

Xác định dòng thông tin giữa các ngành trên thị trường chứng khoán Việt Nam bằng ma trận transfer entropy

Trần Thị Tuấn Anh^(*)

Ngày nhận bài: 17/02/2020 | Biên tập xong: 04/5/2020 | Ngày duyệt: 10/5/2020

TÓM TẮT: Bài viết sử dụng số liệu về giá đóng cửa của chỉ số chứng khoán tổng hợp của 22 ngành trên thị trường chứng khoán (TTCK) Việt Nam trong giai đoạn 01/2012–01/2020 để tính toán và kiểm định dòng thông tin di chuyển giữa các ngành trên TTCK Việt Nam bằng kỹ thuật transfer entropy. Kết quả nghiên cứu cho thấy, những ngành truyền thông tin đáng kể trên TTCK Việt Nam bao gồm ngành vận tải, thương mại, thực phẩm, thép, nhựa, đầu tư xây dựng và công nghệ là những ngành có transfer entropy ròng mang dấu dương. Ngược lại, ngành có vai trò nhận thông tin mạnh nhất trên thị trường là ngành bất động sản với giá trị transfer entropy ròng có giá trị âm nhất trong số các ngành nhận thông tin. Ngoài ra, các ngành nhận thông tin khác trên thị trường còn có ngành sản xuất kinh doanh và vật liệu xây dựng.

TỪ KHÓA: Dòng chảy thông tin, thị trường chứng khoán, transfer entropy, transfer entropy ròng.

Mã phân loại JEL: C00, G00, G14.

1. Giới thiệu

Mỗi doanh nghiệp hiện diện trên thị trường đều bị ảnh hưởng bởi những biến động thông tin trên thị trường đó. Mỗi hoạt động của doanh nghiệp trên thị trường đều tạo ra thông tin đóng góp vào dòng thông tin của thị trường, và ngược lại, doanh nghiệp không ngừng tiếp nhận thông tin từ thị trường. Thông tin đóng một phần quan trọng trong việc hình thành môi trường kinh doanh. Mỗi thông tin doanh nghiệp tiếp nhận từ thị trường hay thông tin mà doanh nghiệp truyền

ra thị trường đều tạo nên một dòng di chuyển thông tin. Các dòng chảy thông tin này có thể xảy ra dưới nhiều cấp độ khác nhau, giữa các doanh nghiệp, giữa các ngành nghề, giữa các địa phương hoặc giữa các thị trường. Vì vậy, một trong những vấn đề nghiên cứu quan trọng khi xem xét hoạt động của một hệ thống kinh tế dạng mạng lưới là xem xét dòng di

... Trần Thị Tuấn Anh - Trường Đại học Kinh tế TP HCM; 59C Nguyễn Đình Chiểu Quận 3, TP. Hồ Chí Minh; Email: anhtt1@ueh.edu.vn.

chuyển thông tin giữa các đối tượng trong mạng lưới đó. Hàm lượng thông tin của mỗi đối tượng trong hệ thống có thể được đo lường bằng Shannon entropy được đề xuất bởi Shannon (1964). Khi thông tin về đối tượng càng ít để có thể dự đoán trạng thái của đối tượng thì Shannon entropy tính được càng lớn và ngược lại. Tuy nhiên, trong cùng một hệ thống, các đối tượng thường có mối liên hệ với nhau. Các mối liên hệ này có thể giúp dự đoán giá trị tương lai cho đối tượng này dựa vào thông tin của các đối tượng khác. Các mối liên hệ như vậy hình thành nên các dòng chảy thông tin trong hệ thống. Schreiber (2000) đề xuất đại lượng transfer entropy để đo lường và xác định dòng thông tin giữa các đối tượng như vậy. Transfer entropy đo lường mối liên hệ thông tin giữa các đối tượng mà không dựa trên một mô hình cụ thể nào và cũng không dựa trên giả thiết liên hệ phải có dạng tuyến tính. Đây là một lợi thế nổi bật của transfer entropy so với các phương pháp thống kê khác trong cùng mục tiêu xác định mối liên hệ giữa các biến số kinh tế. Với ưu điểm này, ngày càng nhiều các nghiên cứu trên thế giới áp dụng transfer entropy trong nghiên cứu các chuỗi thời gian tài chính.

Tại Việt Nam, đã có rất nhiều các nghiên cứu khảo sát mối liên hệ giữa các chuỗi thời gian trên thị trường tài chính trong nước nhưng việc khảo sát dòng di chuyển thông tin trên TTCK bằng transfer entropy còn rất hạn chế. Vì vậy, bài viết này ứng dụng transfer entropy nhằm đo lường mối liên hệ thông tin giữa các ngành trên TTCK Việt Nam và xác định chiều hướng các dòng thông tin trên thị trường. Với mục tiêu này, phần còn lại của bài viết được tổ chức như sau: Mục 2 tóm tắt các nghiên cứu đã sử dụng transfer entropy trong nghiên cứu các chuỗi thời gian tài chính, Mục 3 giới thiệu dữ liệu sử dụng và các phương pháp tính toán thực hiện trong đề tài, Mục 4

xử lý số liệu và thảo luận kết quả nghiên cứu và Mục 5 kết luận những kết quả chính của đề tài và một số hàm ý rút ra từ kết quả đó.

