



TRƯỜNG ĐẠI HỌC THƯƠNG MẠI



HỌC VIỆN CHÍNH SÁCH VÀ PHÁT TRIỂN

KỶ YẾU

HỘI THẢO KHOA HỌC QUỐC GIA

“PHÂN TÍCH ĐỊNH LƯỢNG CÁC VẤN ĐỀ
KINH TẾ VÀ XÃ HỘI TRONG MÔI TRƯỜNG SỐ”



NHÀ XUẤT BẢN THANH NIÊN

Tháng 4 năm 2022

MỤC LỤC

BÁO CÁO ĐỀ DẪN HỘI THẢO KHOA HỌC QUỐC GIA PHÂN TÍCH ĐỊNH LƯỢNG CÁC VẤN ĐỀ KINH TẾ VÀ XÃ HỘI TRONG MÔI TRƯỜNG SỐ	1
---	---

KHOA HỌC DỮ LIỆU TRONG PHÂN TÍCH CÁC VẤN ĐỀ KINH TẾ VÀ XÃ HỘI.....	5
--	---

Hồ Tú Bảo - Viện Nghiên cứu cao cấp về Toán

CHỦ ĐỀ 1: PHÂN TÍCH ĐỊNH LƯỢNG TRONG NGHIÊN CỨU KINH TẾ, XÃ HỘI	7
--	---

ĐÁNH GIÁ HIỆU QUẢ CÁC CHÍNH SÁCH ỨNG PHÓ CỦA CHÍNH PHỦ ĐỐI VỚI KIỂM SOÁT DỊCH BỆNH COVID-19 TẠI VIỆT NAM: TIẾP CẬN MÔ HÌNH ARDL.....	8
--	---

Nguyễn Thị Bích Loan, Vũ Thị Thu Hương - Trường Đại học Thương mại

CẤU TRÚC PHỤ THUỘC GIỮA THỊ TRƯỜNG DẦU VÀ MỘT SỐ THỊ TRƯỜNG KHÁC - TIẾP CẬN BẰNG PHƯƠNG PHÁP COPULA	27
---	----

Trần Trọng Nguyên - Học viện Chính sách và Phát triển, Nguyễn Thu Thủy - Trường Đại học Thương mại

TÁC ĐỘNG CỦA PHÂN CẤP TÀI KHÓA TỐI GIẢM NGHÈO: BẰNG CHỨNG THỰC NGHIỆM TẠI CÁC ĐỊA PHƯƠNG Ở VIỆT NAM	39
---	----

Trần Trọng Nguyên, Trần Hoàng Minh, Đặng Thùy Nhung - Học viện Chính sách và Phát triển

ẢNH HƯỞNG CỦA CƠ CẤU CHI TIÊU THEO NHÓM THỰC PHẨM ĐẾN LƯỢNG CALO TIÊU THỤ VÙNG NÔNG THÔN VIỆT NAM: KẾT QUẢ TỪ PHƯƠNG PHÁP PHÂN TÍCH ĐA HỢP	54
--	----

Phạm Ngọc Anh - Trường Đại học Mỏ - Địa chất, Đào Thị Thanh Bình - Trường Đại học Hà Nội, Nguyễn Minh Quang - Trường Đại học Sư phạm Hà Nội, Trịnh Thị Hương - Trường Đại học Thương mại

TÁC ĐỘNG CỦA ĐẠI DỊCH COVID-19 ĐẾN NGUỒN VỐN FDI VÀO MỘT SỐ NGÀNH KINH TẾ CHỦ LỰC CỦA VIỆT NAM	66
--	----

Nguyễn Thị Hiên, Hoàng Anh Tuấn, Vũ Thị Thùy Trang, Phạm Thị Huyền Trang, Lê Thị Lan Anh - Trường Đại học Thương mại

SỐ HOÁ VÀ TĂNG TRƯỞNG NĂNG SUẤT LAO ĐỘNG Ở VIỆT NAM.....	77
--	----

Nguyễn Thị Cẩm Vân - Trường Đại học Kinh tế quốc dân

TÁC ĐỘNG CỦA HỘI NHẬP KINH TẾ QUỐC TẾ ĐẾN MỐI QUAN HỆ GIỮA THỊ TRƯỜNG CHỨNG KHOÁN MỸ VÀ TRUNG QUỐC ĐẾN THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM	90
--	----

Nguyễn Thu Thủy, Nguyễn Thị Quỳnh Trâm - Trường Đại học Thương mại

TÁC ĐỘNG CỦA ĐA DẠNG SINH KẾ ĐẾN THU NHẬP HỘ GIA ĐÌNH KHU VỰC NÔNG THÔN VIỆT NAM.....	105
---	-----

ẢNH HƯỞNG CỦA CƠ CẤU CHI TIÊU THEO NHÓM THỰC PHẨM ĐẾN LƯỢNG CALO TIÊU THỤ VÙNG NÔNG THÔN VIỆT NAM: KẾT QUẢ TỪ PHƯƠNG PHÁP PHÂN TÍCH ĐA HỢP

Phạm Ngọc Anh - Trường Đại học Mở - Địa chất, Đào Thị Thanh Bình - Trường Đại học Hà Nội, Nguyễn Minh Quang - Trường Đại học Sư phạm Hà Nội, Trịnh Thị Hường - Trường Đại học Thương mại
Email: phamngocanhbmtoan@humg.edu.vn

