0,5	3,2
Chi cho giáo dục (nghìn đồng/học sinh/năm học)		2399,78 (2623,26)	3820,34 (3989,59)	5874,44 (5255,06)
	Thu nhập bình quân (nghìn đồng/ tháng)	3481,52 (2603,37)	3696,37 (4863,24)	3950,96 (2738,51)
Số nam đang đi học		0,98 (0,77)	1,00 (0,77)	0,91 (0,74)
Số nữ đang đi học		1,00 (0,85)	1,02 (0,90)	0,92 (0,79)
Nơi sống (%)	Nông thôn	70,4	69,0	67,1
	Thành thị	29,6	31,0	32,9
Vùng kinh tế - xã hội (%)	Đồng bằng sông Hồng	22,5	22,3	24,1
	Trung du và Miền núi phía Bắc	19,9	20,1	18,3
	Bắc Trung Bộ và Duyên hải miền Trung	24,6	23,2	25,8
	Tây Nguyên	8,1	7,8	6,7
	Đông Nam Bộ	7,1	7,6	8,8
	Đồng bằng sông Cửu Long	17,8	19,1	16,3
Số người đi học trong hộ gia đình (%)	1	23,3	24,8	35,8
	2	59,2	54,6	48,1
	3 hoặc nhiều hơn	17,5	20,5	16,0
Tình trạng hôn nhân của chủ hộ (%)	Kết hôn	88,8	87,7	89,0
	Khác (độc thân, ly hôn)	11,2	12,3	11,0
Dân tộc chủ hộ (%)	Kinh	78,6	79,1	85,7
	Dân tộc khác	21,4	20,9	14,3
Nghề nghiệp chủ hộ (%)	Sản xuất kinh doanh hoặc dịch vụ	23,3	22,0	24,4
	Đi làm để nhận tiền công, tiền lương	51,7	53,0	53,0
	Hoạt động về nông, lâm, thủy sản	25,0	24,9	22,6
Trình độ học vấn của chủ hộ (%)	Không có bằng cấp	17,1	17,3	13,9
	Tiểu học	24,6	25,3	26,8
	Trung học (THCS hoặc THPT)	48,2	48,3	49,2
	Đại học	10,2	9,2	10,1

Ghi chú: Đối với các biến liên tục, thống kê gồm giá trị trung bình và lệch chuẩn (trong ngoặc). Đối với các biến rời rạc, giá trị tương ứng là tỷ lệ phần trăm

**3.2. Đặc điểm cơ cấu chi tiêu giáo dục cho từng cấp học**

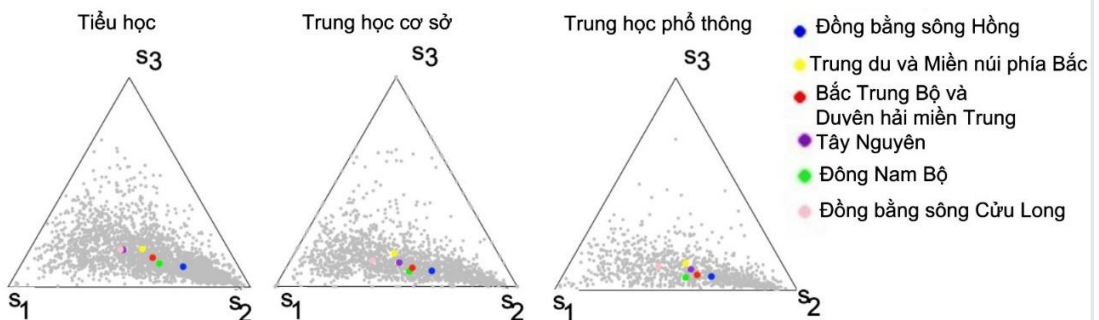
Hình 1 biểu thị biểu đồ hộp cơ cấu chi tiêu theo 8 khoản mục ở ba cấp học như trong phiếu khảo sát của Mức sống dân cư 2020. Như đã hiển hiện ở Bảng 2, cấp càng cao thì chi phí trong 1 năm học học sinh đó cũng cao hơn. Hình 1 cho thấy xu hướng tăng ở tất cả các khoản mục khi lên cấp học cao hơn. Cấp THPT có trung bình chi tiêu các khoản cho giáo dục cao hơn hẳn hai cấp còn lại, đặc biệt cao nhất là khoản chi cho việc học thêm, thấp nhất là các khoản chi cho trái tuyến. Mẫu nghiên cứu không ghi nhận các khoản thu trái tuyến. Trừ chi phí học thêm là khá

