Information flow between stock indices

Okyu Kwon[1,2] and Jae-Suk Yang[1] [(a)]

[1] *Department of Physics, Korea University, Seoul 136-701, Korea*
[2] *National Creative Research Initiative Center for Neuro-dynamics, Korea University, Seoul 136-701, Korea*

Abstract. - Using transfer entropy, we observed the strength and direction of information flow between stock indices. We uncovered that the biggest source of information flow is America. In contrast, the Asia/Pacific region the biggest is receives the most information. According to the minimum spanning tree, the GSPC is located at the focal point of the information source for world stock markets.

Introduction. – Economic systems have recently become an active field of research for physicists striving to transfer concepts and methodologies from statistical physics such as phase transition, fractal theory, spin models, complex networks, and information theory to the analysis of economic problems [1–24].

Among the numerous methodologies put forward, time series analysis has proven to be one of the most efficient methods and is widely applied to the examination of characteristics of stock and foreign exchange markets. In order to analyze financial time series, a range of statistical measures have been introduced, including probability distribution [20–26], autocorrelation [23], multi-fractal [27], complexity [18–20], entropy density [19, 20], and transfer entropy [11].

Information is a keyword in analyzing financial market data or in estimating the stock price of a given company. It is quantified through a variety of methods such as cross-correlation, autocorrelation, and complexity. However, while they may be appropriate measures for the observation of the internal structure of information flow, they fail to illuminate the directionality of information flow. Schreiber [28] introduced transfer entropy, which measures the dependency in time between two variables and notes the directionality of information flow. This concept of transfer entropy has already been applied to the analysis of financial time series. Marschinski and Kantz [29] calculated information flow between the Dow Jones and DAX stock indexes to better observe interactions between the two huge markets. Kwon and Yang [11] measured the di-

rection of information flow between the composite stock index and individual stock prices, while the information flow among individual stocks in a stock market has been estimated by Baek and coauthors [30] in order to measure the internal structure of a stock market.

Through transfer entropy, this paper focuses quantitatively on the direction of information flow between 25 stock markets to determine which market serves as a source of information for global stock indices. Drawn from the economic system, a plethora of empirical data reflecting economic conditions can be obtained. The time series of a composite stock price index provides prime data accurately reflecting economic conditions. Therefore, we analyzed the daily time series of the 25 stock indices listed in Table 1 for the period of 2000 – 2007 using transfer entropy in order to examine the information flow between stock markets and identify the hub.

Transfer entropy. – Transfer entropy, which measures the directionality of a variable with respect to time was recently introduced by Schreiber [28] based on the probability density function (PDF). Let us consider two discrete and stationary process, I and J. Transfer entropy relates k previous samples of process I and l previous samples of process J and is defined as follows:

$$T_{J \to I} = \sum p(i_{t+1}, i_t^{(k)}, j_t^{(l)}) \log \frac{p(i_{t+1} \mid i_t^{(k)}, j_t^{(l)})}{p(i_{t+1} \mid i_t^{(k)})}, \quad (1)$$

where i_t and j_t represent the discrete states at time t of I and J, respectively. $i_t^{(k)}$ and $j_t^{(l)}$ denote k and l dimensional delay vectors of two time consequences I and J,

---

[(a)]Corresponding author.

Americas	1	MERV	Argentina
	2	BVSP	Brazil
	3	GSPTSE	Canada
	4	MXX	Mexico
	5	GSPC	U.S.
	6	DJA	U.S.
	7	DJI	U.S.
Asia/Pacific	8	AORD	Australia
	9	SSEC	China
	10	HSI	China
	11	BSESN	India
	12	JKSE	Indonesia
	13	KLSE	Malaysia
	14	N225	Japan
	15	STI	Singapore
	16	KS11	Korea
	17	TWII	Taiwan
Europe	18	ATX	Austria
	19	BFX	Belgium
	20	FCE.NX	France
	21	GDAXI	Germany
	22	AEX	Holland
	23	MIBTEL	Italy
	24	SSMI	Switzerland
	25	FTSE	U.K.

Table 1: List of 25 markets. We obtain data from the website http://finance.yahoo.com.

respectively. The joint PDF $p(i_{t+1}, i_t^{(k)}, j_t^{(l)})$ is the probability that the combination of i_{t+1}, $i_t^{(k)}$ and $j_t^{(l)}$ have particular values. The conditional PDF $p(i_{t+1} \mid i_t^{(k)}, j_t^{(l)})$ and $p(i_{t+1} \mid i_t^{(k)})$ are the probability that i_{t+1} has a particular value when the value of previous samples $i_t^{(k)}$ and $j_t^{(l)}$ are known and $i_t^{(k)}$ are known, respectively.