2. Tổng quan các nghiên cứu có liên quan

Trong các nghiên cứu trước đây, để khảo sát sự trao đổi thông tin giữa các đối tượng trên thị trường, các nhà nghiên cứu thường sử dụng phân tích nhân quả Granger (GCA-Granger Causality Analysis). Phân tích GCA sử dụng kiểm định thống kê để kiểm tra xem liệu chuỗi thời gian của một đối tượng này có mang lại các thông tin hữu ích để dự đoán cho chuỗi thông tin của một đối tượng khác. Nếu tác động Granger tồn tại, thì xem như một đối tượng cung cấp thông tin đóng vai trò truyền và một đối tượng được dự báo đóng vai trò nhận. Phân tích nhân quả Granger thường được thực hiện hai chiều từ đó xác định dòng chảy cũng như chiều hướng di chuyển thông tin giữa hai đối tượng đó. Tuy nhiên, phân tích nhân quả Granger được xây dựng dựa trên giả định về mối liên hệ dạng tuyến tính giữa chuỗi thời gian và tính chuẩn trong phân phối của chuỗi. Tuy nhiên, các giả định này ít khi được thỏa mãn trong thực tế. Để khắc phục nhược điểm này, đồng thời kế thừa những thành quả quan trọng trong việc ứng dụng lý thuyết thông tin trong phân tích kinh tế-tài chính, Schreiber (2000) đã xây dựng và đề xuất tính toán transfer entropy để định lượng sự gắn kết thông tin giữa các chuỗi thời gian của hai đối tượng. Khắc phục nhược điểm của các phương pháp thống kê truyền thống trước đó khi chưa xem xét được sự bất đối xứng trong dòng thông tin liên hệ giữa các đối tượng, transfer entropy giúp xác định được độ mạnh và chiều hướng thông tin giữa các chuỗi thời gian cũng như cho thấy rõ sự bất cân xứng trong các dòng di chuyển thông tin.

Dimpfl & Peter (2013) sử dụng transfer entropy để xác định dòng di chuyển thông tin

giữa các thị trường tài chính và đề xuất thủ tục suy diễn thống kê bằng bootstrap đối với transfer entropy mà không phụ thuộc vào các giả định tuyến tính trong kiểm định thống kê thông thường. Các tác giả cũng ứng dụng phương pháp do chính các tác giả đề xuất vào xác định, đo lường và kiểm định mối liên hệ thông tin giữa thị trường hoán đổi tín dụng và thị trường trái phiếu để định giá rủi ro tín dụng. Các tính toán được thực hiện cho cả ba giai đoạn trước, trong và sau cuộc khủng hoảng kinh tế năm 2009.

Guo, Yang, & Yu (2015) nhận thấy, sự gia tăng trong tính phức tạp của các quy trình công nghiệp hiện đại làm nảy sinh sự cần thiết phải có các phương pháp phát hiện tính nhấn mạnh nhằm xác định vai trò nguyên nhân-kết quả, từ đó chẩn đoán nguyên nhân gốc rễ khi có sự cố trên hệ thống. Do vậy, các tác giả đề xuất một đại lượng mới mở rộng của transfer entropy và đặt tên là transfer entropy xu thế (trend transfer entropy). Đại lượng transfer entropy xu thế tập trung vào phân tích xu hướng của chuỗi thời gian thay vì chuỗi gốc và do đó, mạnh hơn trong điều kiện phân tích các chuỗi có tính xu thế. Một ưu điểm khác là giảm thiểu được mức độ tính toán cần thực hiện. Kết quả của các tác giả được kiểm chứng trên các chuỗi dữ liệu mô phỏng bằng phần mềm.

Daugherty & Jithendranathan (2015) nghiên cứu mức độ hội nhập của 20 TTCK trên thế giới với TTCK của Mỹ thông qua phương pháp tỷ số phương sai (variance ratios), phương pháp hệ số tương quan có điều kiện (conditional correlations) và transfer entropy. Kết quả nghiên cứu của các tác giả cho thấy, cuộc khủng hoảng nhà đất năm 2008-2009 thực sự có ảnh hưởng đến mức độ hội nhập của các thị trường. Tuy nhiên, tác động của cuộc khủng hoảng nợ châu Âu năm 2011-2012 không có tác động mạnh đến các đại lượng đo lường mức độ hội nhập này.

Teng & Shang (2017) đề xuất một mở rộng của transfer entropy, đặt tên là transfer

entropy hệ số (transfer entropy coefficient) với mục tiêu định lượng mức độ luồng thông tin giữa các chuỗi thời gian tài chính thông qua việc thực hiện tính toán transfer entropy đa lớp (multiscale). Việc thực hiện tính toán entropy hệ số này được minh họa trên chuỗi thời gian mô phỏng và ứng dụng phân tích một số chuỗi thời gian tài chính trên các thị trường thực tế. Kết quả cho thấy mối liên hệ thông tin giữa các chuỗi thời gian có thể được phát hiện bởi transfer entropy hệ số nhưng không được phát hiện bằng transfer entropy thông thường.

Một trong những khó khăn gây trở ngại cho việc sử dụng transfer entropy đó là khó khăn khi tính toán các xác suất đồng thời, xác suất biên và xác suất có điều kiện trong công thức của đại lượng này. Behrendt & ctg (2019) đã xây dựng một gói lệnh trên phần mềm R, đặt tên là RTransferEntropy để thực hiện chức năng tính toán transfer entropy giữa các chuỗi thời gian. Các tác giả cũng đã thử nghiệm kết quả trên các chuỗi thời gian mô phỏng và cho thấy sự ổn định của kết quả tính toán thu được. Các kết quả tính toán và kiểm định transfer entropy trong bài viết này cũng được thực hiện bằng gói lệnh RTransferEntropy trên phần mềm R do Behrendt & ctg (2019) đề xuất.

Với ưu điểm có thể áp dụng được trên các chuỗi có mối quan hệ phi tuyến hoặc phân tích có dạng phi tham số, transfer entropy đã được áp dụng rộng rãi trong xác định dòng di chuyển thông tin giữa các chuỗi thời gian trong các nghiên cứu trên thế giới. Tuy nhiên, phương pháp này ít được đề cập cũng như chưa được áp dụng phổ biến trên thị trường Việt Nam, vì vậy, bài nghiên cứu này xác định mục tiêu nghiên cứu áp dụng transfer entropy để xác định dòng chảy thông tin giữa các ngành trên TTCK Việt Nam.