cao, đặc biệt là giá trị trên trung vị của cấp học tiểu học và THCS cho thấy xu hướng phân biệt rất cao cho việc học thêm ở các cấp học thấp (có khoảng 50% các học sinh là không đi học thêm). Ở cấp THPT, đường trung vị ở giữa hộp tứ phân vị cho thấy chi phí học thêm ở cấp này là ở hầu hết các học sinh. Các khoản chi về học phí, đóng góp, quần áo, sách giáo khoa và chi khác là khá đồng đều giữa các cấp học. Biểu đồ cũng thể hiện học sinh tiểu học được miễn học phí theo chủ trương của Nhà nước. Như đã miêu tả trong mục phương pháp nghiên cứu, 8 khoản mục được chia thành thành 3 khoản mục chính và thể hiện ở biểu đồ tam giác ở Hình 2.



**Hình 1. Biểu đồ hộp cơ cấu chi tiêu cho giáo dục theo khoản mục và theo cấp học**

Ghi chú: Biểu đồ hộp thể hiện 5 giá trị: giá trị nhỏ nhất, tứ phân vị thứ nhất (Q1) – đường dưới, trung vị – đường gạch, tứ phân vị thứ 3 (Q3) – đường trên của hộp chữ nhật và giá trị lớn nhất.



**Hình 2. Biểu đồ tam giác cơ cấu chi tiêu theo 3 khoản mục theo cấp học**

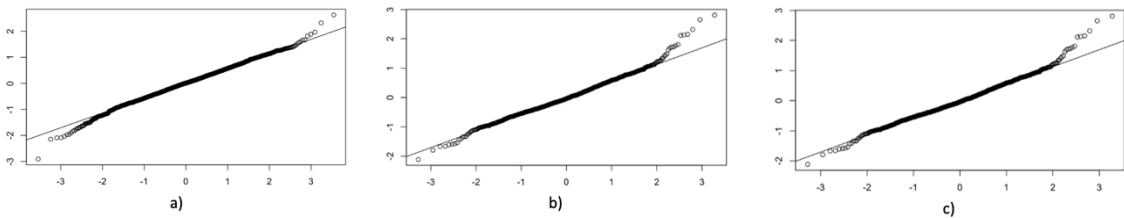
Biểu đồ tam giác thể hiện tỉ trọng của ba thành phần vector đa hợp sau khi đã chuẩn hóa. Trục 100% (hoặc 1 đơn vị) của từng thành phần được tính là đường cao từ đỉnh xuống cạnh đáy tương ứng của đỉnh đó, trong đó giá trị 0 là vị trí giao giữa đường cao và đáy tương ứng. Do đó, ứng với từng thành phần, nếu thành phần đó càng gần đỉnh tương ứng thì có tỉ trọng càng cao. Các điểm màu xám thể hiện từng quan sát và điểm tròn có màu thể hiện giá trị trung bình (hình học) của các vùng kinh tế - xã hội.

Biểu đồ tam giác cho thấy các điểm xám tập trung ở gần đỉnh  $s_1$  (khoản chi tiêu thiết yếu bắt buộc) và  $s_2$  (các khoản đóng góp và học thêm), đồng thời có xu hướng phân bố nhiều hơn gần đỉnh  $s_2$  đối với cả ba cấp học. Điều đó cho thấy tỉ trọng của khoản đóng góp và học thêm là lớn nhất và thấp nhất cho nhóm  $s_3$  (dụng cụ học tập và trái tuyền). Trung bình của 3 khoản chi trên theo từng vùng cũng thể hiện xu hướng như các điểm xám. Trong đó, tọa độ khu vực đồng bằng sông Hồng gần nhất với chi tiêu cho nhóm  $s_2$  và xa khu vực tập trung nhất cho  $s_2$  là đồng bằng sông Cửu Long.