Tóm tắt: Bài báo này đóng góp vào việc phân tích ảnh hưởng của cơ cấu chi tiêu theo nhóm thực phẩm đến năng lượng tiêu thụ tại Việt Nam, sử dụng bộ dữ liệu hạn chế trên khu vực nông thôn từ cuộc khảo sát Điều tra mức sống Dân cư Việt Nam năm 2020. Chúng tôi sử dụng phân tích đa hợp (Compositional Data Analysis, CODA) cho các nhóm chi tiêu và mô hình hồi quy đa hợp giữa các nhóm chi tiêu và lượng năng lượng tiêu thụ bình quân đầu người. Khác với các nghiên cứu trước đây, chúng tôi tập trung đánh giá tác động của tổng chi tiêu thực phẩm và tỉ lệ chi tiêu của 5 nhóm thực phẩm khác nhau (thực phẩm ăn ngoài, thực phẩm chế biến cao, thực phẩm chế biến ít, thực phẩm siêu chế biến, và thực phẩm không chế biến). Chúng tôi sử dụng phương pháp CODA để mô tả các mô hình cơ cấu chi tiêu của 5 nhóm thực phẩm đến năng lượng tiêu thụ bình quân. Chúng tôi tính toán độ co giãn chi tiêu cho thực phẩm với lượng tiêu thụ năng lượng theo tổng chi tiêu hoặc chi tiêu của từng khoản mục. Mô hình phân tích đa hợp làm nổi bật vai trò quan trọng của nhiều yếu tố trong việc xác định lựa chọn chế độ ăn uống. Kết quả nghiên cứu cung cấp thêm bằng chứng khoa học về xu hướng tiêu dùng theo cơ cấu các nhóm thực phẩm và tổng chi tiêu cho các nhà hoạch định chính sách và các nhà kinh tế.

Từ khóa: Compositional Data Analysis CoDa (phân tích đa hợp), Bộ số liệu điều tra mức sống dân cư, tổng chi tiêu thực phẩm, 5 nhóm thực phẩm.

1. Giới thiệu

An ninh lương thực và chất dinh dưỡng ngày càng trở thành mối quan tâm chính của chính phủ và các tổ chức phi lợi nhuận do các ảnh hưởng đối với sức khỏe và tình hình phát triển kinh tế. Chủ đề này đã được liệt kê là một trong 17 mục tiêu phát triển bền vững của thế giới do Liên hợp Quốc đề xuất (United Nations, 2018). Nhiều nghiên cứu thực nghiệm tập trung vào mối quan hệ giữa các đặc điểm kinh tế xã hội của các hộ gia đình và hành vi tiêu dùng thực phẩm của họ (Deaton, 1997). Mức tiêu thụ thực phẩm được đo bằng năng lượng tiêu thụ bình quân đầu người (Per capita calo intake, PCCI), đơn vị: Calo. Một nghiên cứu của Ogundari và Abdulai (2013) cho thấy mối quan hệ giữa lượng calo và thu nhập (hoặc chi tiêu) đã được nghiên cứu kỹ lưỡng ở nhiều quốc gia nhằm thực hiện các chính sách giảm thiểu tình trạng đói và suy giảm dinh dưỡng. Sau đó, sự phát triển kinh tế và đô thị hóa ở các

nước đang phát triển đã ảnh hưởng đến chế độ ăn uống toàn cầu, dẫn đến nhiều nghiên cứu thực nghiệm tập trung vào các nguồn thực phẩm, chẳng hạn như rau, ngũ cốc, thịt, ... Việt Nam là một ví dụ điển hình về một quốc gia có thu nhập trung bình đã đạt được những thành tựu ấn tượng về kinh tế và phúc lợi dân số sau cải cách kinh tế năm 1986. Tuy nhiên, quốc gia này cũng đã trải qua quá trình chuyển đổi dinh dưỡng (nutrition transition) như nhiều quốc gia có thu nhập trung bình khác, trong đó, các vấn đề về suy dinh dưỡng và béo phì diễn ra đồng thời do xu hướng thay đổi khẩu phần ăn và nhiều nguyên nhân khác của quá trình đô thị hóa.

Mục đích của nghiên cứu này là phân tích sự ảnh hưởng của chi tiêu thực phẩm, bao gồm theo 5 loại thực phẩm và cả tổng chi tiêu, đến lượng tiêu thụ PCCI. Theo những gì chúng tôi biết, nghiên cứu của chúng tôi là nghiên cứu đầu tiên sử dụng các công cụ CODA để phân tích ảnh hưởng của cơ cấu chi tiêu theo nhóm thực phẩm đến lượng calo tiêu thụ tại Việt Nam. Ứng dụng của phương pháp CODA đã được nghiên cứu vào nhiều lĩnh vực khác nhau, bao gồm cả nghiên cứu về dinh dưỡng (Van den Boogaart và cộng sự, 2013; Pawlowsky-Glahn và cộng sự, 2015). Thu nhập tăng dẫn đến lượng lương thực hay lượng calo tiêu thụ bình quân (PCCI) cũng tăng theo. Khác với hướng tiếp cận gần đây là mối quan hệ phi tuyến của chi tiêu và lượng calo tiêu thụ bình quân (Trịnh và cộng sự, 2018a). Chúng tôi đưa thêm yếu tố về chi tiêu theo 5 khoản mục vào mô hình đánh giá, từ đó chúng tôi đo lường độ co giãn của tỷ lệ chất dinh dưỡng so với chi tiêu cho thực phẩm và chúng tôi cũng so sánh độ co giãn này với độ co giãn của tổng chi tiêu lên tổng lượng tiêu thụ bằng cách sử dụng các mô hình tuyến tính cổ điển. Nghiên cứu này sử dụng dữ liệu của cuộc khảo sát Điều tra Mức sống Hộ gia đình Việt Nam (MSDC) năm 2020. Nghiên cứu này được chia làm bốn phần chính: Phần một giới thiệu chủ đề nghiên cứu, phần hai thảo luận về phương pháp nghiên cứu, Phần ba phân tích kết quả nghiên cứu, Phần bốn kết luận và đưa ra một số hướng nghiên cứu tiếp theo.