3. Dữ liệu và phương pháp nghiên cứu

3.1. Dữ liệu

Bài viết sử dụng dữ liệu là chỉ số chứng khoán tổng hợp của tất cả các ngành trên TTCK hàng ngày. Theo cách phân ngành của trang web cophieu68.vn, các công ty niêm yết được chia thành 25 ngành theo danh sách được thể hiện ở Bảng 1. Tuy nhiên, chỉ có 22 ngành có đủ số liệu để tính toán nên bài viết thực hiện phân tích số liệu của 22 ngành này như danh sách ở Bảng 3. Chỉ số chứng khoán được thu thập theo giá đóng cửa hàng ngày trong giai đoạn 01/2012-01/2020. Từ số liệu về giá đóng cửa, bài viết tính toán tỷ suất sinh lợi hàng ngày của các ngành theo công thức logarit thông thường:

$$R_{it} = 100 \times \ln \frac{P_{it}}{P_{i,t-1}} \quad (1)$$

Trong đó: R_{it} - tỷ suất sinh lợi của ngành i ở thời điểm t ; và P_{it} - chỉ số chứng khoán của ngành i tại thời điểm t .

Bảng thống kê mô tả về số quan sát cũng như chỉ số đóng cửa hàng ngày của tất cả 25 ngành được thể hiện trong Bảng 2. Từ Bảng 2 có thể nhận thấy, có ba ngành không đủ số quan sát để đưa vào phân tích cùng các ngành

khác, bao gồm dịch vụ công ích, hàng không và xây dựng. Do vậy, với 22 ngành được chọn còn lại, thống kê mô tả tỷ suất sinh lợi hàng ngày trên thị trường được thể hiện trong Bảng 3.

3.2. Phương pháp nghiên cứu

Shannon (1948) giới thiệu ứng dụng entropy trong nhiệt động lực học vào lý thuyết thông tin để đo lường lượng thông tin trung bình của một biến ngẫu nhiên X rời rạc theo công thức như sau:

$$H = - \sum_j p(j) \log_2 p(j) \quad (2)$$

Trong đó: chỉ số j dùng để chỉ trường hợp thứ j có thể xảy ra của X và $p(j)$ là xác suất để biến X rơi vào trường hợp thứ j này. Shannon entropy được tính toán theo công thức 2 càng lớn có nghĩa là biến X càng có ít thông tin để dự đoán giá trị có thể xảy ra của X .

Nếu xem xét hai biến ngẫu nhiên X và Y với xác suất đồng thời $p(x,y)$, lượng thông tin chung của hai biến ngẫu nhiên được xác định bằng công thức entropy như sau:

$$M_n = \sum_{ij} p(i,j) \log_2 \left(\frac{p(i,j)}{p(i)p(j)} \right) \quad (3)$$

Bảng 1: Danh sách ngành trên TTCK Việt Nam

STT	Tên ngành	STT	Tên ngành
1	Bất động sản	14	Dầu khí
2	Cao su	15	Nhựa - Bao bì
3	Chứng khoán	16	Phân bón
4	Công nghệ viễn thông	17	Sản xuất-kinh doanh
5	Dịch vụ - Du lịch	18	Thực phẩm
6	Dịch vụ công ích	19	Thương mại
7	Dược phẩm/y tế/hóa chất	20	Thủy sản
8	Giáo dục	21	Vận tải/cảng/taxi
9	Hàng không	22	Vật liệu xây dựng
10	Khoáng sản	23	Xây dựng
11	Năng lượng điện/khí	24	Đầu tư phát triển
12	Ngân hàng - Bảo hiểm	25	Đầu tư xây dựng
13	Ngành thép		

Nguồn: <https://www.cophieu68.vn/categorylist.php>.

Trong đó: $p(i)$ là xác suất biên theo X và $p(j)$ là xác suất biên theo Y . Thông tin chung giữa X và Y cho biết mức độ phụ thuộc giữa hai biến, nghĩa là khi lượng thông tin chung càng lớn, hai biến càng có mối liên hệ chặt chẽ; nghĩa là biết được thông tin biến này có thể giúp dự đoán thông tin và biến kia và ngược lại. Tuy nhiên, thông tin chung M_{XY} nhưng không cho biết chiều hướng truyền thông tin giữa hai biến được xét. Để khắc phục nhược điểm này, Schreiber (2000) giới thiệu khái niệm transfer entropy giúp tách bạch được lượng thông tin của X hữu ích để dự đoán Y và ngược lại, lượng thông tin của Y hữu ích để dự đoán X . Hai chiều hướng này xảy ra không giống nhau, vì vậy, transfer entropy không có tính đối xứng. Những chiều hướng xảy ra như vậy giữa các biến ngẫu nhiên được xem là dòng chảy thông tin.

Đối với các chuỗi thời gian X_t và Y_t , transfer entropy được sử dụng đo lường lượng thông tin chung được chia sẻ giữa dãy số liệu quá khứ của X và hiện tại của Y , xét điều kiện theo những thông tin quá khứ của Y , thể hiện bằng công thức:

$$TE_{X \rightarrow Y} = \sum \left(p(y_{t+1}, y_t^{(k)}, x_t^{(l)}) \log_2 \frac{p(y_{t+1}, y_t^{(k)}, x_t^{(l)})}{p(y_{t+1}, y_t^{(k)})} \right) \quad (4)$$

Trong đó: $p(y_{t+1} | y_t^{(k)}, x_t^{(l)})$ - xác suất của y_{t+1} , xét điều kiện theo các thông tin trong quá khứ đến độ trễ k của Y và độ trễ l của X ;

$p(y_{t+1} | y_t^{(k)})$ - xác suất của y_{t+1} , xét điều kiện theo các thông tin trong quá khứ đến độ trễ k của Y ; và $p(y_{t+1}, y_t^{(k)}, x_t^{(l)})$ - xác suất đồng thời của $(y_{t+1}, y_t^{(k)}, x_t^{(l)})$.

Nói một cách khác, transfer entropy giúp đo lường xem liệu việc sử dụng thông tin trong quá khứ của cả X và Y để dự báo Y sẽ tốt hơn khi dự báo chỉ với thông tin quá khứ của Y như thế nào.