**3.3. Yếu tố ảnh hưởng cơ cấu chi tiêu**

Bảng 3 thể hiện tác động của các đặc điểm nhân khẩu học cá nhân và hộ gia đình đến cơ cấu chi tiêu giáo dục phổ thông. Biến phụ thuộc ILR1 biểu thị mối quan hệ giữa lôgarít tỉ trọng  $s_1$  (khoản chi tiêu thiết yếu bắt buộc) đối với  $s_2$  (các khoản đóng góp; học thêm và chi giáo dục khác) và  $s_3$  (dụng cụ học tập và trái tuyền). Biến phụ thuộc ILR2 biểu thị mối quan hệ lôgarít giữa  $s_2$  và  $s_3$ . Thu nhập bình quân có tác động âm và ý nghĩa thống kê đối với ILR1 trong cả 3 cấp học, trong khi có tác động dương đối với ILR2. Do đó, nếu thu nhập bình quân hộ gia đình tăng thì làm giảm tỉ lệ % chi tiêu đối với các khoản chi tiêu thiết yếu bắt buộc khi giữ cố định trung bình của  $s_2$  và  $s_3$  (tức giữ cố định mẫu số). Hơn nữa, thu nhập bình quân tăng cũng dẫn đến tăng tỉ lệ chi tiêu của khoản đóng góp; học thêm và chi khác so với khoản chi về dụng cụ học tập và trái tuyền. Khi giữ cố định các nhân tố khác, so với loại

hình trường công lập, học sinh học tại trường dân lập và tư thục có tỉ trọng đóng góp cho các khoản chi bắt buộc cao hơn và có ý nghĩa thống kê so với trung bình tỉ lệ hai khoản chi còn lại. Tương tự, người học có nhận trợ cấp (bao gồm miễn giảm học phí và các khoản đóng góp cho giáo dục) so với người học không nhận trợ cấp làm tăng chi tiêu của khoản chi thiết yếu bắt buộc so với trung bình hai khoản chi còn lại. Hơn nữa, đối với cấp học THCS, người học có nhận trợ cấp có tỉ lệ chi tiêu của  $s_2$  cao hơn so với  $s_3$ . Giới tính của học sinh đi học và số lượng người đang theo học trong mỗi gia đình không có tác động có ý nghĩa thống kê đến cơ cấu chi tiêu. Bằng cấp của chủ hộ có tác động ý nghĩa thống kê đối với chi tiêu của bậc tiểu học (ILR2) và không có tác động ý nghĩa thống kê đối với bậc học THCS và THPT của người học. So với chủ hộ làm lĩnh vực sản xuất và dịch vụ, nghề nghiệp chủ hộ ít có tác động thống kê đối với các tỉ lệ chi tiêu. Nơi sống của hộ gia đình ở nông thôn hay thành thị không có tác động đến tỉ lệ chi tiêu của khoản chi thiết yếu bắt buộc so với hai khoản chi còn lại. Tuy nhiên, so với khu vực thành thị, ở cả 3 cấp học, hộ gia đình ở khu vực nông thôn có xu hướng chi tiêu của khoản  $s_2$  thấp hơn và có ý nghĩa thống kê so với khoản chi  $s_3$ . So với khu vực đồng bằng sông Cửu Long, các khu vực khác đều có mức chi cho tỉ trọng khoản chi tiêu thiết yếu bắt buộc nhỏ hơn hai khoản chi còn lại và xu hướng chi các khoản đóng góp; học thêm và chi khác lớn hơn so với tỉ trọng chi tiêu cho dụng cụ học tập và trái tuyền.



**Hình 3: Biểu đồ xác suất (q-q plot) của chuỗi phần dư cho mô hình ILR1 các cấp học**

Biểu đồ xác suất (q-q plot) của chuỗi phần dư cho mô hình ILR1 các cấp học: a) Cấp tiểu học; b) Cấp THCS và c) Cấp THPT tuân theo phân phối chuẩn.

