

MỤC LỤC

KINH TẾ VÀ QUẢN LÝ

- 1. Nguyễn Việt Thái và Bùi Thị Thanh** - Phân tích tác động không gian của ngành du lịch đến tăng trưởng kinh tế Việt Nam. *Mã số: 137+138.1 TRMg.11* 3
An Analysis of the Spatial Impact of Tourism on Vietnam's Economic Growth
- 2. Nguyễn Mạnh Hùng và Nguyễn Thị Xuân Hồng** - Nghiên cứu hoạt động phát triển nguồn nhân lực du lịch của các tỉnh Trung Du, miền núi Bắc Bộ. *Mã số: 137+138. 1HRMg.11* 10
A Study on Tourism Human Resource Development in Northern Mountainous and Mid-land Provinces
- 3. Đặng Thị Việt Đức** - Cấu trúc cung cầu và các yếu tố ảnh hưởng tới gia tăng sản lượng ngành tài chính ngân hàng Việt Nam giai đoạn 2007-2016. *Mã số: 137+138.1FiBa.11* 28
Input - output structure and sources of output growth of vietnam's banking and finance sector in 2007-2016
- 4. Hoàng Khắc Lịch** - Phân nhóm quốc gia theo tiềm năng và thực tế chi tiêu công. *Mã số: 137+138.1MEco.11* 40
Classifying Countries according to State Spending Potential and Reality
- 5. Nguyễn Thị Cẩm Vân** - Tác động của toàn cầu hóa đến sự phát triển công nghiệp và dịch vụ ở Việt Nam. *Mã số: 137+138.1IIEM.11* 50
The Impact of Globalization on the Development of Industry and Service in Vietnam

QUẢN TRỊ KINH DOANH

- 6. Đỗ Thị Bình** - Nghiên cứu mức độ chủ động trong chiến lược kinh doanh thân thiện với môi trường của các doanh nghiệp chế biến xuất khẩu thủy sản Việt Nam. *Mã số: 137+138.2BMkt.21* 61
A Study on the Activeness in the Environment-Friendly Business Strategy of Vietnam's Aquatic Product Processing and Exporting Enterprises
- 7. Ngô Mỹ Trân và Dương Trọng Nhân** - Các nhân tố ảnh hưởng đến khả năng thành lập các tiểu ban trực thuộc hội đồng quản trị của các công ty niêm yết trên thị trường chứng khoán Việt Nam. *Mã số: 137+138.2OMIS.21* 75
The Factors Affecting the Formation of Subcommittees under Boards of Directors of Listed Companies on Vietnam Stock Market

- 8. Lê Thị Mỹ Phương và Cao Thi Hà Thương** - Phân tích tác động của quản trị tài chính với hiệu quả tài chính tại các doanh nghiệp sản xuất niêm yết trên thị trường chứng khoán Việt Nam. Mã số: 137+138.2FiBa.21 86
An Analysis on the Impact of Financial Administration on Financial Performance at Listed Manufacturing Enterprises on Vietnam Stock Market
- 9. Vũ Thị Thu Hương, Tạ Quang Bình, Hồ Thị Mai Sương và Lương Thị Ngân** - Ảnh hưởng của các công ty zombie đến hiệu quả hoạt động tài chính: Kết quả nghiên cứu thực nghiệm trên các công ty niêm yết nhóm ngành vật liệu xây dựng tại Việt Nam. Mã số: 137+138.2FiBa.21 100
The Impact of Zombie Companies on Financial Performance: Results of Experimental Research at Listed Construction Materials Companies in Vietnam
- 10. Đinh Công Thành, Lê Tấn Nghiêm và Nguyễn Hồng Gấm** - Ảnh hưởng của thuê ngoài dịch vụ đến hiệu quả phi tài chính của doanh nghiệp - nghiên cứu trường hợp các doanh nghiệp vừa và nhỏ tại Đồng bằng Sông Cửu Long. Mã số: 137+138.2BAdm.21 109
The effect of outsourcing on the non-financial performance of smes in the mekong delta

Ý KIẾN TRAO ĐỔI

- 11. Hervé B. BOISMERY** - Entrepreneurship and Credit Crunch in Vietnam: A Recurring Reality? 119
Doanh nghiệp và thắt chặt tín dụng ở Việt Nam: thực trạng tái xuất hiện? Mã số: 137+138.3FiBa.31
- 12. YU-HUI LIN and JIA-CHING JUO** - Risk-Adjusted Productivity Change of Taiwan's Banks in The Financial Holding Companies 133
Thay đổi năng suất điều chỉnh rủi ro của các ngân hàng Đài Loan trong các công ty cổ phần tài chính. Mã số: 137+138.3FiBa.31