2. Phương pháp nghiên cứu

2.1. Dữ liệu nghiên cứu

Bộ dữ liệu sử dụng trong nghiên cứu được lấy từ cuộc khảo sát Điều tra mức sống Dân cư Việt Nam mới nhất năm 2020, thực hiện bởi Tổng cục Thống kê Việt Nam. Đây là các cuộc khảo sát được tiến hành 02 năm một lần trên phạm vi cả nước bằng phương pháp phỏng vấn trực tiếp chủ hộ. Kết quả điều tra phục vụ cho các đánh giá về phát triển kinh tế xã hội của đất nước (Tổng cục thống kê, 2019). Nghiên cứu này sử dụng bộ dữ liệu hơn 9000 hộ gia đình và các kết quả đại diện cấp hộ, chúng tôi hạn chế trên khu vực nông thôn. 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 khu vực sinh thái.

Mục chỉ tiêu: Mục này thu thập thông tin về nhiều loại chỉ tiêu, trong đó có chỉ tiêu thực phẩm thường xuyên trong 30 ngày từ gần 60 loại thực phẩm. Từ đó, chúng tôi tính được tổng chỉ tiêu thực phẩm của hộ gia đình. Đồng thời, các thực phẩm tiêu thụ được chia thành 5 nhóm theo mức độ chế biến là: Thực phẩm không chế biến (Unprocessed food, UN) bao gồm các loại rau và tinh bột chưa chế biến, thực phẩm chế biến ít (Low processed, LP) bao gồm các tinh bột đã sơ chế (như bún gạo khô) loại thịt cá tươi trứng sữa; thực phẩm chế biến cao (High processed, HU) gồm một số loại như sữa đặc; thực phẩm siêu chế biến (Ultra-processed, UL) gồm các dầu ăn, nước ngọt, ... và ăn ngoài (Food away from home, FAFH). Bên cạnh đó, lượng năng lượng (calo) tiêu thụ bình quân PCCI cũng được tính từ thực phẩm tiêu thụ ở mục này. Chúng tôi sử dụng phương pháp tính PCCI có tính điều chỉnh lượng tiêu thụ năng lượng của các thành viên hộ gia đình theo giới tính và tuổi (Deaton, 1997; Trịnh và cộng sự, 2018a). Hơn nữa, PCCI là lượng thực phẩm sẵn có, một trong 4 chỉ tiêu của an ninh lương thực (availability, access, utilization and stability) nên sẽ cao hơn mức thực tế tiêu thụ theo các khuyến cáo của Bộ Y tế.

2.2 Phương pháp phân tích đa hợp

Chúng tôi nhắc một số khái niệm về phân tích số liệu đa hợp (Compositional Data Analysis – CoDa), dựa trên các nghiên cứu của Van den Boogaart và Tolosana-Delgado (2013), Pawlowsky-Glahn và cộng sự (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 và cộng sự, 2015). Phương pháp này gần đây đã được áp dụng trong y tế và nghiên cứu dịch tễ học dinh dưỡng (Dumuid và cộng sự, 2018).

2.2.1. Một số định nghĩa

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

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

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

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

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

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), cụ thể là tỉ lệ bất biến (scale invariance); hoán vị bất biến (permutation invariance) và vectơ con nhất quán (subcompositional coherence).

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 Euclid, 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 vectơ gồm $D-1$ tọa độ trong không gian Euclide thông thường (Egozcue và cộng sự, 2003). ILR được thực hiện dựa trên sự phân nhóm không giao giữa các khoản mục trong vectơ 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:

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

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 quan tâm mô hình có vectơ đa hợp đóng vai trò biến độc lập và tổng dưới dạng tổng đại số.

2.2.2. Mô hình hồi quy đa hợp trong nghiên cứu lượng calo tiêu thụ bình quân

Mô hình này là mở rộng của mô hình nghiên cứu ảnh hưởng của thu nhập đến lượng calo tiêu thụ bình quân, được khởi xướng từ các nghiên cứu của Deaton (1997). Mô hình đề xuất bao gồm mô hình hồi quy bội phi tuyến (log log) và có thể có tính đến tác động phi tuyến qua hàm lôgarít bình phương hoặc các mô hình phức tạp hơn (Trịnh và cộng sự, 2018a). Chúng tôi mở rộng mô hình và thêm yếu tố về cơ cấu chi tiêu theo các nhóm thực phẩm trên nền tảng biến tổng chi tiêu như các nghiên cứu truyền thống.

Kí hiệu V_1, V_2, V_3, V_4, V_5 là độ lớn các khoản của từng khoản mục của chi tiêu thực phẩm, tổng chi tiêu là $T = V_1 + V_2 + V_3 + V_4 + V_5$ và $S = (s_1, s_2, s_3, s_4, s_5) \in$

$S^5, s_i = \frac{V_i}{T}$ là đơn hình của không gian các véctor đa hợp. Sử dụng phép biến đổi tỉ lệ lôgarit đẳng cự ILR như sau:

$$\begin{aligned} Ilr_1 &= \frac{2}{\sqrt{5}} \ln \ln \frac{s_1}{\sqrt{s_2 s_3 s_4 s_5}}; Ilr_2 = \ln \ln \frac{\sqrt{s_2 s_3}}{\sqrt{s_4 s_5}}; Ilr_3 = \frac{1}{\sqrt{2}} \ln \ln \frac{s_2}{s_3}; Ilr_4 \\ &= \frac{1}{\sqrt{2}} \ln \ln \frac{s_4}{s_5}. \end{aligned}$$