Trên TTCK, mỗi chuỗi tỷ suất sinh lợi của

chỉ số chứng khoán từng ngành có thể xem như một chuỗi thời gian và transfer entropy ứng dụng trong trường hợp này giúp xác định dòng chảy thông tin giữa các ngành thông qua tỷ suất sinh lợi. Nếu ký hiệu $TE_{j \rightarrow i}$ là transfer entropy đo lường dòng chảy thông tin từ ngành j đến ngành i , thì ngành j sẽ là ngành truyền thông tin và ngành i là ngành nhận thông tin. Tương tự, $TE_{i \rightarrow j}$ sẽ cho biết lượng thông tin di chuyển theo chiều ngược lại, từ ngành i đến ngành j . Nhìn chung, $TE_{j \rightarrow i} \neq TE_{i \rightarrow j}$, nếu xét thị trường như một hệ thống mạng lưới mà mỗi ngành là một mắt lưới thì mũi tên sẽ được hướng từ i sang j nếu $TE_{j \rightarrow i} - TE_{i \rightarrow j} > 0$.

Thực hiện tính toán tất cả các transfer entropy giữa các ngành và thể hiện dưới dạng ma trận n dòng n cột với n là số ngành được xem xét. Các phần tử trên đường chéo chính của ma trận transfer entropy sẽ để trống vì người ta không xem xét dòng chảy thông tin từ một ngành vào chính ngành đó. Tổng các phần tử trên dòng i , ký hiệu là $TE_i^{out} = \sum_{j=1}^n TE_{i \rightarrow j}$, sẽ cho biết tổng lượng thông tin mà ngành i nhận được từ tất cả các ngành khác. Tương tự, tổng các phần tử trên cột j , ký hiệu là $TE_j^{in} = \sum_{i=1}^n TE_{i \rightarrow j}$, sẽ cho biết tổng lượng thông tin mà ngành j truyền đi đến tất cả các ngành khác. Gọi transfer entropy ròng của ngành i là chênh lệch giữa tổng transfer entropy ngành i nhận vào và transfer entropy ngành i truyền đi. Nghĩa là: $TE_i^{net} = TE_i^{out} - TE_i^{in}$

Khi đó, ngành i sẽ đóng vai trò truyền thông tin trên TTCK nếu $TE_i^{net} > 0$ và ngành i sẽ được gọi là ngành nhận thông tin nếu $TE_i^{net} < 0$.

Tóm lại, bài viết sử dụng tỷ suất sinh lợi của 22 ngành như danh sách ở Bảng 3 để lập ma trận transfer entropy của thị trường, đồng thời tính toán tổng lượng thông tin các ngành

nhân vào và truyền đi cũng như xác định vai trò truyền hay nhận thông tin giữa các ngành trên từng thị trường.

4. Kết quả và thảo luận

4.1. Thống kê mô tả dữ liệu

Số quan sát trong mẫu dữ liệu thu thập được cũng như các thống kê mô tả cơ bản của chỉ số chứng khoán của 25 ngành trên thị trường được thể hiện ở Bảng 2 của bài nghiên cứu. Dựa trên Bảng 2 có thể thấy rằng, trong số 25 ngành trên thị trường, chỉ có 22 ngành có tương đối đầy đủ số liệu trong suốt quãng thời gian thu thập mẫu. Do vậy chỉ có 22 ngành được đưa vào tính toán và kiểm định dòng chảy thông tin trên thị trường. Còn lại ba ngành gồm dịch vụ công ích, hàng không và xây dựng không có đầy đủ số liệu nên chưa

được đưa vào phân tích.

Bảng 3 thể hiện thống kê chuỗi tỷ suất sinh lợi hàng ngày của 22 ngành có đủ số liệu trong giai đoạn 01/2012–01/2020. Kết quả mô tả trong Bảng 3 cho thấy, có ba ngành có tỷ suất sinh lợi trung bình mang dấu âm, bao gồm cao su, khoáng sản và phân bón. Ngành năng lượng có tỷ suất sinh lợi trung bình cao nhất thị trường, tiếp đến là ngành thương mại và ngành giáo dục. Cũng như các thị trường khác trên thế giới, các ngành trên TTCK Việt Nam cũng sẽ có mối liên hệ về mặt thông tin. Bài viết tiến hành đo lường đồng thông tin trao đổi giữa các ngành và thể hiện thành ma trận transfer entropy như ở Bảng 4. Để việc trình bày trong các bảng biểu được ngắn gọn, bài viết sẽ sử dụng số thứ tự ngành trong cột 1 của Bảng 3 để thay cho tên ngành trong các bảng tiếp theo.

Bảng 2: Thống kê mô tả các chỉ số ngành trên TTCK

Ngành	Số quan sát	Trung bình	Độ lệch chuẩn	Giá trị nhỏ nhất	Giá trị lớn nhất
Bất động sản	2.002	513,57	167,44	314,56	939,82
Cao su	2.002	412,33	167,66	186,60	758,57
Chứng khoán	2.002	481,57	102,07	306,85	891,33
Công nghệ	2.002	580,52	195,06	283,20	1047,12
Dầu khí	2.002	557,98	160,51	277,14	1117,80
Dịch vụ	2.002	999,71	289,42	330,39	1599,02
Đầu tư phát triển	2.002	672,92	161,77	347,31	1636,92
Đầu tư xây dựng	2.002	432,37	80,76	237,01	627,04
Dược phẩm	2.002	776,63	215,59	354,57	1206,29
Dịch vụ công ích	1.159	274,19	139,88	121,58	1400,34
Giáo dục	2.002	1036,10	353,49	341,45	1728,07
Hàng không	1.892	2115,74	1374,74	202,52	5430,60
Khoáng sản	2.002	165,86	83,73	71,73	366,09
Năng lượng	2.002	1132,21	363,96	344,29	1634,75
Ngân hàng	2.002	492,43	120,38	309,78	947,30
Nhựa	2.002	1054,34	445,44	338,30	2009,96
Phân bón	2.002	511,98	95,24	286,93	787,63

Ngành	Số quan sát	Trung bình	Độ lệch chuẩn	Giá trị nhỏ nhất	Giá trị lớn nhất
Sản xuất kinh doanh	2.003	895,24	293,61	333,70	1519,06
Thép	2.002	651,07	162,89	350,25	1253,27
Thực phẩm	2.002	645,98	218,44	324,92	1121,92
Thương mại	2.002	704,42	282,19	332,67	1276,53
Thủy sản	2.003	647,86	147,92	349,82	936,23
Vận tải	2.002	624,57	132,81	305,30	813,67
Vật liệu xây dựng	2.002	869,16	358,66	287,83	1539,42
Xây dựng	1.769	758,23	482,04	263,15	2159,66

Nguồn: Tác giả tính toán từ số liệu thu thập.