**Bảng 3. Hệ số hồi quy của mô hình đa hợp tương ứng với biến ILR**

Biến ảnh hưởng		Cấp học						
		Tiểu học		THCS		THPT		
		ILR1	ILR2	ILR1	ILR2	ILR1	ILR2	
Hệ số tự do		1,39 ***	-0,88 ***	1,73 ***	-1,11 ***	1,30 ***	0,11	
Lôgarit thu nhập bình quân		-0,13 ***	0,18 ***	-0,14 ***	0,22 ***	-0,12 **	0,10	
Giới tính người học (mặc định: Nam)	Nữ	0,03	0	0,02	-0,02	0,01	0,07	
Trợ cấp (mặc định: Không)	Có	0,37 ***	0,19	0,27 ***	0,19 **	0,41 ***	0,10	
Loại hình (mặc định: Công lập)	Trường dân lập và tư thục	0,63 ***	0,18	0,87 ***	-0,17	0,48 ***	-0,04	
Số người đi học (mặc định: 1 người)	2 người	-0,03	0,05	-0,01	0,03	0,01	0,02	
	Từ 3 người	-0,02	-0,01	-0,03	0,05	0,01	-0,04	
Hôn nhân chủ hộ (mặc định: khác)	Đã kết hôn	0,01	0,05	-0,03	0	-0,02	-0,09	
Dân tộc chủ hộ (mặc định: Kinh)	Khác	0,17 ***	-0,37 ***	0,18 ***	-0,41 ***	0,25 ***	-0,47 ***	
Nghề nghiệp chủ hộ (mặc định: sản xuất kinh doanh hoặc dịch vụ)	Đi làm để nhận tiền công, tiền lương	0,05	0,01	-0,05	0,01	0,02	0,03	
	Hoạt động về nông, lâm, thủy sản	0,07	-0,01	-0,06	-0,08	-0,04	-0,06	
Trình độ học vấn của chủ hộ (mặc định: không có bằng cấp)	Tiểu học	-0,04	0,19 ***	0,07	-0,04	0,07	-0,04	
	Trung học (THCS hoặc THPT)	-0,06	0,16 **	0,02	0	0,07	-0,03	
	Đại học	-0,03	0,13	0,06	0,04	0,1	-0,04	
Nơi sống (mặc định: thành thị)	Nông thôn	0,04	-0,20 ***	-0,01	-0,20 ***	-0,02	-0,17 **	
	Đồng bằng sông Hồng	-0,45 ***	0,69 ***	-0,48 ***	0,64 ***	-0,40 ***	0,76 ***	
	Trung du và Miền núi phía Bắc	-0,36 ***	0,3 ***	-0,41 ***	0,37 ***	-0,47 ***	0,41 ***	
	Vùng kinh tế - xã hội (mặc định: đồng bằng sông Cửu Long)	Bắc Trung Bộ và Duyên hải miền Trung	-0,27 ***	0,36 ***	-0,33 ***	0,55 ***	-0,27 ***	0,69 ***
		Tây Nguyên	-0,12 *	0,14 *	-0,32 ***	0,41 ***	-0,40 ***	0,45 ***
		Đông Nam Bộ	-0,20 ***	0,43 ***	-0,23 ***	0,48 ***	-0,07	0,64 ***
R2 Hệ số điều chỉnh		0,14	0,23	0,12	0,24	0,09	0,17	

Mức độ ý nghĩa . 10%, \*5%, \*\*1%, \*\*\*0.1%.  $ILR_1 = \sqrt{\frac{2}{3}} \ln \frac{s_1}{\sqrt{s_2 s_3}}$ ;  $ILR_2 = \sqrt{\frac{1}{2}} \ln \frac{s_2}{s_3}$ . Các biến được lựa chọn vào mô hình thông qua thủ tục lùi từng bước. Hệ số phóng đại phương sai đều nhỏ hơn 2. Với mức ý nghĩa 5%, 4 mô hình hồi quy trong 6 mô hình không có hiện tượng phương sai sai số thay đổi bằng kiểm định Breusch Pagan. Với mức ý nghĩa 5%, 3 trong 6 mô hình có phần dư tuân theo phân phối chuẩn bằng kiểm định Shapiro-Wilk.