PHÂN NHÓM QUỐC GIA THEO TIỀM NĂNG VÀ THỰC TẾ CHI TIÊU CÔNG

Hoàng Khắc Lịch

Trường ĐH Kinh Tế, ĐHQGHN

Email: hoangkhalich@gmail.com

Ngày nhận: 01/12/2019

Ngày nhận lại: 02/01/2020

Ngày duyệt đăng: 06/01/2020

Phân loại là một công việc có ý nghĩa trong mọi lĩnh vực của đời sống kinh tế, xã hội. Phân loại giúp các tổ chức và cá nhân có thể xác định được những đối tượng (hoặc sản phẩm, sự việc, vấn đề...) có sự tương đồng về một hoặc nhiều khía cạnh. Từ đó, các giải pháp tương ứng có thể được đề xuất một cách phù hợp nhất với từng nhóm, đảm bảo có tính đặc thù hướng tới nâng cao chất lượng và hiệu quả. Tùy theo từng mục đích mà cách phân loại được sử dụng sao cho phù hợp. Ví dụ, nghiên cứu về mức sống của các nước, người ta thường sử dụng cách phân loại của Ngân hàng Thế giới; Nghiên cứu về tài chính quốc tế, người ta thường sử dụng cách phân loại của IMF; Nghiên cứu về phát triển con người, người ta thường sử dụng cách phân loại của UNDP. Trong bài viết này, tác giả đề xuất một cách phân loại các quốc gia, phù hợp cho các mục đích nghiên cứu về chính sách tài khóa nói chung và chi tiêu công nói riêng. Phương pháp K-means được dùng để phân cụm; Các kỹ thuật phân tích Elbow và chỉ số Silhouette được dùng để xác định số nhóm tối ưu; Kỹ thuật kiểm định Kruskal-Wallis H được dùng để xác định biến số có ý nghĩa trong phân loại. Kết quả là tổng cộng 134 quốc gia có số liệu trong giai đoạn từ năm 2008 (dấu mốc khủng hoảng tài chính toàn cầu) tới nay được tách thành 03 nhóm theo GNI bình quân, tính hiệu quả của chính phủ và quy mô chi tiêu công.

Từ khóa: Phân loại quốc gia, Chi tiêu công, K-means, Elbow, Silhouette.

1. Giới thiệu

Phân loại là việc sắp xếp các đối tượng vào các nhóm có đặc điểm tương đồng như nhau. Sử dụng kết quả phân loại, người nghiên cứu có thể đề xuất được những giải pháp thiết thực, cụ thể và phù hợp cho từng nhóm. Do đó, phân loại là một hoạt động có ý nghĩa quan trọng trong thực tiễn, được áp dụng trong tất cả các lĩnh vực của đời sống, kinh tế, xã hội của quốc gia. Ví dụ, phân loại học sinh dựa vào học lực của từng em. Các trường học có thể dựa vào điểm thi môn Toán để xếp các em vào nhóm học

giỏi, trung bình và kém. Việc này được thực hiện lặp lại đối với môn Lý, Hóa và các môn khác. Căn cứ vào kết quả phân loại, trường học đề xuất ra các giải pháp nhằm bồi dưỡng kiến thức cho những học sinh thuộc nhóm học yếu, trung bình và giỏi theo những mức độ bồi dưỡng khác nhau.

Phân loại được ứng dụng trong nghiên cứu kinh tế khá phổ biến. Cụ thể đối với nghiên cứu ở tầm vĩ mô, người ta thực hiện phân loại quốc gia bằng nhiều cách. Theo kết quả tổng quan của Hoàng Khắc Lịch (2018), các quốc gia trên thế giới có thể

Nghiên cứu này được tài trợ bởi Quỹ Phát triển khoa học và công nghệ Quốc gia (NAFOSTED) trong đề tài mã số 502.01-2018.308

được phân nhóm theo tiêu chí địa lý, thu nhập, hoặc các chỉ số phát triển. Tùy theo từng mục đích mà lựa chọn các tiêu chí sao cho phù hợp. Ví dụ, Ngân hàng Thế giới quan tâm tới mức sống của người dân các quốc gia, do đó lựa chọn tiêu chí là GNI bình quân đầu người để phân loại thành nhóm các nước có thu nhập thấp, trung bình, và cao. IMF quan tâm tới sức khỏe tài chính của các nước nên lựa chọn các tiêu chí kinh tế vĩ mô. UNDP quan tâm tới sự phát triển của con người nên xây dựng công thức tính chỉ số HDI. Điều đó có nghĩa rằng, tùy theo từng mục đích cụ thể, các nhà nghiên cứu kinh tế vận dụng các cách phân loại sẵn có, hoặc có thể tự lựa chọn ra những tiêu chí phù hợp để tiến hành phân loại lại các nước.

Có vẻ như rất thuận tiện khi phân loại các quốc gia thành nhóm phát triển và đang phát triển. Tuy nhiên, không thể có một cách phân loại phù hợp trong mọi trường hợp. Nielsen (2011) đã nêu ra một số lý do, với việc đầu tiên là xác định ranh giới giữa các nước đang phát triển và phát triển không phải lúc nào cũng rõ ràng. Chương trình phát triển quốc gia của Liên hợp quốc đề xuất hệ thống phân loại quốc gia dựa trên chỉ số phát triển con người (HDI) và Báo cáo phát triển con người (HDR). HDI là một chỉ số tổng hợp của ba chỉ số tuổi thọ, giáo dục và thu nhập. Trong những năm qua, chỉ số đã được tinh chỉnh, nhưng cấu trúc cơ bản của chỉ số không thay đổi.

Ngân hàng Thế giới có ba cách phân loại chính các quốc gia, đó là theo khu vực địa lý, theo thu nhập và tình hình vay nợ. Các nhóm được xác định bởi tiêu chí địa lý ít có ý nghĩa kinh tế hơn vì các quốc gia trong cùng khu vực có thể khác nhau đáng kể về cấu trúc và hiệu suất kinh tế. Cách phân loại theo thu nhập có xu hướng được nhiều nhà nghiên cứu vận dụng hơn. Ngân hàng Thế giới sử dụng phương pháp riêng của mình để tính toán GNI bình quân đầu người, thiết lập ba ngưỡng phân định cho bốn nhóm nước. Về mặt tình hình vay nợ, có ba nhóm chính là IDA, IBRD hoặc kết hợp cả hai nhóm này.

Hệ thống phân loại của IMF giống như hệ thống của Ngân hàng Thế giới theo cách nó được sử dụng cho cả mục đích hoạt động và phân tích. Các quốc gia được chia thành hai nhóm chính: nền kinh tế tiên tiến, thị trường mới nổi và nền kinh tế đang phát triển. Trong báo cáo năm 2018, có 39 nền kinh tế tiên tiến, với 7 nền kinh tế thuộc nhóm các nền kinh tế tiên tiến lớn, bao gồm Hoa Kỳ, Nhật Bản, Đức, Pháp, Ý, Anh và Canada. Có 154 quốc gia còn lại trong danh mục thứ hai. Thị trường mới nổi và các nền kinh tế đang phát triển được phân loại theo các tiêu chí phân tích, phản ánh các thành phần nguồn thu từ xuất khẩu và sự khác biệt giữa các nước chủ nợ và nước đi vay.