Mô hình hồi quy đa hợp trong đơn hình có dạng

$$\ln \ln (PCCI_i) = a + \langle b, S_i \rangle_A + \gamma T_i + \sum_k \beta_k Z_{ik} + \epsilon_i \quad (3)$$

với PCCI là lượng calo tiêu thụ bình quân (đơn vị: Calo), S là vector cơ cấu chi tiêu (%), T là tổng chi tiêu thực phẩm (đơn vị: nghìn đồng), Z_k là véctor về đặc điểm nhân khẩu học và nơi sống, i là chỉ số hộ gia đình và ϵ là sai số của mô hình. Chỉ số i có thể được bỏ qua để đơn giản kí hiệu.

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

$$\ln \ln (PCCI_i) = a + \sum_{j=1}^4 \alpha_j ILR_{ji} + \gamma \ln (T_i) + \sum_k \beta_k Z_{ik} + \epsilon_i \quad (4)$$

Thông qua phương pháp bình phương nhỏ nhất, các hệ số của mô hình (4) được ước lượng. Hệ số ước lượng của các biến Z_k được giải thích như mô hình hồi quy bội, tức là xem xét sự thay đổi của 1 yếu tố và giữ cố định các yếu tố còn lại. Đối với ảnh hưởng của chi tiêu và cơ cấu chi tiêu, chúng tôi quan tâm hai trường hợp sau

- **Trường hợp 1:** Tổng T tăng 1 đơn vị và mối quan hệ tương đối giữa các tỉ trọng không thay đổi. Khi đó, chúng mình được rằng mỗi khoản mục $V_j, j = 1, \dots, 5$ đều tăng 0.2 đơn vị và các ILR không thay đổi. Khi đó, hệ số $\hat{\gamma}$ được giải thích như mô hình tuyến tính bội, khi đó, $\hat{\gamma}$ là hệ số co giãn giữa của chi tiêu thực phẩm và lượng calo tiêu thụ bình quân.
- **Trường hợp 2:** Chi tiêu một khoản mục V_1 tăng δ %, và các khoản mục khác không đổi. Do đó, sự thay đổi này kéo theo tổng chi tiêu T cũng thay đổi lượng tương ứng như ở trên. Kí hiệu V'_1 và T' là khoản mục thứ nhất và tổng chi tiêu sau khi thay đổi (tương tự cho các khoản mục khác), ta có

$$V'_1 = (1 + \delta)V_1 \text{ và } T' = T + \delta V_1.$$

Khi đó, ILR_1 sẽ thay đổi và ILR_2, ILR_3, ILR_4 không thay đổi. Hơn nữa, ILR_1 có thể biểu diễn qua tỉ trọng của các khoản mục chi tiêu, nên

$$\begin{aligned} ILR'_1 &= \frac{2}{\sqrt{5}} \ln \ln \frac{V'_1}{\sqrt{V'_2 V'_3 V'_4 V'_5}} = \frac{2}{\sqrt{5}} \ln \ln \frac{(1+\delta)V_1}{\sqrt{V_2 V_3 V_4 V_5}} \\ &= ILR_1 + \frac{2}{\sqrt{5}} \ln \ln (1+\delta) \end{aligned}$$

Gọi $PCCI'$ là mức tiêu thụ năng lượng bình quân sau khi thay đổi, khi đó ta có

$$\begin{aligned} \log(\widehat{PCCI}) &= \hat{a} + \sum_{j=1}^4 \hat{a}_j ILR_j + \hat{\gamma} \ln \ln (T) + \sum_k \hat{\beta}_k Z_k \\ \log(\widehat{PCCI}') &= \hat{a} + \sum_{j=1}^4 \hat{a}_j ILR'_j + \hat{\gamma} \ln \ln (T') + \sum_k \hat{\beta}_k Z_k = \\ &= \hat{a} + \sum_{j=1}^4 \hat{a}_j ILR_j + \hat{\alpha}_1 \frac{2}{\sqrt{5}} \ln \ln (1+\delta) + \hat{\gamma} \ln \ln (T + \delta V_1) + \sum_k \hat{\beta}_k Z_k \end{aligned}$$

Sự thay đổi của $\log(\widehat{PCCI})$ là

$$\begin{aligned} \log(\widehat{PCCI}') - \log(\widehat{PCCI}) &= \hat{\alpha}_1 \frac{2}{\sqrt{5}} \ln \ln (1+\delta) + \hat{\gamma} \ln \ln (T + \delta V_1) - \hat{\gamma} \ln \ln (T) \end{aligned}$$

Khi đó,

$$\begin{aligned} \frac{PCCI'}{PCCI} &= e^{\hat{\alpha}_1 \frac{2}{\sqrt{5}} \ln \ln (1+\delta) + \hat{\gamma} \ln \ln (T + \delta V_1) - \hat{\gamma} \ln \ln (T)} \\ \frac{PCCI'}{PCCI} - 1 &= e^{\hat{\alpha}_1 \frac{2}{\sqrt{5}} \ln \ln (1+\delta) + \hat{\gamma} \ln \ln (T + \delta V_1) - \hat{\gamma} \ln \ln (T)} - 1. \end{aligned}$$