Bảng 3: Thống kê mô tả tỷ suất sinh lời của các chỉ số ngành

STT	Ngành	Số quan sát	Trung bình	Độ lệch chuẩn	Giá trị nhỏ nhất	Giá trị lớn nhất
1	Bất động sản	2.000	0,040	1,380	-6,906	5,227
2	Cao su	2.000	-0,007	1,451	-5,815	4,587
3	Chứng khoán	2.000	0,011	1,754	-8,180	6,496
4	Công nghệ	2.000	0,053	1,237	-7,528	4,994
5	Dầu khí	2.000	0,017	1,877	-8,042	6,225
6	Dịch vụ	2.000	0,054	1,853	-10,035	8,692
7	Đầu tư phát triển	2.000	0,041	2,174	-12,395	10,770
8	Đầu tư xây dựng	2.000	0,021	1,420	-7,439	4,628
9	Dược phẩm	2.000	0,046	1,077	-6,443	5,489
10	Giáo dục	2.000	0,057	1,621	-8,196	9,004
11	Khoáng sản	2.000	-0,077	1,985	-14,172	10,763
12	Năng lượng	2.000	0,070	1,187	-6,776	4,908
13	Ngân hàng	2.000	0,034	1,472	-7,052	5,555
14	Nhựa	2.000	0,052	1,291	-5,717	5,137
15	Phân bón	2.000	-0,009	1,345	-8,218	5,625
16	Sản xuất kinh doanh	2.000	0,056	1,078	-6,236	6,837
17	Thép	2.000	0,031	1,702	-11,273	7,323
18	Thực phẩm	2.000	0,044	1,157	-6,927	4,478
19	Thương mại	2.000	0,063	1,320	-6,522	5,120
20	Thủy sản	2.000	0,047	1,507	-7,416	5,256
21	Vận tải	2.000	0,026	1,144	-5,494	4,499
22	Vật liệu xây dựng	2.000	0,051	1,471	-9,134	4,971

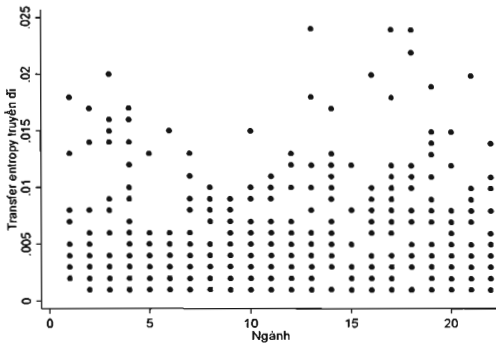
Nguồn: Tác giả tính toán từ số liệu thu thập

4.2 Kết quả kiểm định dòng chảy thông tin bằng transfer entropy

Bảng 4: Ma trận transfer entropy giữa các ngành

Ngành	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	TRUYỀN
1	0,004	0,020	0,016	0,005	0,005	0,003	0,010	0,003	0,004	0,004	0,012	0,024	0,017	0,008	0,020	0,018	0,004	0,014	0,015	0,020	0,010	0,235	
2	0,003		0,008	0,003	0,013	0,003	0,008	0,003	0,003	0,003	0,003	0,008	0,003	0,003	0,003	0,010	0,005	0,003	0,003	0,004	0,003	0,097	
3	0,013	0,005		0,009	0,002	0,001	0,002	0,009	0,005	0,003	0,002	0,005	0,012	0,006	0,002	0,008	0,012	0,011	0,007	0,003	0,008	0,010	0,133
4	0,008	0,001	0,006		0,003	0,001	0,001	0,001	0,001	0,002	0,011	0,005	0,010	0,007	0,008	0,009	0,006	0,001	0,015	0,002	0,007	0,003	0,107
5	0,004	0,006	0,002	0,006		0,001	0,001	0,001	0,001	0,001	0,006	0,006	0,007	0,001	0,001	0,007	0,010	0,001	0,009	0,001	0,009	0,002	0,080
6	0,004	0,002	0,002	0,004	0,001		0,009	0,002	0,009	0,007	0,010	0,007	0,003	0,003	0,002	0,007	0,004	0,004	0,008	0,008	0,005	0,008	0,107
7	0,002	0,017	0,001	0,002	0,006	0,005		0,002	0,005	0,007	0,003	0,004	0,002	0,003	0,003	0,002	0,003	0,004	0,002	0,003	0,004	0,002	0,078
8	0,002	0,001	0,001	0,005	0,001	0,001	0,001		0,003	0,001	0,001	0,001	0,001	0,001	0,002	0,001	0,001	0,001	0,001	0,001	0,001	0,002	0,028
9	0,005	0,001	0,004	0,001	0,002	0,002	0,004	0,003		0,006	0,010	0,002	0,005	0,007	0,002	0,003	0,007	0,022	0,004	0,003	0,003	0,014	0,110
10	0,003	0,002	0,002	0,003	0,002	0,015	0,013	0,002	0,006		0,004	0,005	0,003	0,004	0,002	0,010	0,002	0,007	0,005	0,005	0,002	0,002	0,009
11	0,003	0,001	0,001	0,012	0,002	0,002	0,011	0,001	0,007	0,006		0,003	0,002	0,003	0,012	0,002	0,003	0,004	0,011	0,001	0,002	0,001	0,09
12	0,004	0,002	0,004	0,006	0,002	0,006	0,004	0,003	0,004	0,009	0,006		0,008	0,011	0,005	0,009	0,009	0,005	0,015	0,005	0,009	0,003	0,12
13	0,007	0,003	0,014	0,010	0,002	0,001	0,001	0,007	0,003	0,002	0,003	0,003		0,007	0,005	0,009	0,010	0,008	0,019	0,012	0,008	0,011	0,14
14	0,004	0,002	0,004	0,004	0,002	0,001	0,002	0,004	0,002	0,007	0,002	0,003	0,005		0,002	0,009	0,009	0,004	0,013	0,002	0,008	0,005	0,09
15	0,005	0,002	0,004	0,014	0,003	0,004	0,003	0,005	0,003	0,003	0,009	0,003	0,003	0,003		0,003	0,003	0,003	0,019	0,006	0,003	0,004	0,10
16	0,005	0,001	0,015	0,004	0,002	0,015	0,006	0,001	0,006	0,015	0,006	0,010	0,018	0,013	0,001		0,024	0,024	0,015	0,006	0,009	0,014	0,21
17	0,004	0,008	0,015	0,003	0,004	0,002	0,002	0,002	0,002	0,002	0,002	0,005	0,008	0,009	0,002	0,009		0,007	0,011	0,005	0,008	0,006	0,11
18	0,003	0,005	0,009	0,003	0,001	0,003	0,003	0,001	0,008	0,003	0,007	0,004	0,008	0,005	0,001	0,004	0,008		0,005	0,002	0,003	0,011	0,0E
19	0,002	0,001	0,002	0,017	0,004	0,002	0,003	0,001	0,008	0,005	0,011	0,013	0,006	0,012	0,012	0,010	0,007	0,010		0,004	0,008	0,008	0,14
20	0,005	0,002	0,006	0,009	0,003	0,004	0,002	0,004	0,002	0,004	0,004	0,006	0,003	0,003	0,003	0,003	0,003	0,004	0,003	0,005		0,003	0,005
21	0,004	0,002	0,003	0,004	0,004	0,006	0,005	0,002	0,004	0,007	0,004	0,003	0,004	0,005	0,002	0,006	0,003	0,003	0,007	0,004		0,001	0,01
22	0,018	0,014	0,016	0,007	0,003	0,003	0,005	0,008	0,006	0,010	0,005	0,006	0,018	0,010	0,002	0,009	0,011	0,012	0,007	0,007	0,010		0,11
NHẬN	0,108	0,082	0,139	0,141	0,066	0,083	0,089	0,071	0,090	0,105	0,108	0,109	0,157	0,137	0,078	0,143	0,162	0,141	0,195	0,097	0,132	0,122	
CL	-0,127	-0,015	0,006	0,034	-0,014	-0,024	0,012	0,043	-0,020	0,007	0,017	-0,021	0,012	0,043	-0,024	-0,067	0,047	0,045	0,049	0,015	0,050	-0,066	