Trong nghiên cứu này, tác giả đề xuất một cách phân loại đặc thù, dùng để phục vụ cho các nghiên cứu về chính sách tài khóa, chi tiêu công và tăng trưởng kinh tế ở các quốc gia. Mục đích chính là nhằm xác định các nhóm quốc gia đồng nhất theo các tiêu chí về thực tế và tiềm năng chi tiêu công. Giai đoạn từ 2008 đến nay được lựa chọn thu thập số liệu nhằm tập trung vào khoảng thời gian sau khủng hoảng tài chính toàn cầu. Căn cứ vào kết quả phân loại, các nghiên cứu định tính có thể tìm hiểu đặc điểm chi tiêu công và tăng trưởng ở các nước có đặc điểm tương đồng để rút ra các bài học kinh nghiệm trong thiết kế và thực thi chính sách. Các nghiên cứu định lượng có thể mô tả các con số thống kê theo từng nhóm nước để thấy được đặc điểm chung và riêng trong từng nhóm, hoặc sử dụng kết quả phân loại để phân tích hồi quy cho từng nhóm nước nhằm tìm ra những tác động của chi tiêu công tới tăng trưởng kinh tế.

Tác giả cho rằng, để phục vụ cho các nghiên cứu về chi tiêu công sau này, các quốc gia có thể được phân loại căn cứ vào bốn câu hỏi chính sau: Thứ nhất, quốc gia giàu đến mức nào? Thực vậy, trên cả khía cạnh lý thuyết và thực tiễn, quốc gia càng giàu sẽ càng có tiềm lực tài chính tốt để tài trợ cho các dịch vụ công. Nghiên cứu của Hoàng Khắc Lịch (2019) cho thấy quy mô chi tiêu của chính phủ có xu

hướng tăng theo mức GDP bình quân. Thứ hai, độ mở của nền kinh tế ra sao? Thực tế cho thấy, hội nhập kinh tế có thể khiến nguồn thu ngân sách bị giảm đi do áp lực phải dỡ bỏ các hàng rào thuế quan, tuy nhiên cũng có thể được tăng lên nhờ sự mở rộng về quy mô của các hoạt động (Baunsgaard & Keen, 2010). Thứ ba, chính phủ nước đó quản lý có hiệu quả không? Bergh và Karlsson (2010) lưu ý rằng, quốc gia có quy mô chi tiêu công lớn thường có chất lượng thể chế tốt hơn, ví dụ Thụy Điển và Thụy Sĩ. Thứ tư, quy mô chi tiêu của chính phủ hiện tại như thế nào? Trong đó, các câu hỏi 1-3 đề cập tới tiềm năng chi tiêu công; câu hỏi 4 đề cập tới thực tế chi tiêu công. Toàn bộ số liệu dùng để phân loại được thu thập từ Ngân hàng Thế giới.

Phương pháp phân loại được sử dụng trong nghiên cứu này là kỹ thuật phân cụm K-Means. Đây là một kỹ thuật phân loại đơn giản nhưng được sử dụng rộng rãi đối với yêu cầu phân loại theo nhiều tiêu chí. Để lựa chọn số nhóm tối ưu và những biến phù hợp, nghiên cứu này sử dụng phương pháp Elbow và kiểm định Kruskal-Wallis H. Ngoài ra, chỉ số Silhouette cũng được phân tích nhằm đánh giá mức độ hợp lý của việc phân nhóm. Để khắc phục vấn đề khuyết số liệu trong từng năm cụ thể đối với một số nước, và để không phải loại đi nhiều nước do vấn đề này, tác giả sử dụng số liệu trung bình của cả giai đoạn 2008-2017. Kết quả là 134 quốc gia được chia thành 03 nhóm dựa trên 03 tiêu chí (GNI bình quân, hiệu quả của chính phủ; quy mô chi tiêu của chính phủ). Độ mở của nền kinh tế không tạo ra sự khác biệt giữa các nhóm. Đáng chú ý, GNI thể hiện là một tiêu chí phân loại tốt, giúp phân tách các nhóm khá rõ ràng. Tuy nhiên, dưới sự hiện diện của 02 tiêu chí còn lại, các điểm ngưỡng phân chia giữa các nhóm theo GNI không giống như cách phân loại của Ngân hàng Thế giới. Đa số các nước có thu nhập thấp và trung bình (theo cách phân loại của Ngân hàng Thế giới) được xếp vào cùng một nhóm. Các nước có thu nhập cao được tách thành 02 nhóm nhỏ với sự khác biệt rõ ràng.

Phần còn lại của bài viết được bố cục như sau: Mục 2 trình bày về phương pháp nghiên cứu và số liệu. Mục 3 trình bày kết quả phân cụm. Mục 4 là kết luận.

2. Phương pháp nghiên cứu và số liệu

2.1. Phương pháp Elbow và chỉ số Silhouette

Như đã đề cập trong mục Giới thiệu, nghiên cứu này sử dụng kỹ thuật phân cụm K-Means để thực hiện phân loại các quốc gia. Theo Hartigan và Wong (1979), cơ chế phân loại quốc gia được thực hiện như sau: Các cụm được hình thành căn cứ vào sự gần gũi của các quan sát bên trong, trong khi đó bản thân các cụm cần phải có sự khác biệt nhất có thể. Giả sử mỗi quan sát có p đặc điểm, được thể hiện thông qua vector $x = (x_1, x_2, \dots, x_p)$. Sự khác biệt giữa hai quan sát i và j được đo lường bằng khoảng cách Euclidean, $d_{ij} = \sqrt{\sum_{l=1}^p (x_{il} - x_{jl})^2}$. Việc phân cụm các quan sát được thực hiện dựa vào khoảng cách tương đối này. Giả sử ban đầu có k cụm được hình thành. Các quan sát sau đó được sắp xếp lại giữa các nhóm sao cho tổng khoảng cách giữa các quan sát trong nhóm đạt mức nhỏ nhất, thể hiện sự đồng nhất trong các cụm.

