

# 金融市場の網羅的高精度データを用いた市場状態の定量化 —エージェント集団行動の偏りと類似性の観点から—

## Quantifying market states with comprehensive high-resolution data of financial markets: From viewpoint on diversification and similarity of agents' collective behavior

佐藤彰洋<sup>1\*</sup>

<sup>1</sup> 京都大学大学院情報学研究科数理工学専攻

<sup>1</sup> Department of Applied Mathematics and Physics,  
Graduate School of Informatics, Kyoto University

**Abstract:** This study proposes methods to quantify states of market participants with comprehensive high-resolution data of financial markets. The number of quotations and transactions are focused from information transmission perspectives. The proposed methods are demonstrated by using comprehensive high-resolution data of the foreign exchange market. It is found that the states of market participants drastically changed around recent financial crisis which had happened in the middle of 2008.

### 1 はじめに

近年の情報伝達技術の発展と、それに伴う社会システムの電子化により世界規模で社会状態の観測が可能となってきた。しかしながら、一個人の内部状態に比して、その総体以上である我々の社会の状態数ははるかに多く、複数の人間が協力したとしても、その全貌を理解することが不可能(不可知)であるという意味で、我々の社会は高い複雑性を有する。そのため、我々が社会状態を理解するためには、大自由度の社会システムを、一個人の処理能力の範疇におさまる程度の情報量に縮約する方法論が必要である。

例えば、多くの社内データベースには、取引関係や、商品購買に関する記録等がテキストデータなどのデジタルデータとして保管されている。そして、それらのデジタルデータの規模は近年増大傾向にある。しかしながら、膨大なレコード数のテキストデータや大量の画像データを、そのまま人間が人力でながめたとしても、そこに意味や傾向を見出すことは不可能(不可知)である。

そのため、何らかの数理的、または、アルゴリズム的なルールにより大量のテキストデータから傾向や特徴を抽出し、我々が理解できる程度の情報量にする必要がある。ここに、大容量の経済・社会データに対して知識処理的あるいは、大量の計算を通じた計算科学的な考え方

を取り入れる意義と必要性が見出される。

更に、これまで不可知である状況を、もし、技術的先進性により可知とすることが可能であるならば、その可知性は我々に予見性や先行性を与えるため、これまで不可知性から隠されてきた裁定機会の発見や、不可知性によりもたらされてきた様々な不幸的状况の改善をもたらす可能性ある。

本報告では、金融市場の状態を網羅的観点から計量し、市場の状態を定量化することにより、網羅的観点から金融市場状態の変化をとらえる方法論について議論する。

### 2 金融市場の高精度データ

近年の電子取引システムの発展と普及の結果、注文や取引単位の金融市場のデータを入手することが可能になりつつある。

例えば、ロンドン証券取引所(London Stock Exchange) [1] は、Rebuild Order Book と呼ばれるヒストリカルデータを販売している。このデータには、1ms の時間解像度で収集された注文と約定に関する全板データが含まれている。また、ニューヨーク証券取引所(New York Stock Exchange) [2] で商品化されている NYSE Trades and Quotes(TAQ) や Orders/Depth-of-Book を利用することで、どのようなタイミングで注文と取引が生起しているかを高精度でとらえることができる。一方、CQG社 [3] の CQG's Data Factory を通じ、Time & Sales

\*連絡先：京都大学大学院情報学研究科数理工学専攻  
〒606-8501 京都市左京区吉田本町  
E-mail: aki@i.kyoto-u.ac.jp

データやヒストリカル DOM データを購入することが可能となっており、これらのデータを用いることにより網羅的な高精度データの収集が可能となりつつある。CQG 社では、CQG API と呼ばれる API サービスを提供しており、リアルタイムでの市場データの収集が可能である。一方、外国為替市場のデータとして、ICAP EBS 社 [4] が販売している EBS Data Mine では、高精度の外国為替市場の注文と取引に関するデータを利用することができる。

### 3 本研究のねらい

金融市場は一般に複数の商品が同時に取引されており、いくつかの商品間では有意の価格変化の傾向が似ていることが知られている。この相関傾向を利用することで、ポートフォリオを構築し、リスク資産総額の価格変動を軽減することが行われている。