Nguồn: Tác giả tính toán từ số liệu thu thập.



Nguồn: Tác giả tính toán từ số liệu thu thập.

Hình 1: Đồ thị phân tán của transfer entropy theo ngành truyền thông tin

Bảng 4 thể hiện kết quả tính toán transfer entropy giữa mỗi hai ngành được xem xét trong số 22 ngành được lựa chọn để phân tích. Mỗi ngành được đại diện bằng số thứ tự ngành như đã mô tả ở Bảng 3. Phần tử ở dòng i cột j của Bảng 4 thể hiện transfer entropy của dòng thông tin mà ngành i nhận được từ ngành j , nói một cách khác, đây cũng chính là dòng thông tin truyền từ ngành j đến ngành i (tức là $TE_{j \rightarrow i}$). Ngành j là ngành truyền thông tin và ngành i là ngành nhận thông tin. Nếu xét theo dòng, dòng i của Bảng 4 sẽ cho biết lượng thông tin mà ngành i nhận được từ các ngành khác, không xét chính nó. Vì vậy các phần tử trên đường chéo chính được bỏ trống. Tương tự, khi xét theo cột, cột j thể hiện lượng thông tin mà ngành j đã truyền đến các ngành khác. Để kết quả được thể hiện rõ ràng hơn, bài viết dùng biểu đồ phân tán ở Hình 1 và Hình 2 để thể hiện dòng thông tin truyền đi và nhận được của từng ngành. Trục hoành của đồ thị thể hiện số thứ tự của ngành theo Bảng 1 và trục tung là độ lớn của transfer

entropy.

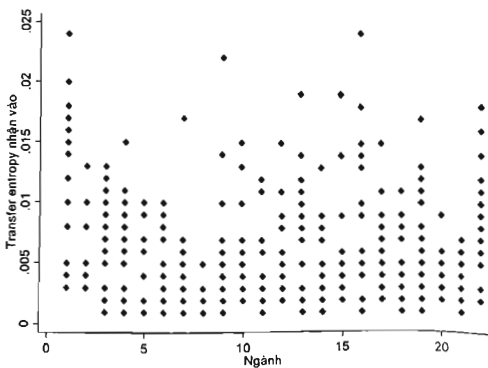
Hình 1 thể hiện transfer entropy truyền đi của từng ngành. Biểu đồ dạng phân tán cho biết độ trải rộng của transfer entropy từng ngành truyền đi các ngành khác. Mỗi một chấm tròn tương ứng với một giá trị j ở trục hoành cho biết mức truyền từ ngành j đến một ngành nào đó. Nếu các chấm tròn càng tập trung ở vị trí thấp thì đó là dấu hiệu cho thấy vai trò truyền thông tin của ngành đó yếu, trong khi nếu những chấm tròn tập trung ở vị trí cao sẽ thể hiện vai trò truyền thông tin mạnh đến các ngành khác. Nếu các chấm tròn phân tán rộng, có nghĩa là ngành tương ứng truyền thông tin mạnh đến một số ngành này nhưng yếu đối với một số ngành khác. Khi ngành i nhận dòng thông tin từ ngành j thì có nghĩa là nhà đầu tư có thể sử dụng các thông tin trong chuỗi tỷ suất sinh lợi chứng khoán của ngành j để dự báo ngành i bên cạnh những thông tin trong quá khứ của chính ngành i . Vì transfer entropy không có tính đối xứng nên

việc sử dụng thông tin ngành j để dự báo ngành i sẽ khác với việc dùng thông tin ngành i để dự báo ngành j .

