Comprehensive Analysis of Information Transmission among Agents: Similarity and Heterogeneity of Collective Behavior

Aki-Hiro Sato

Abstract This contribution considers comprehensive methods to characterize market states based on high-resolution data of financial markets protecting the anonymity of agents. Specifically ways to estimate diversification of agents in the financial markets and to evaluate similarity of market states between different observation periods are proposed. The bipartite network representation consisting of participants and groups where they exchange one kind of financial commodity is considered. The bipartite network modeling provides us with an idea that the relative occurrence rates of activities, which computed from the number of participant's actions at each group, may be fingerprints of agents' collective behavior during the observation period. I propose the method to evaluate similarity between market states based on Jensen-Shannon divergence between their relative occurrence rates. I calculated the similarities of participant's activities on different observations in the foreign exchange market with high-resolution data. As a result it is found that the timings when the drastic change took place are related to economic events with a large impact.

1 Introduction

The recent world economy tightly couples and information about international trades is transmitted across the globe with light speed. Moreover due to development and spread of Information and Communication Technology (ICT) a larger amount of data on human activities than human cognitive capacity is accumulated in digital storage devices every day and everywhere. As a result the conditions that one can detect and trace flows of money, substances, information, and energy across the world is being put into place. Such technological advantages provide us with abilities to pay attention to the future of our world based on high-resolution data.

Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, JAPAN, e-mail: aki@i.kyoto-u.ac.jp

Furthermore we may be able to manage our global society based on high-resolution and comprehensive analysis on states of our world.

Specifically financial markets over the world can be observed and collective behavior of participants in financial markets is also becoming detectable through high-frequency financial data [1, 2]. Several researchers in econophysics, econometrics, and finance pay significant attention to fat-tailsness in probability density functions for log returns of financial commodities [3, 4]. A relationship between trades and price movements has been considered in the literature of various concepts of time scales. In the early 1960s, Allais proposed the concept of psychological time [5]. Mandelbrot and Taylor suggested that cumulative transaction volume gives a new time scale (the transaction clock) [6]. Clark took a similar approach and proposed that the cumulative transaction volume or cumulative number of transactions can be one of subordinators of Wiener processes [7].

More recently in order to understand the price movements in the financial markets more precisely, several researchers noticed an importance of understanding of quotation behavior as precursor of transactions and studied on price formation mechanisms through high-resolution data of financial markets [8, 9, 10].

The comprehensively investigation on quotation/transaction activities provides us with a whole picture of financial markets. The interesting characteristics could be found and physical methodologies would be useful to do so.

The aim of the present contribution is to propose methods to quantify and visualize attentions of market participants to financial commodities with protecting the anonymity of agents. Specifically in this contribution I focus on the foreign exchange market and attempt to comprehensively visualize market states of the foreign market with high-resolution data recorded in an electronic broking system.

This contribution is organized as follows. In Sec. 2 I propose a model of a financial market where N kinds of commodities are traded by M participants and methods to capture states of market participants in a practical manner. In Sec. 3 I show results of empirical analysis with high-resolution data of the foreign exchange market. Sec. 4 is devoted to conclusions.

2 Methods

Consider networks consisting of nodes between which constituents (money, substances, information, energy, and so on) convey. We assume that constituents are transmitted from one node to the other. If a constituent is exchanged between two nodes then we suppose that these nodes should be connected. Menezes and Barabási propose the method to quantify centrality of nodes for such a network from multiple time series of constituents' flows [11]. This estimator originates with random walkers on a network proposed by Noh and Rieger [12]. The centrality is based on stationary probability distributions for constituents to exist at each node. If many constituents exist at a node then it should seem to be an important node. Therefore the existence probability distribution of constituents provides us with information on the ratio of degree to its entity. It may be related to a relative importance of nodes in the network.

Let us consider a financial market where N kinds of financial commodities are traded by M participants. It can be described as a bipartite graph shown in Fig. 1. If one observes a quotation or transaction from the j-th participant to the i-th group then one draws a link between the j-th participant and the i-th group. Then the degree centrality of the i-th group can be calculated by the ratio of the number of links which the i-th group possesses to the total number of links over the groups. According to [11] since the degree centrality is equivalent to the existence probability distribution, the patterns of degree centrality seems to be a characteristic property to express market states from comprehensive point of view.