Nếu δ nhỏ, ta có

$$\begin{aligned} &e^{\hat{\alpha}_1 \frac{2}{\sqrt{5}} \ln \ln (1+\delta) + \hat{\gamma} \ln \ln (T + \delta V_1) - \hat{\gamma} \ln \ln (T)} - 1 \\ &\approx \hat{\alpha}_1 \frac{2}{\sqrt{5}} \ln \ln (1+\delta) + \hat{\gamma} \ln \ln (T + \delta V_1) - \hat{\gamma} \ln \ln (T) \\ &\approx \hat{\alpha}_1 \frac{2}{\sqrt{5}} \delta + \hat{\gamma} \ln \ln \frac{T + \delta V_1}{T} \approx \hat{\alpha}_1 \frac{2}{\sqrt{5}} \delta + \hat{\gamma} \delta \frac{V_1}{T}. \end{aligned}$$

Do đó, tỉ lệ thay đổi (%) của PCCI phụ thuộc vào $\frac{V_1}{T}$ và tỉ lệ là, với δ nhỏ:

$$\delta \left(\hat{\alpha}_1 \frac{2}{\sqrt{5}} + \hat{\gamma} \frac{V_1}{T} \right). \quad (5)$$

Công thức (5) cho thấy sự thay đổi của PCCI trong trường hợp 2 này phụ thuộc vào từng quan sát i .

Trong trường hợp quan tâm đến sự thay đổi của 1 khoản khác, chúng ta có thể thay thế vai trò của V_1 bởi các khoản mục khác và đạt được kết quả tương tự do vai trò của các khoản mục là tương đương nhau trong ILR và tổng T.

Các trường hợp thay đổi khác giữa tổng và các khoản mục, người đọc có thể tham khảo nghiên cứu của Coenders và cộng sự (2017), Morais và Thomas-Agnan (2021), Pawlowsky-Glahn và cộng sự (2015). Mô hình (4) có thể được thêm bình phương lôgarit của tổng T, tuy nhiên kết quả tính tỉ lệ thay đổi của PCCI sẽ không đạt được công thức tường minh như (5), nhưng có thể tính toán mô phỏng qua phần mềm thống kê.

Các tính toán trong nghiên cứu này được tiến hành trên phần mềm R, phiên bản 4.0.2 và các gói lệnh *tidyverse*, *tableone*, *compositions* và *zCompositions*.

3. Kết quả nghiên cứu

3.1. Đặc điểm đối tượng quan sát

Thông tin nhân khẩu học và đặc điểm kinh tế xã hội của đối tượng khảo sát theo vùng sinh thái được thể hiện ở Bảng 1. Nghiên cứu hạn chế trên vùng nông thôn nên số quan tại 4 vùng sinh thái tương đối đồng đều, trừ khu vực Tây Nguyên và Đông Nam Bộ. Chi tiêu thực phẩm hàng tháng của hộ gia đình nhiều nhất là ở vùng Đông Nam Bộ và ít nhất là khu vực Trung du và miền núi phía Bắc. Từ đó, khu vực càng phát triển thì mức chi tiêu cho thực phẩm càng nhiều. Hơn nữa, vùng Đông Nam Bộ cũng có tỉ lệ chi tiêu ăn uống ngoài gia đình (FAFH) lớn nhất. Khoản mục thực phẩm chế biến ít (LP) và thực phẩm không chế biến (UN) chiếm tỉ trọng cao nhất đối với hộ gia đình ở cả 4 vùng sinh thái. Hình 1 thể hiện tứ phân vị của các khoản mục và theo vùng miền. Các khoản mục FAFH, LP and UN cũng có độ giao động nhiều nhất, đặc biệt ở Đông Nam Bộ. Nghiên cứu hạn chế trên khu vực nông thôn, nên trình độ giáo dục phổ biến nhất là Không bằng cấp hoặc tiểu học (ngoại trừ khu vực Đồng Bằng sông Hồng). Chủ hộ là nam giới chiếm đa số, tuổi chủ hộ khoảng 50 tuổi, gia đình có 4 thành viên. Chủ hộ làm công ăn lương chiếm đa số, tiếp theo là tỉ lệ làm việc trong lĩnh vực nông nghiệp. Tính trên tỉ lệ các thành viên trong độ tuổi lao động, tỉ lệ phụ nữ đi làm chiếm từ 40-45%.

Bảng 1: Đặc điểm đối tượng quan sát

Biến quan sát	Vùng sinh thái					
Vùng sinh thái	Đồng bằng sông Hồng	Trung du và miền núi phía Bắc	Bắc Trung Bộ và duyên hải miền Trung	Tây Nguyên	Đông Nam Bộ	Đồng bằng sông Cửu Long
Số quan sát	1201	1128	1127	380	358	1129