一方、世界中の金融市場で取引されている証券、通貨、債権や商品の総数は膨大であり、時々刻々変化するリスク資産の価格変化の全体像を把握することは、人間の情報処理能力の上限を越えており、極めて困難な問題である。特に、現在の市場がどのような状態にあるかを知ることが、価格の変化が生じることをみこして行動を取るために有益な情報を与えると思われる。

更に、近年の研究から金融市場で生じる商品価格の変化は完全な確率事象ではなく、その背後にある市場への情報到着と深く関係していることが認識されつつある。特に、商品価格の変化と注文回数、取引量などの変数の間に強い相関があることが確認されている。例えば、Clark は市場価格の変化と取引量との間には強い相関関係があることを示し、対数収益率の価格変動に見られる裾野の厚みを、取引回数を考慮した劣後化過程 (subordination process) を用いることで説明する仮説を提案している [5]。特に取引量や約定回数は、情報到着の代理変数とみなされ、情報到着を媒介として価格変化やボラティリティクラスタリングをモデル化し、市場価格変化と取引量や約定回数との関係を定量分析する研究が行われている [5, 6, 7]。

本研究では、市場状態を、高解像度データでは計量が難しい市場価格の変化ではなく、注文回数と約定回数の市場全体での偏りから網羅的観点から計量する方法を議論する。そして、市場状態が市場への情報到着により、どれくらい偏りを生じ、またこれまでの状況とどれほど乖離しているかの類似性を定量化する方法を提案する。

### 4 分析方法

本節では、金融市場から得られる高精度データ全体を用いることにより、市場の動向を定量化する方法を考察する。

Noh と Rieger はネットワーク上をランダムウォークするモデルを考察し、ランダムウォークする粒子のノード上に存在する定常確率は、ノードの degree centrality と一致することを示した [8]。更に、Menezes と Barabási は、担架体がネットワーク上を移動する様々なシステム (交通流、コンピュータネットワーク、電子回路網) の実証分析を行い、各ノードで観測される流量から計算されるノードごとの担架体の相対存在割合とシステムの挙動との関係を調査している [9, 10]。これらの先行研究から、ネットワーク構造とネットワーク上を流れる担架体の存在確率との間には強い関連があることが示されている。更に、担架体の存在確率を知ることができれば、詳細なネットワーク構造を知らなくともシステムの挙動を巨視的観点から定量化することができることが示唆されている。

一方、複数の商品が取引される金融市場は、図 1 に示すように、複数の市場参加者 (participants) と取引される商品を媒介として形成される複数の場 (groups) から構成される 2 部グラフ構造を有する。

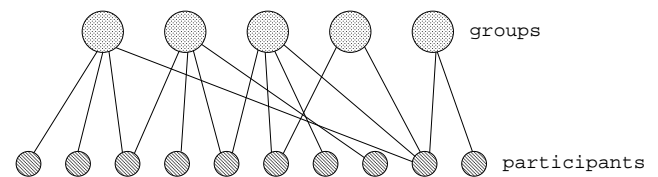


図 1: 金融市場の二部グラフ表現。

ここで高精度データで記録されているデータは各場への注文の生起と取引 (約定) の発生である。各場での注文あるいは取引の発生回数から計算される相対頻度は、各場の注文の生れやすさを計量していることから、ある種の確率を推定していると考えられる。確率の大きい商品は多くの参加者を引き付ける魅力が高く、一方、確率が小さい商品は魅力が低いことに対応する。Noh と Rieger の方法論 [8] を 2 部グラフにあてはめると、注文あるいは取引の発生回数から計算される相対頻度はある場に参加する参加者の割合とおおよそ対応している。

$N$  種類の商品  $i$  ( $i = 1, \dots, N$ ) が取引されている金融市場を考える。観測期間  $T : [t - \Delta, t]$  ( $\Delta > 0$ ) に観測された総注文回数と総取引回数をそれぞれ、 $f_i(T)$ 、 $g_i(T)$  と表記する。このとき、期間  $T$  における商品  $i$  の注文と取引の相対頻度はそれぞれ

$$p_i(T) = \frac{f_i(T)}{\sum_{i=1}^N f_i(T)}, \quad (1)$$

$$q_i(T) = \frac{g_i(T)}{\sum_{i=1}^N g_i(T)}, \quad (2)$$

により定義される。  $p_i(T)$  と  $q_i(T)$  は定義より、  $0 \leq p_i(T) \leq 1, 0 \leq q_i(T) \leq 1, \sum_{i=1}^N p_i(T) = 1, \sum_{i=1}^N q_i(T) = 1$  を満足する。