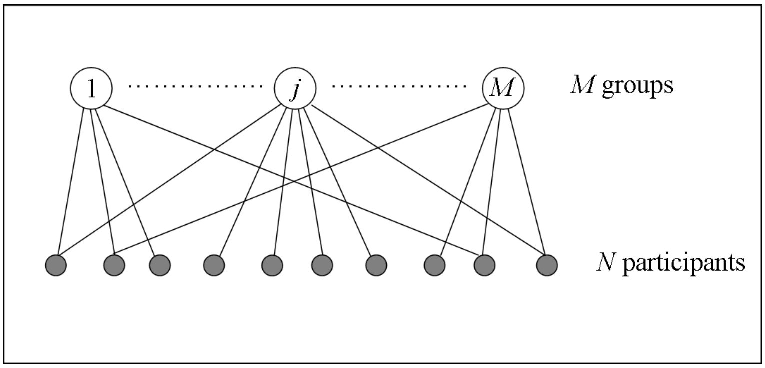


Fig. 1 A bipartite network description of a financial market. N financial commodities (groups) are traded by M market participants.

This quantity can be approximated by relative occurrence rates of quotations or transactions for currency pairs. We consider the relative occurrence rates of quotations/transactions and apply the concept of the degree centrality to estimating the relative importance of currency pairs and currencies.

Suppose one can observe quotations/transactions about the i-th currency and the j-th currency and count their arrival of quotations or occurrence of transactions for each currency pair on a brokerage system with an interval of $\Delta(>0)$. The quotation/transaction activity is defined as the number of quotations/transactions which market participants enter into the electronic broking system per Δ. We define $f_{ij}(t;S)$ as the quotation's activities between the i-th currency and j-th currency $(i,j=1,2,\ldots,N)$ in $[t\Delta,(t+1)\Delta] \quad (t=1,2,\ldots,T)$ on the S-th observation period. In this analysis we adopt the definition that the activities should be counted in symmetric way $f_{ij}(t;S)=f_{ji}(t;S)$ and satisfy no self-dealing condition $f_{ii}(t;S)=0$. Then the density of quotations between the i-th currency and j-th currency can be estimated as

$$A_{ij}(S)=\frac{\sum_{t=1}^{T} f_{ij}(t;S)}{\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}(t;S)}. \quad (1)$$

Obviously it has $\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(S) = 1$, $A_{ii}(S) = 0$, and $0 \le A_{ij}(S) \le 1$.

Under the assumption that the attention of market participants to the exchangeable currency pairs can be estimated as the centrality of currency pairs, $A_{ij}(S)$ can be empirically estimated by using quotation/transaction frequencies extracted from high-resolution data without knowledge on network structure of market participants. The reason why the centrality is adopted in order to quantify the attention of market participant is because currency pairs (currencies) which are quoted/traded by many participants are focused by many participants. Moreover relative occurrence rates of the i-th currency on the S-th observation period are defined as

$$K_i(S) = \sum_{j=1}^{N} A_{ij}(S), \tag{2}$$

where it has $\sum_{i=1}^{N} K_i(S) = 1$ and $0 \le K_i \le 1$.

Since both $A_{ij}(S)$ and $K_i(S)$ may be regarded as fingerprints representing the market states on the observation period S, their shape may describe market states at S. Furthermore since they are probability distributions from their definition, the similarity between them can be evaluate by means of several kinds of information-theoretical divergences.

In order to estimate a total diversification of quotation/transaction activities in a financial market one can adopt the normalized Shannon entropy of the centralities [16] for currencies/currency-pairs defined as

$$H_{cp}(S) = -\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(S) \log A_{ij}(S)}{\log N(N-1)}. \tag{3}$$

$$H_c(S) = -\frac{\sum_{j=1}^{N} K_j(S) \log K_j(S)}{\log N}, \tag{4}$$

If one currency-pair/currency is only traded then $H_{cp}(S)/H_c(S)$ takes the minimum value 0. Contrarily every currency/currency pair is equivalently traded then $H_{cp}(S)/H_c(S)$ takes the maximum value 1. Therefore $H_{cp}(S)/H_c(S)$ can be one of candidates to measure monopolization for financial commodities in the market.