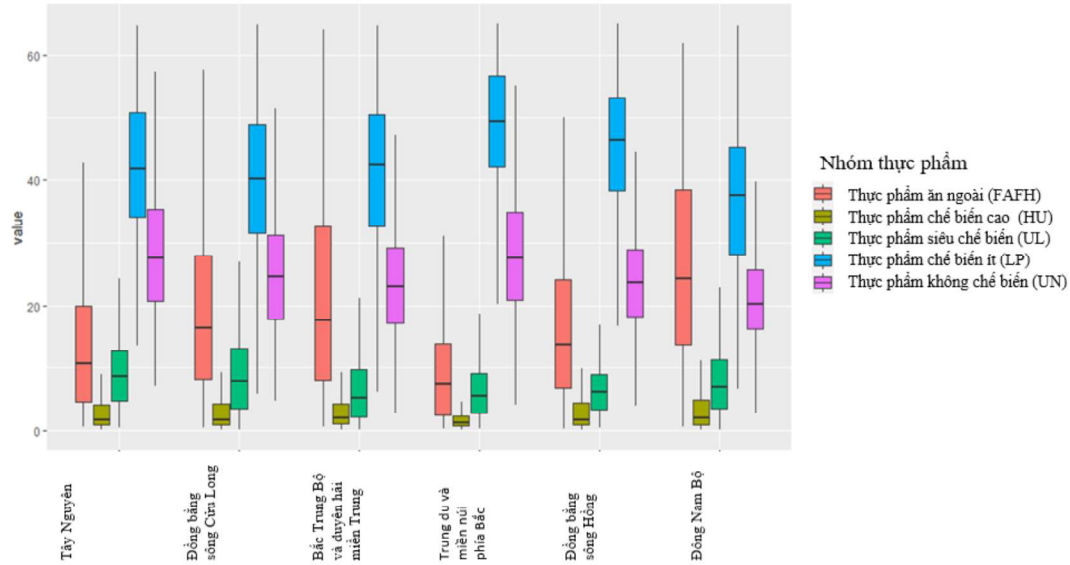
PHÂN TÍCH ĐỊNH LƯỢNG CÁC VẤN ĐỀ KINH TẾ VÀ XÃ HỘI TRONG MÔI TRƯỜNG SỐ

Chi tiêu thực phẩm hộ gia đình (tháng/người đồng)		3819.28 (2017.53)	3063.99 (1563.52)	3775.77 (2226.95)	3391.72 (2012.24)	4569.27 (2355.11)	3624.69 (1931.06)
Lượng calo tiêu thụ bình quân (PCCI)		3265.19 (858.28)	3395.09 (873.47)	3507.82 (904.37)	3567.50 (876.53)	3592.55 (891.21)	3661.59 (888.30)
Cơ cấu chi tiêu khoản mục (%)	Thực phẩm ăn ngoài (FAFH)	17.43 (13.79)	10.32 (10.71)	22.70 (18.00)	15.62 (15.35)	28.16 (17.82)	20.09 (15.59)
	Thực phẩm chế biến cao (HU)	4.81 (7.43)	3.05 (5.64)	4.35 (6.26)	3.83 (5.45)	4.72 (6.94)	4.57 (6.72)
	Thực phẩm chế biến ít (LP)	46.90 (11.90)	51.41 (12.40)	42.31 (13.56)	42.89 (12.88)	37.54 (12.54)	40.77 (12.32)
	Thực phẩm siêu chế biến (UL)	6.66 (4.48)	6.37 (4.67)	6.83 (6.15)	9.27 (6.54)	7.95 (5.67)	9.27 (7.28)
	Thực phẩm không chế biến (UN)	24.19 (8.10)	28.84 (10.99)	23.81 (9.23)	28.39 (10.47)	21.64 (8.57)	25.31 (9.79)
Giới tính chủ hộ	Nam	81.4	83.3	79.3	79.5	73.5	73.9
	Nữ	18.6	16.7	20.7	20.5	26.5	26.1
Tuổi chủ hộ	Năm	52.20 (14.55)	47.19 (12.79)	51.99 (14.11)	48.11 (13.75)	50.70 (13.64)	52.46 (13.15)
Số thành viên hộ gia đình (%)	Một hoặc 2 thành viên	30.2	19.2	25.2	20.5	24.9	23.1
	3 thành viên	15.1	19.4	19.5	15	24	22.6
	4 thành viên	26.6	27.9	27.8	30	28.5	24.5
	5 thành viên	16.2	18.5	16.1	17.6	13.1	16
	6 thành viên trở lên	11.9	14.9	11.4	16.8	9.5	13.7
Chỉ số phụ thuộc hộ gia đình		20.74 (21.55)	24.96 (21.19)	20.55 (21.62)	23.62 (21.15)	18.20 (20.07)	19.95 (20.11)
Dân tộc (%)	Kinh	97 (16)	35 (48)	88 (32)	64 (48)	89 (31)	92 (27)
Bằng cấp giáo dục cao nhất của chủ hộ	Không bằng cấp hoặc tiểu học	23.2	49.4	46.2	55.8	61.7	74.3
	Trung học cơ sở và trung học phổ thông	71.4	47	48.5	42.4	35.8	22.5
	Đại học	5.3	3.6	5.2	1.8	2.5	3.2
Tỉ lệ phụ nữ đi làm trong gia đình (%)		44.64 (28.69)	46.73 (24.21)	42.95 (28.64)	44.65 (24.07)	41.11 (28.70)	39.94 (28.43)
Chủ hộ là người làm công ăn lương (%)		68 (47)	63 (48)	68 (47)	57 (50)	70 (46)	66 (47)
Nghề nghiệp chủ hộ	Làm công ăn lương	68	63.3	68.3	57.4	69.6	66.4
	Nông nghiệp	18.4	33.5	23.9	35.8	14.8	22.9
	Dịch vụ	13.6	3.2	7.8	6.8	15.6	10.6
Tình trạng hôn nhân (%)	Kết hôn	84 (37)	84 (36)	82 (38)	81 (39)	80 (40)	80 (40)

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

Nguồn: Tác giả tính toán từ bộ số liệu MSDC 2020.

Hình 1: Biểu đồ hộp về cơ cấu chi tiêu thực phẩm theo nhóm ở các vùng miền.