**4. KẾT LUẬN**

Kết quả nghiên cứu về 8 khoản mục chi tiêu chi tiết của cả ba cấp học phổ thông, các khoản chi tiêu về học phí chiếm tỉ lệ tương đối thấp trong tổng chi tiêu cho 1 học sinh trong 1 năm học. Kết quả này là

hợp lý cho thấy các chính sách hỗ trợ của Chính phủ và các nhà hoạch định chính sách góp phần làm giảm gánh nặng cho các hộ gia đình về tổng chi tiêu cho giáo dục; tuy nhiên, còn rất nhiều các khoản đóng góp khác và chi khác, đặc biệt là học thêm đối với cả ba cấp học. Học thêm có tỉ trọng chi tiêu lớn,



đặc biệt ở cấp THPT thể hiện sự đầu tư của các gia đình cho các con trong giai đoạn cuối cấp học phổ thông, nhưng đó cũng là gánh nặng và áp lực đối với một bộ phận gia đình. Kết quả nghiên cứu cho thấy giới tính của người học không ảnh hưởng đến cơ cấu chi tiêu ở tất cả các cấp học, thể hiện xu hướng bình đẳng giới đối với tiếp cận giáo dục của học sinh hiện nay và phù hợp với mục tiêu phát triển bền vững giáo dục của Chính phủ Việt Nam. Hiện nay, chính sách hỗ trợ giáo dục của Việt Nam hướng tới đối tượng khó khăn và đối tượng chính sách xã hội. Do đó, các đối tượng này được hưởng lợi từ chính sách hỗ trợ này có thể có tỉ trọng chi tiêu cho các khoản mục khác ít hơn, từ đó giảm gánh nặng chi tiêu giáo dục cho các hộ gia đình. Các đặc điểm hộ gia đình cũng có tác động đến cơ cấu chi tiêu và có ý nghĩa thống kê, tuy nhiên các tác động có xu hướng khác nhau. Thu nhập hộ gia đình tăng lên có tác động làm tăng chi tiêu cho các tỉ trọng đóng góp, học thêm và dụng cụ học tập và chi khác trong cơ cấu chi tiêu. Xu hướng này phù hợp với các nghiên cứu trước về ảnh hưởng của thu nhập hộ gia đình tới chi tiêu giáo dục (Trang, 2020; Ngoan và ctv., 2021). Đồng thời, xu hướng này phù hợp với chính sách xã hội hóa giáo dục, trong đó có tăng cường đầu tư cho giáo dục của hộ gia đình. Đặc điểm nhân khẩu học như dân tộc, trình độ học vấn và nghề nghiệp của chủ hộ có tác động ý nghĩa thống kê đến cơ cấu chi tiêu. Kết quả nghiên cứu này cũng đạt được ở các nghiên cứu khác khi sử dụng mô hình hồi quy Tobit (Tiền và ctv., 2014; Ngoan và ctv., 2021).

Một kết quả khá lưu ý là tác động của khu vực sinh sống và vùng kinh tế - xã hội đến cơ cấu chi tiêu giáo dục. Do chính sách về học phí trên toàn quốc, nghiên cứu chỉ ra không có sự khác biệt trong chi tiêu cho khoản mục thiết yếu giữa khu vực nông thôn và thành thị, tuy nhiên hộ gia đình ở nông thôn có tỉ trọng chi tiêu cho các khoản đóng góp; học thêm thấp hơn so với hộ gia đình tại khu vực thành thị. Bên cạnh đó, hộ gia đình sống tại khu vực đồng bằng sông Cửu Long có tỉ trọng chi tiêu cho các khoản thiết yếu bắt buộc thấp hơn các khu vực sinh thái khác. Tỉ trọng chi tiêu thiết yếu thấp trong tổng cơ cấu chi tiêu thể hiện tỉ trọng của các khoản mục khác (không thiết yếu) chiếm tỉ trọng cao. Cụ thể, đồng bằng sông Cửu Long vẫn còn nhiều khó khăn về giáo dục đào tạo, chi ngân sách địa phương cho giáo dục của đồng bằng sông Cửu Long thấp nhất cả nước, thấp hơn cả những vùng khó khăn như Tây Bắc, Tây Nguyên, nguồn ngân sách địa phương chi cho giáo dục thấp thì nguồn vốn Trung ương hỗ trợ các địa phương khu vực đồng bằng sông Cửu Long thông qua các chương trình, đề án nói chung chiếm