Kỹ thuật phân cụm k-means dễ áp dụng tuy nhiên chất lượng của việc phân loại phụ thuộc vào số nhóm (k) dự kiến ban đầu. Có nghĩa là, tương ứng với mỗi giá trị của k ($k=1,2,\dots, K$) chúng ta đều có thể thực hiện việc phân loại các quan sát, tuy nhiên mức độ đồng nhất giữa các quan sát tùy theo số nhóm k mà chúng ta xác định. Do đó, lựa chọn số nhóm tối ưu luôn là một nhiệm vụ quan trọng đầu tiên cần thực hiện khi áp dụng kỹ thuật này.

Trong nghiên cứu này, số nhóm tối ưu được lựa chọn bằng phương pháp Elbow. Theo đó, chúng ta cần tìm điểm ngoặt trên đường cong thể hiện các giá trị của tổng bình phương sai số bên trong (WSS) hoặc giá trị logarithm của nó $[\log(WSS)]$ cho tất cả các giá trị của k . Lưu ý rằng, WSS càng nhỏ thì tính đồng nhất trong các nhóm càng cao. Vì không nên phân ra quá nhiều nhóm, do đó đồ thị thể hiện WSS được vẽ với $k \leq 10$. Số nhóm tối ưu được chọn tại vị

trí điểm ngoặt bởi vì kể từ vị trí đó, tăng số nhóm k không làm giảm đáng kể giá trị của WSS. Ngoài ra, các tiêu chí khác cũng có thể được sử dụng để xác định số nhóm tối ưu như hệ số η_k^2 , tương tự như R^2 , hoặc tỷ lệ giảm của sai số (*PRE*). Trong đó, hệ số η_k^2 đo lường sự sụt giảm của WSS trên tổng bình phương sai số (TSS) đối với mỗi lời giải cho k nhóm. Ngược lại, PRE_k biểu diễn tỷ lệ sụt giảm của WSS đối với lời giải cho k nhóm so với giá trị của nó nếu giải cho k-1 nhóm (Schwarz, 2008)¹.

Sau khi thực hiện phân loại các quốc gia theo kỹ thuật k-means, tác giả phân tích chỉ số Silhouette để xem xét mức độ hợp lý của kết quả phân loại. Chỉ số Silhouette cho biết mức độ phù hợp của việc tạo nhóm và ước lượng khoảng cách trung bình bên trong các cụm. Đồ thị Silhouette thể hiện sự gần gũi, phù hợp của mỗi quan sát đối với cụm mà nó được sắp xếp vào, cũng như cho biết khả năng phù hợp của nó với cụm lân cận. Chỉ số Silhouette có giá trị lớn nhất bằng 1, thể hiện việc sắp xếp nhóm hoàn toàn hợp lý. Ngược lại, giá trị nhỏ nhất của chỉ số Silhouette là -1, thể hiện sự sắp xếp hoàn toàn không hợp lý, có nghĩa là, quan sát đó nên được sắp xếp vào nhóm khác. Do đó, Silhouette analysis đôi khi cũng được sử dụng trong các nghiên cứu để lựa chọn số nhóm tối ưu.

2.2. Kiểm định Kruskal-Wallis H

Bên cạnh việc xác định số nhóm tối ưu, lựa chọn các tiêu chí (biến số) phù hợp để phân loại cũng là một nhiệm vụ quan trọng. Kết quả phân loại số liệu đòi hỏi các nhóm phải khác biệt theo các tiêu chí (biến số) đã lựa chọn. Nếu biến số nào không khác biệt về mặt giá trị giữa các nhóm, thì nó không phù hợp để làm căn cứ phân loại xét theo khía cạnh kỹ thuật phân tích số liệu. Để loại bỏ những biến số không phù hợp, nghiên cứu này sử dụng kiểm định Kruskal-Wallis H. Kiểm định Kruskal-Wallis H là phương pháp kiểm định phi tham số dựa trên thứ

bậc. Nó được xem là phương pháp phi tham số tương tự như phương pháp ANOVA một chiều (đôi khi được gọi là ANOVA một chiều về thứ bậc). Sau khi chọn được biến số phù hợp dùng để phân loại quốc gia bằng kiểm định Kruskal-Wallis H, nghiên cứu này lặp lại bước xác định số nhóm tối ưu và tiến hành phân tích kết quả.

2.3. Số liệu

Như đã trình bày trong mục giới thiệu, mục đích phân loại các quốc gia trong nghiên cứu này là để xác định những nhóm nước có sự đồng nhất về mặt thực tế và tiềm năng chi tiêu công. Thực tế chi tiêu công ở các nước thể hiện ở tỷ trọng chi tiêu của chính phủ so với tổng GDP. Tiềm năng chi tiêu công thể hiện ở GNI bình quân đầu người, hiệu quả của chính phủ, và độ mở của nền kinh tế thể hiện qua kim ngạch thương mại. Tất cả những số liệu này đều được thu thập từ hệ thống số liệu của Ngân hàng Thế giới.