Kết quả tính toán transfer entropy trên Hình 1 cho thấy có một số ngành có vai trò truyền thông tin yếu hơn các ngành khác như ngành dầu khí, đầu tư xây dựng, dược phẩm, và phân bón. Hầu hết các ngành khác có biểu đồ phân tán trải rộng cho thấy mối quan hệ đa dạng trong dòng thông tin giữa các ngành được xét và các ngành khác. Điển hình trong số này có ngành ngân hàng, thép và thực phẩm. Tra cứu kết quả ở Bảng 4, ngành ngân hàng truyền thông tin rất mạnh đến ngành bất động sản, sản xuất kinh doanh và vật liệu xây dựng. Điều đó có nghĩa là thông tin trong chuỗi tỷ suất sinh lợi của ngành ngân hàng có thể giúp dự đoán tỷ suất sinh lợi của ngành bất động sản, sản xuất kinh doanh và vật liệu xây dựng. Tương tự, dòng thông tin từ ngành thép di chuyển nhiều đến ngành sản xuất kinh doanh và bất động sản. Ngoài ra, dòng thông tin từ ngành thực phẩm di chuyển mạnh đến

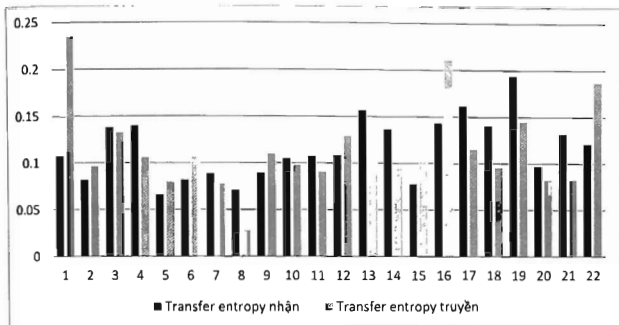
ngành dược phẩm và sản xuất kinh doanh.

Hình 2 thể hiện biểu đồ phân tán xét theo ngành nhận thông tin, nhưng sử dụng các dấu chấm hình kim cương để phân biệt với Hình 1. Các thông tin thể hiện trên Hình 2 cũng chính là các thông tin ở Bảng 4 nhưng được sắp xếp theo độ mạnh transfer entropy của ngành nhận thông tin. Tương tự như Hình 1, những ngành có các chấm tập trung ở vùng thấp là những ngành ít nhận thông tin từ các ngành khác và những ngành có dấu chấm tập trung ở vị trí cao là ngành nhận nhiều dòng thông tin từ các ngành khác. Kết quả thể hiện trên Hình 2 cho thấy, một số ngành nhận rất ít thông tin từ các ngành khác, điển hình như ngành đầu tư xây dựng, vận tải và ngành thủy sản. Trong khi đó, các ngành như bất động sản, sản xuất kinh doanh và dược phẩm nhận thông tin nhiều từ các ngành khác, thể hiện ở việc transfer entropy của dòng thông tin nhận được của các ngành này có giá trị khá cao so với các ngành khác. Kết quả này cũng khớp với những nhận định đã rút ra từ Hình 1.



Nguồn: Tác giả tính toán từ số liệu thu thập.

Hình 2: Đồ thị phân tán transfer entropy theo ngành nhận thông tin



Nguồn: Tác giả tính toán từ số liệu thu thập.

Hình 3: Biểu đồ tổng transfer entropy truyền đi và nhận vào của từng ngành

Trong Bảng 4, tổng trên từng dòng của ma trận transfer entropy sẽ cho biết tổng lượng thông tin mà ngành tương ứng đã nhận từ những ngành khác. Trong số các ngành được phân tích, Bất động sản là ngành tiếp nhận dòng thông tin từ các ngành khác nhiều nhất, điều này có nghĩa là, trong số các ngành được phân tích, ngành bất động sản là ngành có thể dự báo thông tin tỷ suất sinh lợi tốt nhất từ thông tin của các ngành khác với tổng lượng transfer entropy nhận là 0,235. Ngành có tổng dòng thông tin nhận đứng vị trí thứ hai là ngành sản xuất kinh doanh với số tổng thông tin nhận là 0,211. Ngược lại, ngành đầu tư xây dựng là ngành có tổng lượng transfer entropy vào thấp nhất (0,024), thấp hơn rất nhiều so với tất cả các ngành còn lại. Điều này cho thấy rằng việc dùng thông tin chuỗi thời gian của các ngành khác để dự báo cho giá trị tỷ suất sinh lợi tương lai của ngành đầu tư xây dựng sẽ ít có hiệu quả và ít thuyết phục.

Tương tự, tổng trên mỗi cột sẽ cho biết lượng thông tin mà ngành ở cột đó đã truyền đi đến các ngành khác. Nó cho biết mức độ

đóng góp thông tin của một ngành trong việc dự báo cho các ngành khác. Theo kết quả ở Bảng 4, ngành thương mại có tổng lượng transfer entropy truyền đi lớn nhất (0,195) và ngành ngân hàng (0,157) đứng ở vị trí thứ hai khi xét lượng thông tin truyền đi này. Trong khi đó, ngành đầu tư xây dựng (0,071), chứng khoán (0,082) và đầu tư phát triển (0,083) có tổng lượng transfer entropy truyền đi rất thấp, cho thấy lượng thông tin của các ngành này ít có tiềm năng để sử dụng trong dự báo tỷ suất sinh lợi các ngành khác.

Hình 3 biểu diễn tổng transfer entropy truyền đi và nhận vào của từng ngành. Cột màu đậm thể hiện tổng lượng transfer entropy truyền đi và cột màu nhạt thể hiện tổng lượng transfer entropy nhận vào của từng ngành. Biểu đồ dạng cột giúp dễ dàng so sánh tổng transfer entropy nhận vào và truyền đi của từng ngành, đồng thời cũng giúp so sánh giữa các ngành với nhau. Những ngành có lượng thông tin nhận vào nhiều hơn lượng thông tin truyền đi được xem là các ngành nhận thông tin, Ngược lại, các ngành



Nguồn: Tác giả tính toán từ số liệu thu thập.