The transfer entropy with index $J \rightarrow I$ measures the extent to which the dynamics of process J influences the transition probabilities of another process I. Reverse dependency is calculated by exchanging i and j of the joint and conditional PDFs. The transfer entropy is explicitly asymmetric under the exchange of i_t and j_t. It can thus provide information regarding the direction of interaction between two time series.

The transfer entropy is quantified by information flow from J to I. The transfer entropy can be calculated by subtracting the information obtained from the last observation of I exclusively from the information about the latest observation I obtained from the final joint observation of I and J. This is the nexus of transfer entropy. Therefore, transfer entropy can be rephrased as

$$T_{J \rightarrow I} = h_I(k) - h_{IJ}(k, l), \qquad (2)$$

where

$$h_I(k) = -\sum p(i_{t+1}, i_t^{(k)}) \log p(i_{t+1} \mid i_t^{(k)}) \quad (3)$$

$$h_{IJ}(k, l) = -\sum p(i_{t+1}, i_t^{(k)}, j_t^{(l)}) \log p(i_{t+1} \mid i_t^{(k)}, j_t^{(l)}). \quad (4)$$

Empirical data analysis. – We analyzed the daily data records of 25 stock exchange indices for the period January 2000 – December 2007. We adopted the log returns to describe a financial time series $x(n) \equiv \ln(S(n)) - \ln(S(n-1))$. Where S_n means the index of n-th trading day. We partitioned the real value $x(n)$ into discretized value $A(n)$; $A(n) = 0$ for $x(n) \leq -d$ (decrease), $A(n) = 1$ for $-d < x(n) < d$ (intermediate), $A(n) = 2$ for $x(n) \geq d/2$ (increase). We set $d = 0.04$ in our analysis. At that value, the probabilities of three states are approximately same for all indices. Additionally, we set $k = l = 1$.
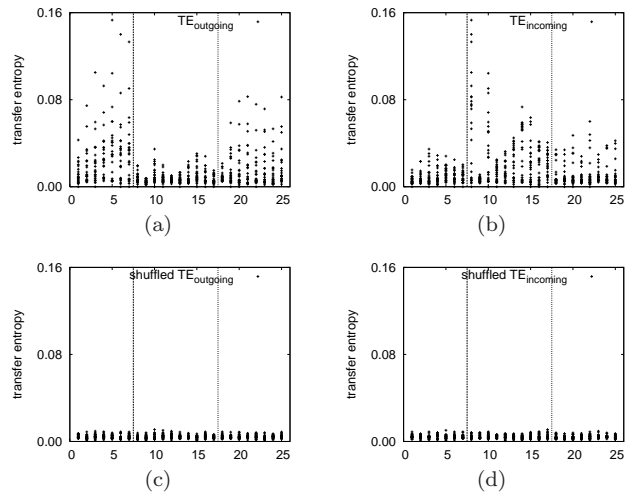

Fig. 1: Plots of (a) the outgoing and (b) the incoming transfer entropy. Plots of (c) the outgoing and (d) the incoming transfer entropy for the shuffled data. Numbers in x-axis indicate markets listed in Table 1.

Fig. 1 is the plots of the outgoing and the incoming transfer entropy. Numbers in x-axis indicate markets listed in Table 1 while the dotted lines in the graphs serve to distinguish the continents. Each stock market interacts with 24 other stock markets, and the transfer entropy for all possible pairs can be calculated. Fig. 1(a) represents the transfer entropy from the market indicating x-axis to the other 24 markets, that is, the outgoing transfer entropy. Fig. 1(b) represents the transfer entropy from the other 24 markets to the market in x-axis, that is, the incoming transfer entropy. The markets in the Americas have high values of outgoing transfer entropy, while the value of the transfer entropy for the markets in Asia remains low. Moreover, the incoming transfer entropy for Asia is higher than that of the Americas and Europe. From Fig. 1, we can determine that the directionality of influence for a market index is generally from the Americas to Asia. To clarify the tendencies of information flow, we shuffled the data and verified the transfer entropy. In Fig. 1(c) and Fig. 1(d), the transfer entropy for the shuffled data is generally approaching zero, information flow

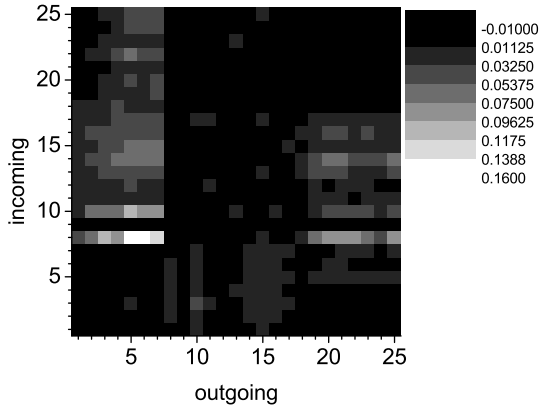between markets has nearly disappeared.