Lý do lựa chọn các biến thể hiện cho tiềm năng chi tiêu công được giải thích như sau. Thứ nhất, quốc gia giàu sẽ có những mối quan tâm khác so với các quốc gia nghèo. Ví dụ, quốc gia giàu quan tâm nhiều hơn tới an sinh xã hội, cải thiện môi trường, duy trì và phát triển các giá trị văn hóa..., trong khi các nước nghèo quan tâm nhiều tới tiếp nhận chuyển giao công nghệ, phát triển cơ sở hạ tầng cho sản xuất, xóa đói giảm nghèo... Thứ hai, độ mở của nền kinh tế một mặt cho thấy sự hội nhập, đóng góp của xuất nhập khẩu vào GDP, mặt khác đòi hỏi các quốc gia phải tháo dỡ các hàng rào thuế quan theo cam kết của các hiệp định song phương và đa phương. Do đó, độ mở của nền kinh tế ảnh hưởng tới nguồn thu ngân sách, tác động tới việc cân đối chi tiêu công ở các nước. Thứ ba, tính hiệu quả trong quản lý công có thể là một yếu tố quan trọng ảnh hưởng tới niềm tin của dân chúng đối với chính phủ. Theo đó, chi tiêu công có thể tiếp tục được mở rộng ngay

1. $\eta_k^2 = 1 - \frac{WSS(k)}{WSS(1)} = 1 - \frac{WSS(k)}{TSS}$ và $PRE_k = \frac{WSS(k-1) - WSS(k)}{WSS(k-1)}$, trong đó $WSS(1)$ là WSS cho lời giải với $k = 1$ (tức là số liệu không được phân cụm).

cả khi quy mô hiện tại của nó đã ở mức cao, trường hợp Thụy Điển là một điển hình (Bergh và Henrekson, 2011).

3. Kết quả

Những phân tích ban đầu bằng phương pháp Elbow (sử dụng tất cả các biến số, Hình 1a) cho

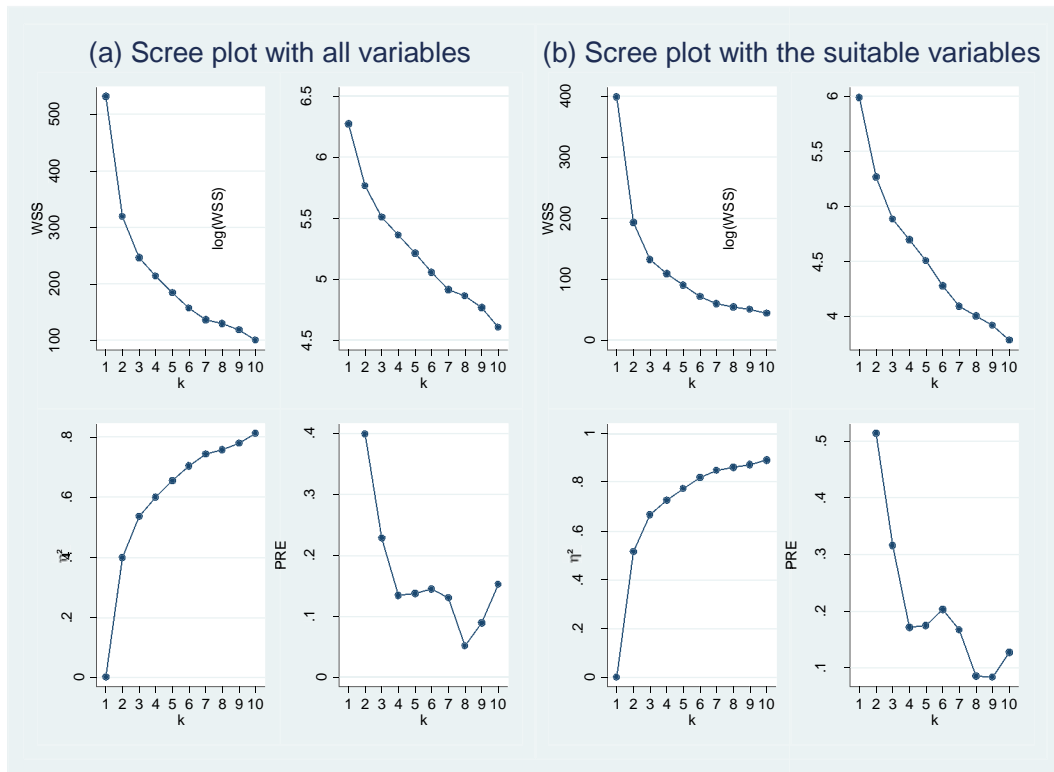
Bảng 1: Danh sách các biến dùng cho phân loại

Tên biến	Định nghĩa	Lý do sử dụng
TOTL	Tổng chi tiêu của chính phủ (% GDP)	Thực tế chi tiêu công
TRADE	Kim ngạch thương mại (% của GDP)	Biến số này có thể gây ra những biến động đối với thu ngân sách khi mà hội nhập vừa mang lại cơ hội tăng cường xuất khẩu ròng, nhưng cũng yêu cầu các quốc gia phải dỡ bỏ các hàng rào thuế quan để thực hiện tự do hóa thương mại.
GE_EST	Hiệu quả của chính phủ	Chỉ số hiệu quả của chính phủ thể hiện chất lượng của dịch vụ công, chất lượng dịch vụ dân sự và mức độ độc lập khỏi áp lực chính trị, chất lượng xây dựng và thực thi chính sách và độ tin cậy của cam kết của chính phủ đối với các chính sách đó.
GNI	GNI bình quân đầu người, theo phương pháp Atlas (đồng Đôla Mỹ hiện hành)	GNI tự nó đã là một chỉ số được các tổ chức (WB, IMF, UNDP) sử dụng để phân loại các quốc gia.