tỷ lệ thấp trong tổng số vốn Trung ương hỗ trợ các địa phương - thấp nhất so với các vùng trong cả nước, do chính sách hỗ trợ tập trung cho các địa bàn đặc biệt khó khăn, vùng núi, vùng đồng bào dân tộc thiểu số mà các tỉnh đồng bằng sông Cửu Long có ít đối tượng thụ hưởng so với các vùng khó khăn khác (Dung, 2018).

Loại hình trường học, dân lập/tư thục hoặc công lập, cũng ảnh hưởng đến cơ cấu chi tiêu giáo dục. Các trường dân lập và tư thục thường có mức thu học phí cao và đồng thời cũng bao gồm nhiều hỗ trợ và cơ sở vật chất tốt cho người học. Do đó, học sinh học trong trường dân lập/tư thục có tỉ trọng khoản chi tiêu thiết yếu bắt buộc cao hơn so với học sinh học trong các trường công lập.

Thông qua thống kê mô tả và mô hình hồi quy đa hợp, thực trạng và các yếu tố tác động đến cơ cấu chi tiêu giáo dục của học sinh theo hệ giáo dục phổ thông được đánh giá. Nền giáo dục của Việt Nam đang đạt được những thành tựu quan trọng như phổ cập giáo dục tiểu học, miễn học phí cho học sinh tiểu học, bình đẳng giới trong tiếp cận giáo dục và sự đa dạng hóa trong các loại hình trường học ở tất cả các địa phương. Bên cạnh đó, khoản chi tiêu về đóng góp, học thêm và chi khác chiếm tỉ trọng lớn trong cả ba cấp học được nghiên cứu. Đồng thời, đây cũng là khoản chi tiêu có thể thay đổi và đa dạng giữa các loại hình trường và khu vực. Điều kiện kinh tế xã hội như thu nhập bình quân hay vùng kinh tế, khu vực sống có tác động lớn đến cơ cấu chi tiêu của hộ gia đình đối với con em đang theo học ở các cấp học phổ thông. Xu hướng này thể hiện giáo dục được các hộ gia đình quan tâm đầu tư, tuy nhiên cũng có thể là gánh nặng đối với nhiều gia đình vì các chi phí giáo dục bên ngoài khoản chi tiêu thiết yếu. Do đó, các chính sách của Nhà nước về đầu tư giáo dục, xã hội hóa giáo dục và quan tâm các đối tượng yếu thế sẽ phát huy hiệu quả về đầu tư giáo dục. Nghiên cứu này còn một số hạn chế như chưa liên kết được cơ cấu chi tiêu giáo dục hộ gia đình đối với hiệu quả và chất lượng của đào tạo. Hơn nữa, hệ số  $R^2$  hiệu chỉnh của mô hình đối với hệ THPT là thấp hơn nhiều so với các cấp học còn lại nên các biến đưa vào mô hình chưa giải thích được nhiều về thực trạng cơ cấu chi tiêu của hộ. Về mặt phương pháp nghiên cứu, giải thích ý nghĩa của mô hình hồi quy trong không gian Euclide cũng có sự phức tạp do sự phụ thuộc lẫn nhau giữa các tỉ trọng, do đó, các hướng nghiên cứu tiếp sâu hơn để giải thích ý nghĩa mô hình hồi quy trong đơn hình, như hướng tiếp cận của Morais and Thomas-Agnan (2021) nên tiếp tục được sử dụng.