Fig. 2: Gray-scale map of the transfer entropy between stock markets.

Fig. 2 is the gray-scale map of the transfer entropy between stock markets using the same data as in Fig. 1. The direction of information flow is from the x-axis to y-axis. The darker the lattice, the lower the transfer entropy. The markets in the Americas influence the markets in both Asia and Europe, but influence on the Asian markets is particularly strong. The European markets also influence the Asian markets, but to a lesser degree than do the American markets.
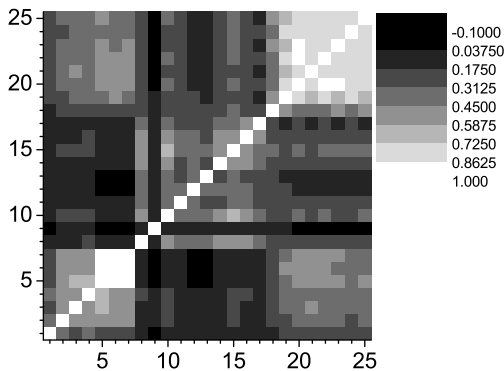

Fig. 3: Gray-scale map of the cross-correlations between stock markets.

Fig. 3 is the gray-scale map of the cross-correlations between stock markets from log-return time series $x(n)$. The cross-correlation does not represent directionality between markets. Therefore, the graph is symmetrical, centering around the diagonal. In this graph, a light lattice means

that the cross-correlation is high. The American and European markets are highly correlated between themselves, while the Asian markets are not when compared with the American and European markets. Also, the American and Asian markets are rarely correlated with each other. From these results, we can induce (or that the comparatively mature American and European markets are well clustered between themselves, while the relatively emerging Asian markets are not. Furthermore, between the American and Asian markets there exists low cross-correlation and high transfer entropy, demonstrating an imbalance of information flow.
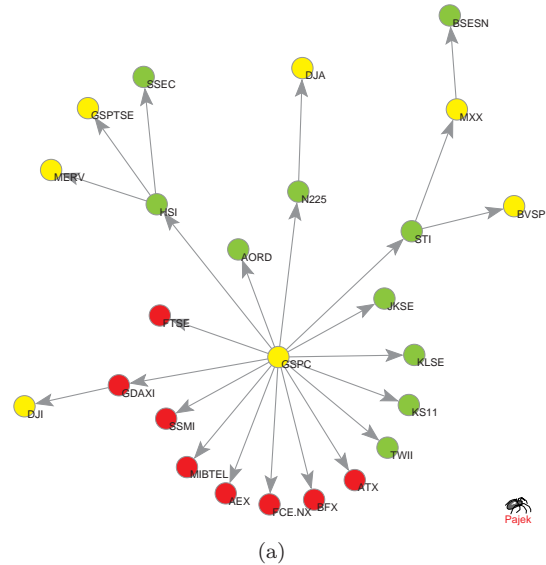
(a)
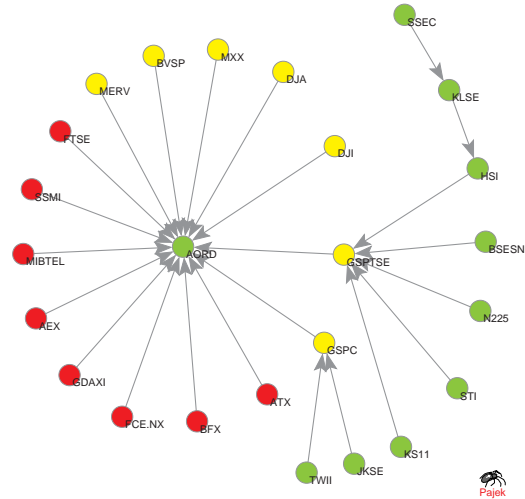
(b)

Fig. 4: Minimum spanning tree for (a) the outgoing transfer entropy and (b) the incoming transfer entropy. The minimum spanning tree is drawn by Pajek (http://vlado.fmf.uni-lj.si/pub/networks/pajek/).