Để khắc phục vấn đề khuyết số liệu, tác giả sử dụng giá trị trung bình của cả giai đoạn 2008-2017. Bằng cách này, số lượng quốc gia còn lại trong bộ số liệu là nhiều nhất, đạt 134 quốc gia trong tổng số 217 quốc gia trên thế giới. Hơn nữa, việc dùng giá trị trung bình của cả giai đoạn để phân loại cung cấp một góc nhìn về sự ổn định của các quốc gia, thay vì những biến động trong từng năm. Trong bộ số liệu này, quốc gia nghèo nhất là Malawi với mức thu nhập bình quân là \$399, trong khi quốc gia giàu nhất là Norway với \$91.328. Nước có chỉ số hiệu quả chính phủ thấp nhất là Cộng hòa Trung Phi, trong khi Singapore có chỉ số hiệu quả chính phủ cao nhất. Singapore cũng là nước có độ mở kinh tế lớn nhất (254,1% GDP), và Timor-Leste có độ mở nhỏ nhất (14,1% GDP). Về thực tế chi tiêu công, quốc gia ở Tây Phi là Equatorial Guinea là nước có tỷ trọng chi tiêu công thấp nhất, chỉ đạt 7,2% GDP, trong khi có những nước có tỷ lệ chi tiêu công cao trên 50% GDP như Hy Lạp, Marshall Islands và Kiribati.

thấy, điểm ngoặt của WSS và $\log(WSS)$ bắt đầu tại $k=3$. Điểm ngoặt cũng thể hiện rõ trên đồ thị của η_2 . Với 3 nhóm này, kiểm định Kruskal-Wallis H cho thấy các biến số có ý nghĩa trong phân loại gồm có GNI, GE_EST, và TOTL ($pvalue < 0.05$). Trong khi TRADE không tạo ra được sự khác biệt giữa các nhóm ($pvalue = 0.3219$). Do vậy, nghiên cứu thực hiện lặp lại bước xác định số nhóm tối ưu với số lượng biến đã rút gọn (gồm 03 biến kể trên, loại trừ TRADE). Kết quả cho thấy các biến số này đều phù hợp và số lượng nhóm tối ưu được gợi ý bằng phương pháp Elbow trở nên rõ ràng hơn tại vị trí $k=3$ (Hình 1b). Kết quả phân loại cho thấy có 21 nước thuộc nhóm 1, 28 nước thuộc nhóm 2, và 85 nước thuộc nhóm 3 (danh sách các nhóm được trình bày trong Bảng 2, Phụ lục).

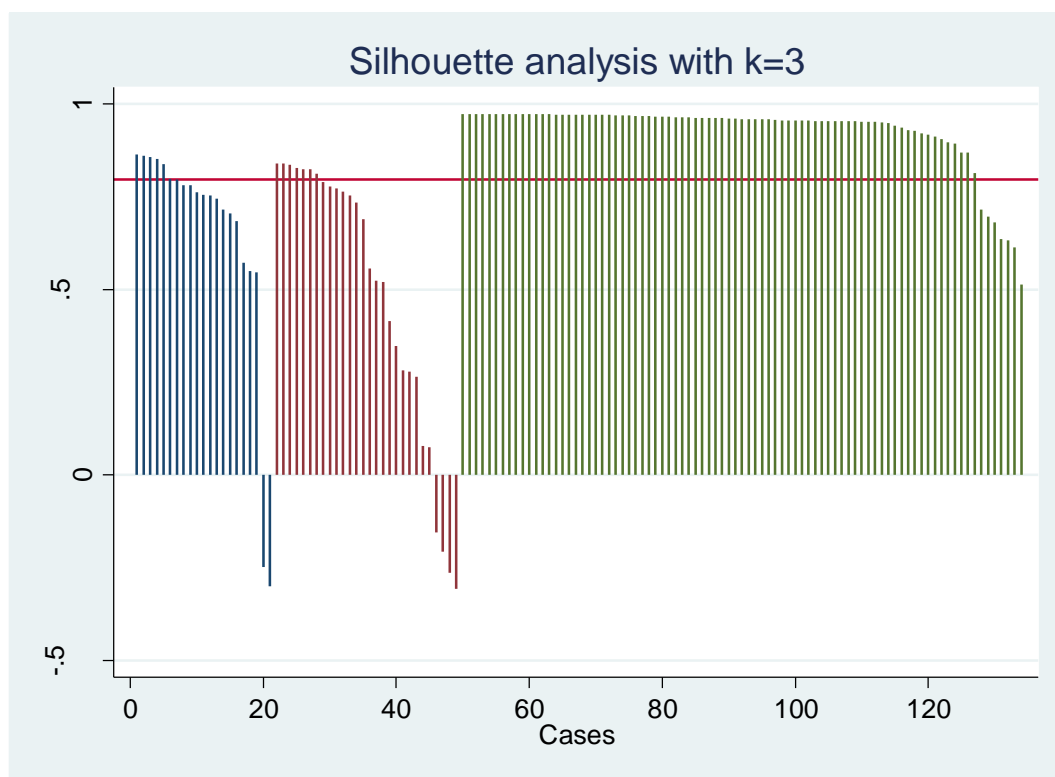
Ngoài ra, Hình 2 cho thấy chỉ số Silhouette trung bình đạt mức cao, xấp xỉ 0.79. Chỉ số Silhouette của các nhóm (sắp xếp theo thứ tự từ 1 tới 3) lần lượt là 0.65, 0.48, và 0.93. Điều này có nghĩa là tính đồng



Hình 1: WSS, log(WSS), η^2 , và PRE đối với các giá trị của k

nhất của nhóm 3 là cao nhất, tiếp theo lần lượt là nhóm 1, và nhóm 2. Chỉ có 06 trong tổng số 134 trường hợp có chỉ số Silhouette nhỏ hơn 0, cụ thể gồm: Italy, New Zealand (thuộc nhóm 1), Argentina, Equatorial Guinea, Russian Federation, và Turkey (nhóm 2). Những trường hợp có chỉ số Silhouette thấp thường là những nước có các chỉ số phân loại đạt mức nhỏ nhất, hoặc thấp nhất trong nhóm phân bố, cho nên chúng nằm ở ranh giới phân định giữa các nhóm. Mặc dù tồn tại một vài trường hợp có chỉ số Silhouette thấp (có nghĩa là, mức độ phù hợp thấp đối với nhóm hiện tại), nhưng đa số các trường hợp còn lại đều có chỉ số cao hơn 0.5 (121 trường hợp, chiếm trên 90% của tổng thể). Như vậy, chúng ta có thể nhận định rằng việc phân loại đã xác định được các nhóm quốc gia đồng nhất về các tiêu chí GNI bình quân, tính hiệu quả của chính phủ và tỷ trọng chi tiêu của chính phủ trên GDP.