Furthermore in order to compare the shape of probability distribution we can choose a divergence from f-divergence, Kullback-Leibler divergence, Jensen-Shannon divergence, and so on [17]. In the case of the f-divergence the similarity between two observation periods is defined as follow. Let $f(u)$ be a convex function satisfying $f(1) = 0$. Then the similarity between market states on the S_1-th observation period and those on the S_2-th observation period is defined as

$$D_A(S_1,S_2) = \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(S_1) f\Big(\frac{A_{ij}(S_2)}{A_{ij}(S_1)}\Big), \tag{5}$$

$$D_K(S_1,S_2) = \sum_{i=1}^{N} K_i(S_1) f\Big(\frac{K_i(S_2)}{K_i(S_1)}\Big). \tag{6}$$

They have the following properties

$$D_A(S_1,S_2) \geq 0, \tag{7}$$
$$D_A(S_1,S_2) = 0, \quad \text{iff} \quad A_{ij}(S_1) = A_{ij}(S_2), \tag{8}$$
$$D_K(S_1,S_2) \geq 0, \tag{9}$$
$$D_K(S_1,S_2) = 0, \quad \text{iff} \quad K_i(S_1) = K_i(S_2). \tag{10}$$

If we choose $f(u) = -\log u$, then they give the Kullback-Leibler divergence

$$D_A(S_1,S_2) = \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(S_1)\log\frac{A_{ij}(S_1)}{A_{ij}(S_2)}, \tag{11}$$
$$D_K(S_1,S_2) = \sum_{i=1}^{N} K_i(S_1)\log\frac{K_i(S_1)}{K_i(S_2)}. \tag{12}$$

As an alternative symmetric divergence the Jensen-Shannon divergence was introduced [18]. They are defined as

$$D_A(S_1,S_2) = H_A\Big(\frac{1}{2}\sum_{k=1}^{2} A_{ij}(S_k)\Big) - \frac{1}{2}\sum_{k=1}^{2} H_A\Big(A_{ij}(S_k)\Big), \tag{13}$$
$$D_K(S_1,S_2) = H_K\Big(\frac{1}{2}\sum_{k=1}^{2} K_i(S_k)\Big) - \frac{1}{2}\sum_{k=1}^{2} H_K\Big(K_i(S_k)\Big), \tag{14}$$

where $H_A(A_{ij}) and H_K(K_i)$ respectively denote the Shannon entropies defined as

$$H_A(A_{ij}) = -\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}\log A_{ij} \tag{15}$$
$$H_K(K_i) = -\sum_{i=1}^{N} K_i\log K_i. \tag{16}$$

From the definitions they have the following properties

$$D_A(S_1,S_2) = D_A(S_2,S_1) \tag{17}$$
$$D_A(S_1,S_2) \geq 0, \tag{18}$$
$$D_A(S_1,S_2) = 0 \quad \text{iff} \quad A_{ij}(S_1) = A_{ij}(S_2), \tag{19}$$
$$D_K(S_1,S_2) = D_K(S_2,S_1) \tag{20}$$
$$D_K(S_1,S_2) \geq 0, \tag{21}$$
$$D_K(S_1,S_2) = 0 \quad \text{iff} \quad K_i(S_1) = K_i(S_2), \tag{22}$$

3 Results

The analysis is conducted by using high-resolution data collected by ICAP EBS plathome (ICAP EBS Data Mine Level 1.0) [19]. In the exchangeable currency pairs consisting of 24 currencies and 5 commodities 47 kinds of currencies pairs are included in the data set during a period from May 2008 to June 2009 [1]. The relative occurrence rates of quotations and transactions are counted for each week.

Fig. 2 shows the normalized Shannon entropies of quotation activities for currency pairs and currencies, and of transaction activities for currency pairs and currencies for each week. The normalized Shannon entropies obtained from the number of quotations changed around the week beginning from 21st July 2008, from 15th December 2008, 23rd February 2009. The normalized Shannon entropies computed from the transactions changed around the week beginning from 10th November 2008, from 22th December 2008, 9th March 2009.