3.2. Kết quả mô hình hồi quy

Bảng 2 thể hiện hệ số hồi quy của mô hình, kết quả kiểm định tính phân phối chuẩn của sai số ngẫu nhiên là phù hợp ($p\text{-value} < 0.1$). Các biến đổi tỉ lệ lôgarit đẳng cự ILR đều có ý nghĩa thống kê, trừ ILR4 về tỉ lệ giữa FAFH và UL. Lôgarit của chi tiêu thực phẩm có tác động dương và có ý nghĩa thống kê. Do đó, mỗi quan hệ giữa các tỉ trọng không thay đổi và các yếu tố khác cố định, tăng 1% cho chi tiêu thực phẩm thì lượng calo tiêu thụ bình quân tăng thêm 0.296%. Phân tích tác động của các biến đổi tỉ lệ lôgarit đẳng cự ILR cần tính đến sự thay đổi của tổng, tức lôgarit của chi tiêu thực phẩm. Ở đây, tỉ trọng thực phẩm không chế biến (Unprocessed food, UN) đóng vai trò như V_1 . Do hệ số ước lượng ILR1 và lôgarit của chi tiêu thực phẩm đều dương, nên khi mức chi khoản mục UN tăng δ (%) dẫn đến tỉ lệ tăng calo bình quân dương. Bảng 3 minh hoạt một số gia đình có cùng mức chi tiêu thực phẩm (3440 ngàn đồng/tháng) nhưng có mức chi khoản mục thực phẩm không chế biến khác nhau và trong trường hợp các gia đình cùng tăng 1% tỉ lệ chi tiêu cho khoản mục UN và giữ cố định chi tiêu các khoản mục khác. Khi đó, hộ 3 có mức chi tiêu cho khoản mục UN nhỏ nhất nên mức tăng calo bình quân là 4.27%, trong khi hộ 1 và hộ 2 có mức tăng calo bình quân cao hơn, tương ứng 7.16% và 9.78%. Hộ 2 cũng có tỉ lệ chi tiêu cho nhóm UN lớn nhất.

So với gia đình có 1 hoặc hai thành viên, các gia đình có nhiều thành viên có mức tiêu thụ calo ít hơn và có ý nghĩa thống kê. Tương tự, so với các gia đình có chủ hộ là người dân tộc thiểu số, chủ hộ là người Kinh cũng có mức tiêu thụ calo ít hơn. Chủ hộ có trình độ giáo dục càng cao cũng có mức tiêu thụ calo ít hơn. Các kết quả này đạt được giống như các nghiên cứu trước đây (Trịnh và cộng sự, 2018b). Calo giảm thể hiện xu đa dạng hóa thực phẩm và phù hợp với sự phát triển của thị hiếu

người tiêu dùng, khi người tiêu dùng lựa chọn nhiều loại thực phẩm khác nhau và không có xu hướng tăng calo như các loại rau. Bên cạnh đó, tỉ lệ người phụ thuộc trong hộ gia đình cũng có tác động dương và có ý nghĩa thống kê. So với chủ hộ là người làm công ăn lương, chủ hộ làm trong lĩnh vực dịch vụ có mức tiêu thụ năng lượng bình quân ít hơn và có ý nghĩa thống kê. Các khu vực sinh thái đều có mức tiêu thụ calo bình quân ít hơn so với khu vực Tây Nguyên và có ý nghĩa thống kê.

Bảng 2: Hệ số của mô hình hồi quy

Hệ số hồi quy		Hệ số hồi quy	Độ lệch chuẩn
Hệ số chặn		5.969 ***	-0.07
Biến đổi tỉ lệ lôgarit đăng cự ILR	Ilr1 (UN: (FAFH, UL, LP, HU))	0.12 ***	-0.009
	Ilr2 ((LP, HU): (FAFH, UL))	-0.067 ***	-0.005
	Ilr3 (LP:HU)	-0.058 ***	-0.006
	Ilr4 (FAFH:UL)	-0.004	-0.003
Lôgarit của chỉ tiêu thực phẩm		0.296 ***	-0.009
Giới tính (mặc định: Nữ)	Nam	0.018.	-0.011
Số thành viên hộ (mặc định: nhỏ hơn bằng 2)	3 thành viên	-0.08 ***	-0.011
	4 thành viên	-0.137 ***	-0.013
	5 thành viên	-0.212 ***	-0.014
	6 thành viên trở lên	-0.278 ***	-0.016
Dân tộc (Mặc định: Thiểu số)	Kinh	-0.044 ***	-0.01
Bằng cấp giáo dục cao nhất của chủ hộ (mặc định: Không bằng cấp hoặc tiểu học)	Trung học cơ sở và trung học phổ thông	-0.026 ***	-0.007
	Đại học	-0.055 **	-0.017
Tỉ lệ phụ nữ đi làm trong gia đình (%)		0.	0
Tình trạng hôn nhân (mặc định: Khác)	Kết hôn	-0.041 ***	-0.012
Chỉ số phụ thuộc hộ gia đình		0.001 ***	0
Nghề nghiệp chủ hộ (mặc định: Làm công ăn lương)	Nông nghiệp	0.006	-0.008
	Dịch vụ	-0.032 **	-0.012
Vùng sinh thái (mặc định: Tây Nguyên)	Bắc Trung Bộ và duyên hải miền Trung	-0.028 *	-0.014
	Đồng bằng sông Hồng	-0.091 ***	-0.014
	Trung du và miền núi phía Bắc	-0.033 *	-0.014
	Đông Nam Bộ	-0.08 ***	-0.018
	Đồng bằng sông Cửu Long	0.007	-0.014
R bình phương hiệu chỉnh		0.252	

Bảng 3: Minh họa tỉ lệ thay đổi mức tiêu thụ calo đối với chi tiêu khoản mục UN

Hộ gia đình	Chi tiêu thực phẩm hộ gia đình (tháng/ngàn đồng)	Chi tiêu thực phẩm khoản mục UN (tháng/ngàn đồng)	Tỉ trọng thực phẩm không chế biến (UN)	Tỉ lệ thay đổi PCCI (%)
Hộ 1	3440	82000	23.84	7.16
Hộ 2	3440	100900	29.33	8.78
Hộ 3	3440	48400	14.07	4.27