Để xác định đặc điểm của các nhóm quốc gia, chúng ta sử dụng biểu đồ hộp để thể hiện ngũ phân vị trong mỗi nhóm (gồm có: giá trị nhỏ nhất, lớn nhất, trung vị, phần tư thứ nhất và phần tư thứ ba). Hình 3 cho thấy GNI bình quân phân biệt các nhóm rất rõ ràng. Trên \$35.000 là các nước thuộc nhóm 1, gồm các nước cực giàu; từ \$11.000 đến \$35.000 là nhóm 2; và dưới \$11.000 là nhóm 3. Các điểm ngưỡng này không giống như các điểm ngưỡng mà Ngân hàng Thế giới đã xác định (\$1.025, \$4.035, \$12.475 lần lượt là các điểm ngưỡng phân biệt 04 nhóm nước có thu nhập thấp, trung bình thấp, trung bình cao và thu nhập cao). Hầu hết những nước có mức thu nhập thấp, trung bình thấp và trung bình cao (theo cách phân loại của Ngân hàng Thế giới, còn được gọi là các nước đang phát triển) đều thuộc nhóm 3 (theo cách phân



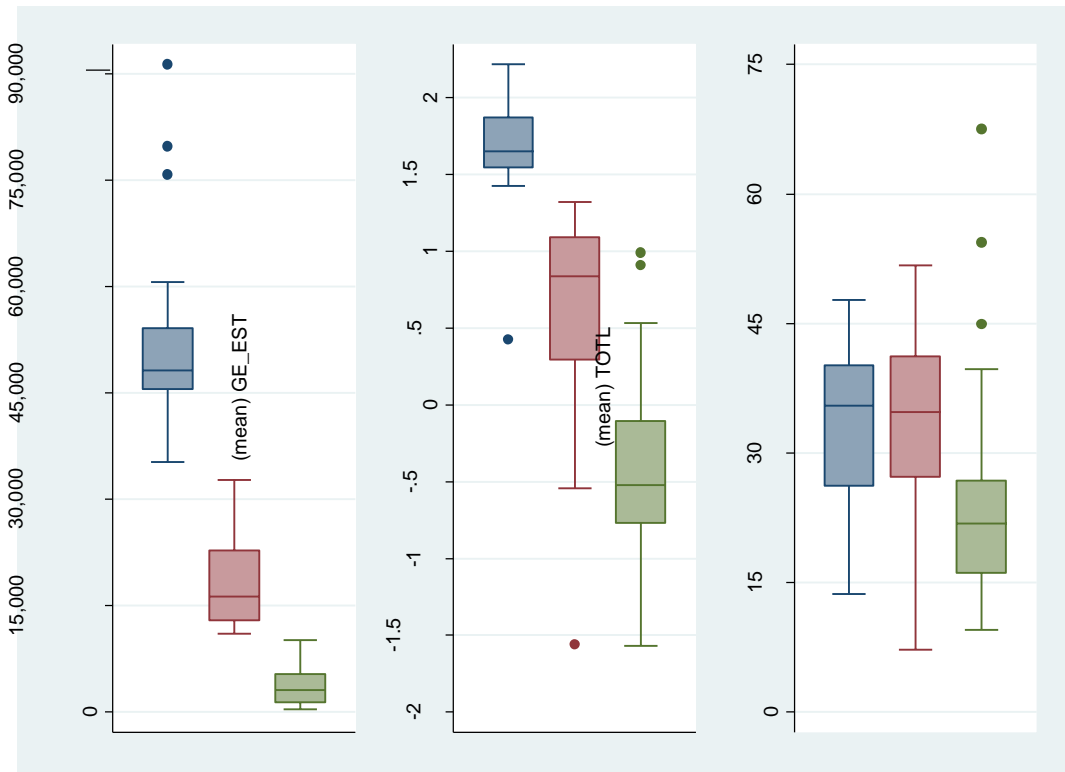
Hình 2: Biểu đồ chỉ số Silhouette đối với 03 nhóm

loại của nghiên cứu này). Đối với các nước có thu nhập cao (theo cách phân loại của Ngân hàng Thế giới) được sắp xếp vào nhóm 1 và nhóm 2 trong nghiên cứu này.

Rõ ràng, sự khác biệt về kết quả phân loại của nghiên cứu này so với của Ngân hàng Thế giới đến từ cách thức phân loại đa tiêu chí. Những tiêu chí được bổ sung thêm ngoài thu nhập bình quân đã làm thay đổi các điểm ngưỡng phân định giữa các nhóm. Khi xem xét các biến số khác, chúng ta cũng thấy sự khác biệt trong các nhóm, tuy nhiên các điểm ngưỡng không được rõ ràng như đối với biến GNI bình quân. Điều này có nghĩa là hai nước có cùng một mức chỉ số hiệu quả của chính phủ (hoặc quy mô chi tiêu của chính phủ) có thể được sắp xếp vào hai nhóm khác nhau. Không giống như cách phân loại dựa trên 1 biến số duy nhất, cách phân loại này

đã xác định được những nhóm quốc gia có mức độ tương đồng về nhiều tiêu chí. Hay nói cách khác, các quốc gia thuộc cùng một nhóm trong nghiên cứu này có sự đồng nhất về nhiều khía cạnh khác nhau liên quan tới thực tế và tiềm năng chi tiêu công.