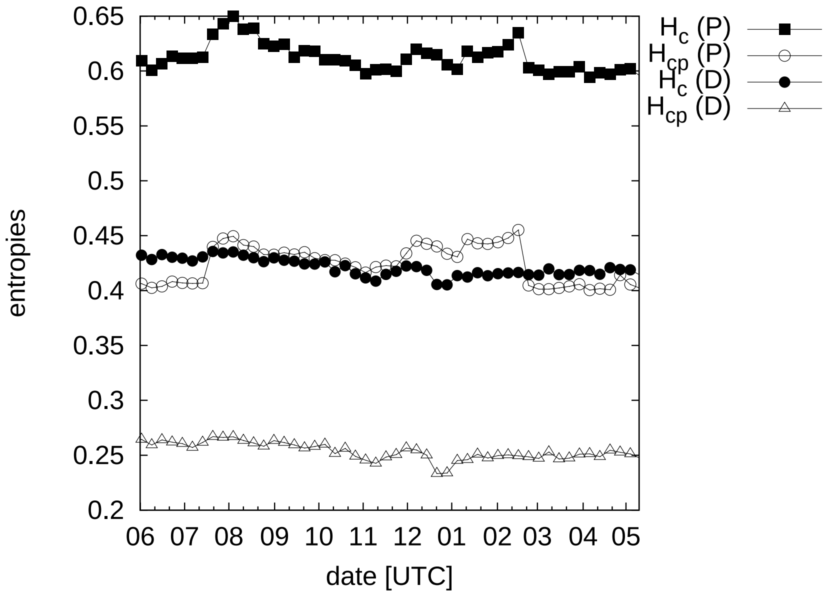
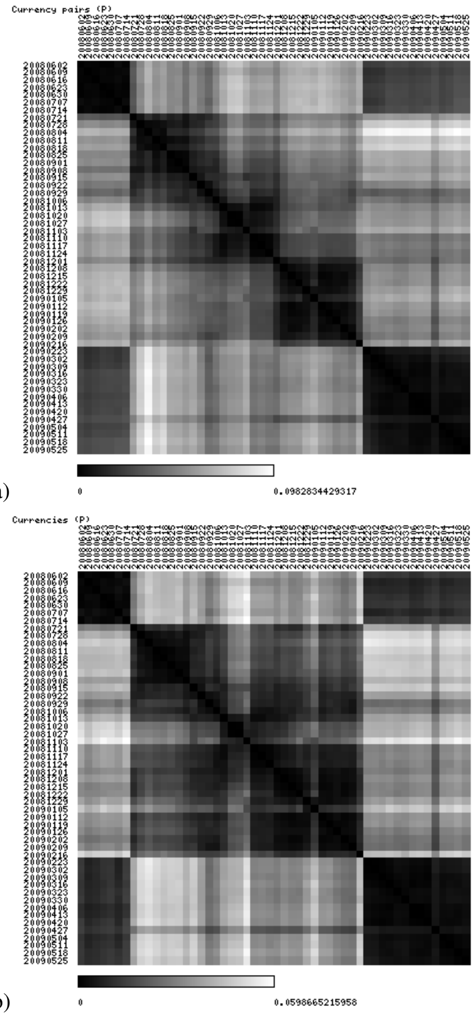


Fig. 2 The normalized Shannon entropies of quotation activities for currency pairs ($H_{cp}(P)$) and currencies ($H_c(P)$), and of transaction activities for currency pairs ($H_{cp}(D)$) and currencies ($H_c(D)$).