Hình 4: Biểu đồ chênh lệch giữa transfer entropy nhận và truyền giữa các ngành

có lượng thông tin truyền đi nhiều hơn được gọi là ngành truyền thông tin, những ngành như vậy được xem là có đóng góp vào dòng thông tin chung của toàn thị trường. Nếu gọi transfer entropy ròng là chênh lệch giữa lượng thông tin truyền đi và lượng thông tin nhận vào thông thì một ngành là truyền thông tin nếu transfer entropy ròng mang dấu dương; và ngược lại, nếu transfer entropy ròng mang dấu âm thì ngành được xác định là ngành nhận thông tin. Dòng cuối cùng của Bảng 4 thể hiện giá trị transfer entropy ròng của cả 22 ngành và được biểu diễn trên đồ thị ở Hình 4.

Theo Hình 4, những ngành có cột giá trị transfer entropy ròng phía trên trục hoành, ứng với trường hợp transfer entropy ròng mang dấu dương, là ngành truyền thông tin; và những ngành có cột transfer entropy ròng phía dưới trục hoành, ứng với trường hợp transfer entropy ròng mang dấu âm là những ngành nhận thông tin. Như vậy có thể nhận thấy những ngành truyền thông tin trên TTCK Việt Nam bao gồm ngành vận tải,

thương mại, thực phẩm, thép, nhựa, đầu tư xây dựng và công nghệ.

Ngược lại, ngành có vai trò nhận thông tin mạnh nhất trên thị trường là ngành bất động sản với giá trị transfer entropy ròng có giá trị âm nhất trong số các ngành nhận thông tin. Ngoài ra, các ngành nhận thông tin khác trên thị trường còn có ngành sản xuất kinh doanh và vật liệu xây dựng.

5. Kết luận và gợi ý chính sách

Bài viết sử dụng số liệu về giá đóng của chỉ số chứng khoán tổng hợp của 22 ngành trên TTCK Việt Nam trong giai đoạn 01/2012–01/2020 để tính toán và kiểm định dòng thông tin di chuyển giữa các ngành trên TTCK Việt Nam bằng kỹ thuật transfer entropy. Kết quả nghiên cứu cho thấy, những ngành đóng vai trò truyền thông tin trên TTCK Việt Nam bao gồm ngành vận tải, thương mại, thực phẩm, thép, nhựa, đầu tư xây dựng và công nghệ, là những ngành có transfer entropy ròng mang dấu dương; và ngược lại, ngành có vai

trở nhận thông tin mạnh nhất trên thị trường là ngành Bất động sản với giá trị transfer entropy ròng có giá trị âm nhất trong số các ngành nhận thông tin. Ngoài ra, các ngành nhận thông tin khác trên thị trường còn có ngành sản xuất kinh doanh và vật liệu xây dựng. Như vậy, thông qua kết quả của bài viết, có thể nhận thấy rằng, có một số ngành tương tác thông tin với nhau mạnh hơn các ngành khác và cũng có những ngành cung cấp thông tin ra bên ngoài nhiều hơn là dòng thông tin bản thân ngành đó nhận vào. Việc giải thích cho hiện tượng này cần đòi hỏi những nghiên cứu chuyên sâu hơn về mặt lý thuyết kinh tế. Đây cũng là một hướng mở rộng cần thiết cho bài viết trong các bước nghiên cứu

tiếp theo.

Kết quả nghiên cứu có ý nghĩa với nhà đầu tư để giúp họ xác định được mối liên hệ truyền và nhận thông tin giữa các ngành trên TTCK, để từ đó lựa chọn sử dụng thông tin phù hợp để thực hiện dự báo cho ngành mà họ quan tâm đầu tư. Ngoài ra, kết quả nghiên cứu cũng có thể giúp các nhà hoạch định chính sách thấy được mối liên hệ giữa các ngành trên TTCK, để từ đó, khi xây dựng một chính sách hoặc ban hành một quyết định liên quan đến ngành này, họ có thể lường trước hoặc nhận định được các ngành khác sẽ bị ảnh hưởng gián tiếp như thế nào với quyết định mà nhà làm chính sách đưa ra dựa trên mạng lưới liên kết thông tin giữa các thị trường.

Tài liệu tham khảo

- Behrendt, S., Dimpfl, T., Peter, F. J., & Zimmermann, D. J. (2019) RTransferEntropy — Quantifying information flow between different time series using effective transfer entropy. *SoftwareX*, 10, 100265.
- Daugherty, M. S., & Jithendranathan, T. (2015). A study of linkages between frontier markets and the U.S. equity markets using multivariate GARCH and transfer entropy. *Journal of Multinational Financial Management*, 32-33, 95–115.
- Dimpfl, T. & Peter, F. (2013). Using transfer entropy to measure information flows between financial markets. *Studies in Nonlinear Dynamics and Econometrics*, 17(1), pp. 85-102.
- Guo, C., Yang, F., & Yu, W. (2015). A Causality Capturing Method for Diagnosis Based on Transfer Entropy by Analyzing Trends of Time Series. *IFAC-PapersOnLine*, 48(21), 778–783.
- Schreiber, T. (2000). Measuring Information Transfer. *Physical Review Letters*, 85(2), 461–64.
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27,379–423.
- Teng, Y., & Shang, P. (2017). Transfer entropy coefficient: Quantifying level of information flow between financial time series. *Physica A: Statistical Mechanics and Its Applications*, 469, 60–70.

Determining the Information Flow Among Vietnam Stock Market Industries by Transfer Entropy Matrix

Tran Thi Tuan Anh⁽¹⁾

Received: 17 February 2020 | Revised: 04 May 2020 | Accepted: 10 May 2020

ABSTRACT: This paper uses data on closing daily indices of 22 industries on Vietnam's stock market from January 2012 to January 2020 to calculate and reveal the information flows by transfer entropy. The finding shows that the transportation, trading, food, steel, plastic, construction investment, and technology industries have positive net transfer entropy, and contribute significantly to the global information market flow. In contrast, the industry has the strongest role of receiving information is Real Estate. Other industries are also information recipients on the market include manufacturing, business and construction materials.

KEYWORDS: Information flow, Net transfer entropy, Stock market, Transfer-entropy.

JEL classifications: C00, G00, G14.



Tran Thi Tuan Anh

Email: anhttt@ueh.edu.vn.

⁽¹⁾ University of Economics Ho Chi Minh City;

59C Nguyen Dinh Chieu, District 3, Ho Chi Minh City.