Có thể phát biểu một cách tương đối về đặc điểm của các nhóm quốc gia như sau: Nhóm có thu nhập cao thường có chính phủ hoạt động hiệu quả và dành một tỷ trọng lớn hơn cho chi tiêu của chính phủ. Sự tương quan giữa mức độ giàu có và tính hiệu quả của chính phủ thể hiện rõ (Hình 3). Nhưng riêng về tiêu chí quy mô chi tiêu công, các nước thuộc nhóm 2 có sự phân tán, mặc dù có thể thấy có sự khác biệt giữa nhóm 1 và nhóm 3, giá trị trung vị cũng được sắp xếp lần lượt từ cao xuống thấp.



Hình 3: Biểu đồ hộp thể hiện giá trị của 03 biến dùng để phân cụm

4. Kết luận

Tùy theo yêu cầu vận hành và phân tích, mỗi tổ chức quốc tế (WB, IMF, hay UNDP) đều có những lựa chọn khác nhau về bộ tiêu chí phân loại quốc gia, từ đó xác định ra các nhóm quốc gia đồng nhất. Các nhà nghiên cứu kinh tế nên hiểu rõ về mục đích phân loại để sử dụng những kết quả phân loại này. Thậm chí, họ có thể căn cứ vào những tiêu chí khác để tự phân loại quốc gia, phục vụ cho các nghiên cứu đặc thù.

Trong nghiên cứu này, tác giả đã vận dụng kỹ thuật phân cụm K-Means để phân loại các quốc gia dựa trên các tiêu chí về thực tế và tiềm năng chi tiêu công ở các quốc gia trong giai đoạn 2008-2017. Mục đích phân tích là để xác định các nhóm quốc gia đồng nhất, phục vụ cho các nghiên cứu liên quan tới chi tiêu công sau kể từ sau năm 2008,

đánh dấu thời điểm khủng hoảng tài chính toàn cầu. Ở đây, thực tế chi tiêu công thể hiện ở quy mô chi tiêu công đo lường bằng tỷ trọng phần trăm của GDP. Tiềm năng chi tiêu công thể hiện ở GNI bình quân, tính hiệu quả của chính phủ, và độ mở của nền kinh tế. Kết quả phân tích đã xác định được 03 nhóm quốc gia có sự đồng nhất theo 03 tiêu chí phù hợp nhất (kiểm định đối với độ mở của nền kinh tế cho thấy biến này không có ý nghĩa trong phân loại các nước). Cụ thể, nhóm 1 gồm 21 nước, nhóm 2 gồm 28 nước, và nhóm 3 gồm 85 nước. GNI bình quân thể hiện là tiêu chí tốt nhất để phân loại vì nó có các điểm ngưỡng giúp phân định các nhóm một cách rõ ràng.

Mặc dù vậy, nghiên cứu này mới chỉ phân loại được 134 trong tổng số 217 quốc gia trên thế giới. Nhiều nước không được phân loại vì bị thiếu số liệu.

Vì vậy, trong một chừng mực nhất định, nếu thực hiện một số kỹ thuật bổ sung số liệu bị thiếu có thể sẽ có được một bộ số liệu bao gồm nhiều nước hơn, từ đó có được một bức tranh đầy đủ hơn về các nhóm nước. Các nghiên cứu đặc thù tiếp theo có thể áp dụng kỹ thuật và cách tiếp cận tương tự để xác định ra các nhóm nước có sự tương đồng theo một bộ tiêu chí cụ thể khác. ♦

Tài liệu tham khảo:

1. Baunsgaard, T. & Keen, M. (2010), *Tax Revenue and (or?) Trade Liberalization*, Journal of Public Economics, 94(9-10), pp.563-577.
2. Bergh, A. & Karlsson, M. (2010), *Government Size and Growth: Accounting for Economic Freedom and Globalization*, Public Choice, 142(1-2), pp.195-213.
3. Bergh, A. and Henrekson, M., 2011, *Government size and growth: a survey and interpretation of the evidence*, Journal of Economic Surveys, 25(5), pp.872-897.
4. IMF World Economic Outlook, April 2018: *Cyclical Upswing, Structural Change*, April 17, 2018
5. Hartigan, J.A. and Wong, M.A.(1979), *Algorithm AS 136: A k-means clustering algorithm*, Applied Statistics, 100-108.
6. Hoàng Khắc Lịch (2019), *Optimal Public Expenditure in Developing Countries*, VNU Journal of Science: Economics and Business, 35(2).
7. Nielsen, L. (2011), *Classifications of Countries Based on Their Level of Development: How it is Done and How it Could be Done*, IMF Working Paper, No. WP/11/31.
8. Schwarz, A. 2008, *Lokale Scoring-Modelle*, Lohmar, Germany: Eul Verlag.
9. Stefan, R.M. (2012), *A Comparison of Data Classification Methods*, Procedia Economics and Finance, No. 3, pp. 420 - 425.

Summary

Classification is a significant task in any field of the socio-economic life. Classification helps organizations and individuals to determine subjects (such as products, problems or issues) which are similar in one or some aspects, so that appropriate solutions can be suggested for each group to ensure specificity in order to raise quality and efficiency. Depending on specific goals, different ways of classification can be used. For example, in studies on living standards of nations, the classification methods proposed by the World Bank are often employed; in studies on international finance, the classification methods proposed by the IMF are often chosen; in studies on human development, the classification methods of UNDP are often used. In this paper, the author suggests a method of classification for nations in accordance with the research objectives of fiscal policy in general and state spending in particular. K-means methods are used to cluster; Elbow analysis technique and Silhouette indicators are used to determine optimal group numbers; Kruskal-Wallis H testing techniques are used to determine significant variables. The results are a total of 134 nations with statistics in the period since 2008 (the year with the onset of global financial crisis) are classified into 3 groups under average GNI, government efficiency and state spending scales.