[1] The analyzed data include quotations or transactions for USD/CHF, EUR/USD, USD/JPY, EUR/JPY, EUR/CHF, NZD/USD, AUD/USD, GBP/USD, USD/CAD, AUD/NZD, EUR/GBP, XAU/USD, XAG/USD, EUR/SEK, CHF/JPY, XPD/USD, XPT/USD, USD/SGD, USD/HKD, EUR/NOK, USD/ZAR, USD/MXN, EUR/DKK, EUR/CZK, EUR/PLN, EUR/HUF, EUR/ISK, USD/RUB, GBP/JPY, EUR/SKK, USD/PLN, GBP/CHF, AUD/JPY, USD/TRY, USD/THB, NZD/JPY, CAD/JPY, ZAR/JPY, EUR/ZAR, EUR/RUB, EUR/CAD, GBP/AUD, USD/SEK, EUR/AUD, EUR/RON, BKT/RUB, and USD/NOK.

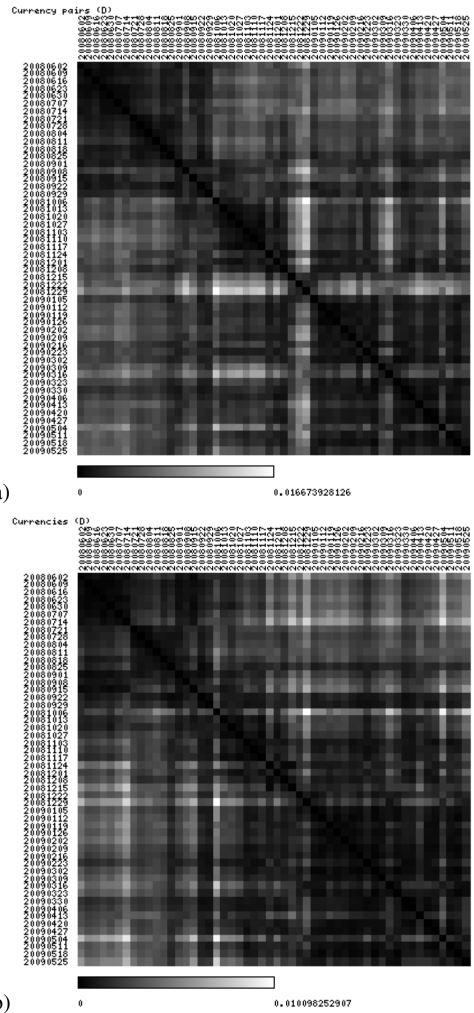
The similarity of market states between two periods is computed for each week. The Jensen-Shannon divergence is employed in order to compute similarities since it has resistance characteristic for zero probabilities.

Figs. 3 and 4 show similarities obtained from A_{ij} of quotations (a), K_i of quotations (b) and similarities obtained from A_{ij} of transactions (a), and K_i of transactions (b), respectively.



(a)
(b)

Fig. 3 The similarities of market states between observation periods based on activities of currency pairs (a) and of currencies (b) computed from quotations. The black pixel represents a similar relation, and the white pixel represents a dissimilar relation.



(a)

(b)

Fig. 4 The similarities of market states between observation periods based on activities of currency pairs (a) and of currencies (b) computed from transactions. The black pixel represents a similar relation, and the white pixel represents a dissimilar relation.

On the basis of quotation activities it is found that the transaction activities of currency pairs show drastic changes around the week beginning from 21st July 2008, from 29th September 2008, from 1st December 2008, and 16th February 2009. Furthermore it is found that the states during a period from 14th July 2008 to 23rd February 2009 are different from other periods. This period corresponds to the period when the normalized Shannon entropies for currency pairs and currencies remained larger values than before 14th July 2008 and after 23rd February 2009 (see

Fig. 2). It is confirmed that the states before 14th July 2008 and those after 23rd February 2009 were very similar from Fig. 3.

The transaction activities for currency pairs slightly differ from their quotation activities. From Fig. 4 the situations of the week beginning from 8th September 2008, the weeks from 15th December 2008 to 29th December 2008, the weeks from 9th March 2009 to 23rd March 2009, and the week beginning from 4th May 2009 are slightly different from other weeks. Specifically it is found that the transaction activities show drastic changes around the week beginning from 29th September 2008, from 25th December 2008, from 9th March 2009, and from 4th May 2009. The transaction activities for currencies show drastic changes at the week beginning from 6th October 2008, the week 24th November 2008, the weeks from 15th December 2008 to 29th December 2008, and the week 4th May 2009. It is found that these periods are confirmed from the normalized Shannon entropies for transaction activities on the basis of the changes of $H_{cp}(D)$ and $H_c(D)$ as shown in Fig. 2.