To observe schematically the relation of information flow between the stock markets, the minimum spanning tree is depicted in Fig. 4. The colors of nodes indicate

the continent housing the markets. The minimum spanning tree is constructed by connecting the highest outgoing (Fig. 4(a)) and incoming (Fig. 4(b)) transfer entropy. In Fig. 4(a), the GSPC is located at the focus of the network, connected with the Asian and European markets. The GSPC plays the role of information source among stock markets. On the other hand, in Fig. 4(b) the Australian market, AORD, is located in the center of the minimum spanning tree for incoming transfer entropy.

Conclusions. – We observed the transfer entropy and the cross-correlation of the daily data of 25 stock markets to investigate the intensity and the direction of information flow between stock markets. The value of the outgoing transfer entropy from the American markets is high, as it is from the European markets. The information flow mainly streams to the Asia/Pacific region. However, the transfer entropy between markets on the same continent is not high. We also reveal that markets on the same continent have higher values of cross-correlation for log-return. This indicates that intra-continental market indices fluctuate simultaneously with common deriving mechanisms. The cross-correlation between America and Europe is relatively high. Therefore, we can see that American and European stock markets fluctuate in tune with common deriving mechanisms. However, the Asia/Pacific markets fluctuate simultaneously with other common deriving mechanisms. Consequently, we can predict that the deriving mechanism for America and Europe does not depend on the Asia/Pacific market situation but rather that Asia is affected by the American and European market situation. From Fig. 4, we can confirm the above description. The GSPC is located in the center of the star network, influencing all the other markets. Therefore, the Asian markets represent similar patterns of time series, although they do not actively interact with each other.

Transfer entropy is a proper measure for the investigation of the information flow between global stock markets. Through this measure we can quantify information transportation between underlying mechanisms that govern stock markets even though we cannot concretely identify the mechanism.

$$* * *$$

REFERENCES

[1] Eguiluz V. M. and Zimmermann M., *Phys. Rev. Lett.*, **85** (2000) 5659.
[2] Krawiecki A., Hołyst J. A., and Helbing D., *Phys. Rev. Lett.*, **89** (2002) 158701.
[3] Chowdhury D. and Stauffer D., *Eur. Phys. J. B*, **8** (1999) 477.
[4] Takaishi T., *Int. J. Mod. Phys. C*, **16** (2005) 1311.
[5] Kaizoji T., *Physica A*, **287** (2000) 493.
[6] Kaizoji T. and Kaizoji M., *Adv. Complex Syst.*, **6** (2003) 303.
[7] Palágyi Z. and Mantegna R. N., *Physica A*, **269** (1999) 132.
[8] Kaizoji T., Bornholdt S., and Fujiwara Y., *Physica A*, **316** (2002) 441.
[9] Jung W.-S., Chae S., Yang J.-S., and Moon H.-T., *Physica A*, **361** (2006) 263.
[10] Jung W.-S., Kwon O., Yang J.-S., and Moon H.-T., *J. Korean Phys. Soc.*, **48** (2006) S135.
[11] Kwon O., and Yang J.-S., *Physica A*, **doi:10.1016/j.physa.2008.01.007** (2008) .
[12] Arthur W. B., Durlauf S. N., and Lane D. A, *The Economy as an Evolving Complex System II* (Perseus Books, Jackson) 1997.
[13] Mantena R. N. and Stanley H. E., *An Introduction to Econophysics* (Cambridge University Press, Cambridge) 2000.
[14] Bouchaud J.-P. and Potters M., *Theory of Financial Risks* (Cambridge University Press, Cambridge) 2000.
[15] Mandelbrot B. B., *Quant. Finance*, **1** (2001) 124.
[16] Kullmann L., Kertész J., and Mantegna R. N., *Physica A*, **287** (2000) 412.
[17] Giada L. and Marsili M., *Physica A*, **315** (2002) 650.
[18] Park J. B., Lee J. W., Yang J.-S., Jo H.-H., and Moon H.-T., *Physica A*, **379** (2007) 179.
[19] Lee J. W., Park J. B., Jo H.-H., Yang J.-S., and Moon H.-T., *physics/0607282*, (2006) .
[20] Yang J.-S., Kwak W., Kaizoji T., and Kim I.-m., *physics/0701179*, (2007) .
[21] Kaizoji T., *Econophysics of Stock and other Markets: Proceedings of the Econophys-Kolkata II Series, New Economic Windows* (Springer) 2006, p. 3.
[22] Matal K., Pal M., Salunkay H., and Stanley H. E., *Europhys. Lett.*, **66** (2004) 909.
[23] Yang J.-S., Chae S., Jung W.-S., and Moon H.-T., *Physica A*, **363** (2006) 377.
[24] Silva A. C., Prange R. E., and Yakovenko V. M., *Physica A*, **344** (2004) 227.
[25] Stanley H. E., Amaral L. A. N., Gabaix X., Gopikrishnan P., and Plerou V., *Physica A*, **299** (2001) 1.
[26] McCauley J. L. and Gunaratne G. H., *Physica A*, **329** (2003) 178.
[27] Kim K. and Yoon S.-M., *Physica A*, **344** (2004) 272.
[28] Schreiber T., *Phys. Rev. Lett.*, **85** (2000) 461.
[29] Marschinski R. and Kantz H., *Eur. Phys. J. B*, **30** (2002) 275.
[30] Baek S. K., Jung W.-S., Kwon O., and Moon H.-T., *physics/0509014*, (2005) .