## TÀI LIỆU THAM KHẢO

- Aitchison, J. (1986). *The statistical analysis of compositional data*. Chapman and Hall, London. <https://doi.org/10.1007/978-94-009-4109-0>
- Aslam, M., & Kingdon, G. G. (2008). Gender and household education expenditure in Pakistan. *Applied Economics*, 40(20), 2573-2591. <https://doi.org/10.1080/00036840600970252>
- Egozcue, J. J., Pawlowsky-Glahn, V., & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35, 279-300. <https://doi.org/10.1023/A:1023818214614>
- Dung, Đ. T. T. (2018). Những cơ hội và thách thức trong việc giáo dục nhận thức cho học sinh vùng đồng bằng sông Cửu Long sẵn sàng thích ứng với biến đổi khí hậu. *Tạp chí Khoa học - Trường Đại học Sư phạm Tp. Hồ Chí Minh*, 15(4), 168.
- Dumuid, D., Stanford, T. E., Martin-Fernández, J. A., Pedišić, Ž., Maher, C. A., Lewis, L. K., Hron, K., Katzmarzyk, P. T., Chaput, J. P., Fogelholm, M., & Hu, G. (2018). Compositional data analysis for physical activity, sedentary time and sleep research. *Statistical methods in medical research*, 27(12), 3726-3738. <https://doi.org/10.1177/0962280217710835>
- Kuvat, Ö., & Kizilgöl, Ö.A., 2020. An analysis of out of pocket education expenditures in Turkey: logit and tobit models. *Ege Academic Review*, 20(3), 231-244.
- Le Van, T., Thu, T. D. T., & Thi, H. T. (2020). Các nhân tố ảnh hưởng cơ cấu chi tiêu của du khách quốc tế đến Việt Nam: Kết quả từ phương pháp phân tích số liệu đa hợp CoDA. *Tạp chí Khoa học Trường Đại học Cần Thơ*, 56(4), 208-218. <https://doi.org/10.22144/ctu.jvn.2020.100>
- Morais, J., & Thomas-Agnan, C. (2021). Impact of covariates in compositional models and simplicial derivatives. *Austrian Journal of Statistics*, 50(2), 1-15. <https://doi.org/10.17713/ajs.v50i2.1069>
- Ngoan, N. T., Mai, N. T. T., Trang, Đ. T. T., & Huong, T. T. (2021). Nghiên cứu các yếu tố ảnh hưởng đến việc chi tiêu cho giáo dục của hộ gia đình tại các tỉnh đồng bằng sông Hồng. *TNU Journal of Science and Technology*, 226(4), 53-61.
- Pawlowsky-Glahn, Vera, Juan José Egozcue, and Raimon Tolosana-Delgado. (2015). *Modeling and Analysis of Compositional Data*. John Wiley & Sons. <https://doi.org/10.1002/9781119003144>
- Tiến, D. K., & Lê, T. P. (2014). Các yếu tố Ảnh hưởng đến chi tiêu cho giáo dục của người dân Ở Đồng bằng sông Cửu Long. *Tạp chí Khoa học Trường Đại học Cần Thơ*, 31, 81-90.
- Trang, Đ. T. (2020). Yếu tố ảnh hưởng đến chi tiêu cho giáo dục tại 5 thành phố lớn. *Kỳ yếu Hội thảo Khoa học Quốc tế thường niên các nhà khoa học trẻ của các trường khối kinh tế và kinh doanh* (trang 1453-1463).
- Trinh, H. T., Morais, J., Thomas-Agnan, C., & Simioni, M. (2019). Relations between socio-economic factors and nutritional diet in Vietnam from 2004 to 2014: New insights using compositional data analysis. *Statistical Methods in Medical Research*, 28(8), 2305-2325. <https://doi.org/10.1177/0962280218770223>
- Tổng cục Thống kê. (2021). *Kết quả sơ bộ Khảo sát mức sống dân cư Việt Nam năm 2020*. <https://www.gso.gov.vn/du-lieu-va-so-lieu-thong-ke/2021/05/cong-bo-ket-qua-so-bo-ket-qua-khao-sat-muc-song-dan-cu-nam-2020/>.
- Van den Boogaart, K. G., & Tolosana-Delgado, R. (2013). *Analyzing compositional data with R (Vol. 122)*. Springer, Berlin. <https://doi.org/10.1007/978-3-642-36809-7>