These peculiar periods detected throughout this empirical analysis correspond to the beginning of financial crisis, stimuli by emergency economic package of G20 countries, and crisis in March 2009, respectively.

4 Conclusions

I proposed the methods to estimate states of market participants from comprehensive point of view. The uncertainty of quotation/transaction activities based on the normalized Shannon entropy of the relative occurrence rates and similarity of market states between two observation periods based on their Jensen-Shannon divergence are proposed, respectively. By means of the proposed methods I empirically analyzed market states of the foreign exchange market with high-resolution data. It is found that the timings when the drastic change took place are related to economic events with a large impact. Finally I expect the direction of my study to contribute for understanding environmental situations from comprehensive point of view.

Acknowledgement

This work was supported by the Grant-in-Aid for Young Scientists (B) (# 21760059) from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT).

References

1. V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Nunes, N. Amaral and H.E. Stanley, "Universal and Nonuniversal properties of cross correlations in financial time series", Physical Review Letters, 83 (1999) pp. 1471–1474.
2. A.-H. Sato, "Application of spectral methods for high-frequency financial data to quantifying states of market participants", Physica A, 387 (2008) pp. 3960–3966.
3. R. Mantegna and H.E. Stanley, "An Introduction to Econophysics Correlations and Complexity in Finance" (Cambridge University Press, Cambridge, 2000).
4. T. Mizuno, S. Kurihara, M. Takayasu, and H. Takayasu, "Analysis of high-resolution foreign exchange data of USD-JPY for 13 years", Physica A, 324 (2003) pp. 296–302.
5. M. Allais, "The psychological rate of interest", Journal of Money, Credit and Banking, 3 (1974) pp. 285–331.
6. B.B. Mandelbrot and H.M. Taylor, "On the distribution of stock prices differences", Operations Research, 15 (1967) pp. 1057–1062.
7. P.K. Clark, "A subordinated stochastic process model with finite variance for speculative prices", 15 (1973) pp. 15–36.
8. M. Melvin, and X. Yin, "Public information arrival, exchange rate volatility, and quote frequency", The Economic Journal, 110 (2000) pp. 644–661.
9. D. Challet and R. Stinchcombe, "Limit order market analysis and modelling: on a universal cause for over-diffusive prices", Physica A, 324 (2003) pp. 141–145.
10. A.-H. Sato, "Characteristic time scales of tick quotes on foreign currency markets: an empirical study and agent-based model", The European Physical Journal B, 50 (2006) pp. 137–140.
11. M.A. de Menezes, and A.-L. Barabá si, "Separating internal and external dynamics of complex systems", Physical Review Letters, 93 (2004) 068701.
12. J.D. Noh and H. Rieger, "Random walks on complex networks", Physical Review Letters, 92 (2004) 118701.
13. Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity in 2007: http://www.bis.org/publ/rpfxf07t.pdf
14. T. Aste, and T.D. Matteo, "Dynamical networks from correlations", Physica A, 370 (2006) pp. 156–161.
15. M.J. Naylor, L.C. Rose, B.J. Moyle, "Topology of foreign exchange markets using hierarchical structure methods", Physica A, 382 (2006) pp. 199-208.
16. A.-H. Sato, "Detecting environmental changes through high-resolution data of financial markets", New Advances in Intelligent Decision Technologies: Results of the First KES International Symposium IDT'09 (Studies in Computational Intelligence), Springer(Berlin), Ed. by K. Nakamatsu, G. Phillips-Wren, L.C. Jain, and R.J. Howlett, (2009), pp. 595–603.
17. S. Amari, H. Nagaoka, "Methods of Information Geometry" (AMS and Oxford University Press, Oxford, 2000).
18. J. Lin, "Divergence Measures Based on the Shannon Entropy", *IEEE transactions on information theory*, Vol. 37, No. 1, pp. 145–150.
19. The data is purchased by ICAP EBS: *http://www.icap.com*