4. Kết luận và một số mở rộng

Nghiên cứu tập trung phân tích ảnh hưởng của cơ cấu chi tiêu theo nhóm thực phẩm đến lượng calo tiêu thụ bình quân tại Việt Nam tại khu vực nông thôn Việt Nam. Chúng tôi sử dụng phương pháp phân tích đa hợp CODA, xem xét tác động đồng thời của chi tiêu thực phẩm hộ gia đình và chi tiêu theo các nhóm thực phẩm. Hạn chế trên khu vực nông thôn và bộ số liệu điều tra năm 2020, chúng tôi tìm được các kết quả như các nghiên cứu trước đây như xu hướng giảm calo tiêu thụ bình quân khi tăng số thành viên hộ, trình độ giáo dục cao hơn. Đối với tác động của chi tiêu thực phẩm, tăng chi tiêu thực phẩm sẽ dẫn đến tăng calo tiêu thụ bình quân của hộ gia đình. Cơ cấu chi tiêu thực phẩm cũng có tác động có ý nghĩa thống kê đến lượng calo tiêu thụ. Chúng tôi minh họa tác động khi tăng chi tiêu của khoản mục thực phẩm không chế biến – bao gồm chủ yếu chi tiêu các loại ngũ cốc và rau xanh. Trong trường hợp này, calo bình quân của hộ gia đình sẽ tăng. Kết quả nghiên cứu cung cấp thêm bằng chứng cho mục tiêu giảm nghèo đói vì sự phát triển bền vững thiên niên kỷ mà chính phủ Việt Nam cam kết thực hiện. Hơn nữa, kết quả nghiên cứu cũng cho thấy xu hướng giảm calo tiêu thụ vì hướng tới đa dạng hóa thực phẩm, vì một mức sống thịnh vượng, phù hợp thị hiếu người tiêu dùng.

Phương pháp phân tích đa hợp cung cấp chi tiêu ảnh hưởng của các thành phần trong một tổng thể. Tuy nhiên, quá trình phân tích và giải thích tác động không trực quan như trong mô hình hồi quy bội. Các nghiên cứu tiếp theo sử dụng phương pháp phân tích đa hợp và ứng dụng trong các lĩnh vực khác nhau nên được khuyến khích để tăng sự phát triển cả về lý thuyết và ứng dụng.

TÀI LIỆU THAM KHẢO

- [1]. Aitchison, J. (1986). The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability. Chapman & Hall Ltd., London (UK). (Reprinted in 2003 with additional material by The Blackburn Press). P. 416.
- [2]. Coenders, G., Martín-Fernández, J.A. and Ferrer-Rosell, B., 2017. When relative and absolute information matter: compositional predictor with a total in generalized linear models. Statistical Modelling, 17(6), pp.494-512.
- [3]. Deaton, A., 1997. The analysis of household surveys: a microeconomic approach to development policy. World Bank Publications.

- [4]. Dumuid, D., Stanford, T. E., Martin-Fernández, J. A., Pedišić, Ž., Maher, C. A., Lewis, L. K., ... & Olds, T. (2018). Compositional data analysis for physical activity, sedentary time and sleep research. *Statistical methods in medical research*, 27(12), 3726-3738.
- [5]. Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G. and Barcelo-Vidal, C., (2003), "Isometric logratio transformations for compositional data analysis". *Mathematical Geology*, 35(3), pp.279-300.2003; 35: pp.279–300.
- [6]. Morais, J. and Thomas-Agnan, C., 2021. Impact of covariates in compositional models and simplicial derivatives. *Austrian Journal of Statistics*, 50(2), pp.1-15.
- [7]. Ogundari, K., and Abdulai, A. (2013). Examining the heterogeneity in calorie–income elasticities: A meta-analysis. *Food Policy*, 40, 119-128.
- [8]. Pawlowsky-Glahn, V., Egozcue, J.J. and Lovell, D., 2015. Tools for compositional data with a total. *Statistical Modelling*, 15(2), pp.175-190.
- [9]. SDG, UN. "Sustainable development goals." United Nations (2018).
- [10]. Tổng cục thống kê, T. (2019). Kết quả Khảo sát mức sống dân cư Việt Nam năm 2018. Nhà xuất bản thống kê.
- [11]. Trinh, H.T., Simioni, M. and Thomas-Agnan, C., 2018a. Assessing the nonlinearity of the calo-income relationship: An estimation strategy—With new insights on nutritional transition in Vietnam. *World Development*, 110, pp.192-204.
- [12]. Trinh, H.T., Simioni, M. and Thomas-Agnan, C., 2018b "Decomposition of changes in the consumption of macronutrients in Vietnam between 2004 and 2014." *Economics & Human Biology* 31: 259-275.
- [13]. Van den Boogaart, K.G. and Tolosana-Delgado, R., 2013. Analyzing compositional data with R (Vol. 122). Heidelberg: Springer.



KỶ YẾU

**HỘI THẢO KHOA HỌC QUỐC GIA
“PHÂN TÍCH ĐỊNH LƯỢNG CÁC VẤN ĐỀ
KINH TẾ VÀ XÃ HỘI TRONG MÔI TRƯỜNG SỐ”**

**TRƯỜNG ĐẠI HỌC THƯƠNG MẠI
HỌC VIỆN CHÍNH SÁCH VÀ PHÁT TRIỂN**

815-2022/CXBIPH/6-21/TN

ISBN: 978-604-358-602-2



SÁCH KHÔNG BÁN