epl draft

Information flow between stock indices

OKYU KWON[1,2] and JAE-SUK YANG[1] (a)

[1] *Department of Physics, Korea University, Seoul 136-701, Korea*
[2] *National Creative Research Initiative Center for Neuro-dynamics, Korea University, Seoul 136-701, Korea*

Abstract. - Using transfer entropy, we observed the strength and direction of information flow between stock indices. We uncovered that the biggest source of information flow is America. In contrast, the Asia/Pacific region the biggest is receives the most information. According to the minimum spanning tree, the GSPC is located at the focal point of the information source for world stock markets.

Introduction. – Economic systems have recently become an active field of research for physicists striving to transfer concepts and methodologies from statistical physics such as phase transition, fractal theory, spin models, complex networks, and information theory to the analysis of economic problems [?,?,?,?,?,?,?,?,?,?,?,?,?,?, ?,?,?,?,?,?,?,?,?,?].

Among the numerous methodologies put forward, time series analysis has proven to be one of the most efficient methods and is widely applied to the examination of characteristics of stock and foreign exchange markets. In order to analyze financial time series, a range of statistical measures have been introduced, including probability distribution [?,?,?,?,?,?,?], autocorrelation [?], multi-fractal [?], complexity [?,?,?], entropy density [?,?], and transfer entropy [?].

Information is a keyword in analyzing financial market data or in estimating the stock price of a given company. It is quantified through a variety of methods such as cross-correlation, autocorrelation, and complexity. However, while they may be appropriate measures for the observation of the internal structure of information flow, they fail to illuminate the directionality of information flow. Schreiber [?] introduced transfer entropy, which measures the dependency in time between two variables and notes the directionality of information flow. This concept of transfer entropy has already been applied to the analysis of financial time series. Marschinski and Kantz [?] calculated information flow between the Dow Jones and DAX stock indexes to better observe interactions between the two huge markets. Kwon and Yang [?] measured the direction of information flow between the composite stock index and individual stock prices, while the information flow among individual stocks in a stock market has been estimated by Baek and coauthors [?] in order to measure the internal structure of a stock market.

Through transfer entropy, this paper focuses quantitatively on the direction of information flow between 25 stock markets to determine which market serves as a source of information for global stock indices. Drawn from the economic system, a plethora of empirical data reflecting economic conditions can be obtained. The time series of a composite stock price index provides prime data accurately reflecting economic conditions. Therefore, we analyzed the daily time series of the 25 stock indices listed in Table **??** for the period of 2000 – 2007 using transfer entropy in order to examine the information flow between stock markets and identify the hub.

Transfer entropy. – Transfer entropy, which measures the directionality of a variable with respect to time was recently introduced by Schreiber [?] based on the probability density function (PDF). Let us consider two discrete and stationary process, I and J. Transfer entropy relates k previous samples of process I and l previous samples of process J and is defined as follows:

$$T_{J \to I} = \sum p(i_{t+1}, i_t^{(k)}, j_t^{(l)}) \log \frac{p(i_{t+1} \mid i_t^{(k)}, j_t^{(l)})}{p(i_{t+1} \mid i_t^{(k)})}, \quad (1)$$

where i_t and j_t represent the discrete states at time t of I and J, respectively. $i_t^{(k)}$ and $j_t^{(l)}$ denote k and l dimensional delay vectors of two time consequences I and J,

(a)Corresponding author.