更に、担架体の相対頻度はネットワーク構造と関係しているため、相対頻度の分布形状はネットワーク構造を捉える一つの手掛かりとなりうる。もし、それぞれの場の相対頻度が一樣であるならば、各場に参加する参加者の割合は一樣である。他方、ある場の相対頻度が突出しているならば、その場に集中して参加者が存在することになる。

この参加者の分布の様子を情報理論の枠組により、  $p_i(T), q_i(T)$  の Shannon エントロピー

$$H_p(T) = - \sum_{i=1}^N p_i(T) \log p_i(T), \quad (3)$$

$$H_q(T) = - \sum_{i=1}^N q_i(T) \log q_i(T), \quad (4)$$

を用いて計量することで、市場全体にわたる商品取引の偏りを通じた観測期間毎に捉える方法を提案する。また、相対エントロピー  $KL(s_i, t_i) = \sum_{i=1}^N s_i \log \frac{s_i}{t_i}$  を用いることにより、

$$dpp(T_1, T_2) = KL(p_i(T_1), p_i(T_2)), \quad (5)$$

$$dpq(T_1, T_2) = KL(p_i(T_1), q_i(T_2)), \quad (6)$$

$$dqq(T_1, T_2) = KL(q_i(T_1), q_i(T_2)), \quad (7)$$

や Jensen-Shannon divergence  $JS(s_i, t_i) = H\left(\frac{s_i+t_i}{2}\right) - \frac{1}{2}\left(H(s_i) + H(t_i)\right)$  ( $H(\cdot)$  は Shannon エントロピー) を用いることにより、

$$d_{pp}(T_1, T_2) = JS(p_i(T_1), p_i(T_2)), \quad (8)$$

$$d_{pq}(T_1, T_2) = JS(p_i(T_1), q_i(T_2)), \quad (9)$$

$$d_{qq}(T_1, T_2) = JS(q_i(T_1), q_i(T_2)), \quad (10)$$

により期間間の分布形状の差異や注文と取引との間の差異を計量することが可能である。

## 5 外国為替市場における分析方法

ここでは外国為替市場の高精度データを用いて実際に計算を行うためのデータの形成方法について述べる。

期間  $[t\Delta, (t+1)\Delta]$  ( $t = 0, 1, \dots, T-1; \Delta > 0$ ) で  $M$  種類の通貨からなる通貨ペア  $j/k$  ( $j, k = 1, \dots, M$ ) に対して観測される注文回数/約定回数を  $f_{jk}(t)$  と表記する。

一般に  $j/k$  と記載した場合の売値 (Offer/Ask) と買値 (Bid) は、第一通貨  $j$ , 1 単位を第二通貨  $k$  で売り/買いつくときの第二通貨  $j$  換算での金額と定義される。例えば、USD/JPY の売値が 92.60、買値が 92.80 であるとは、1 米ドルを 92.60 円で売り、92.80 円で買うことを意味する。

外国為替市場では通貨取引は対称的であり、第一通貨を第二通貨で売り/買いつくことは第二通貨を第一通貨で買い/売りすることと等価である。本分析では、通貨ペア  $j/k$  の注文/約定と通貨ペア  $k/j$  の注文/約定を同一視し、その回数  $f_{jk}(t) = f_{kj}(t)$  なる定義を採用する。また、同一通貨間の交換は行われなことから、明らかに  $f_{jj}(t) = 0$  である。

更に、注文回数/約定回数から計算される通貨ペア  $(j, k)$  と通貨  $k$  の相対頻度を  $p_{jk}, q_k$

$$p_{jk} = \frac{\sum_{t=0}^{T-1} f_{jk}(t)}{\sum_{t=0}^{T-1} \sum_{j=1}^M \sum_{k=1}^M f_{jk}(t)}, \quad (11)$$

$$q_k = \sum_{j=1}^M p_{jk}, \quad (12)$$

と定義する。定義より、  $0 \leq p_{jk} \leq 1, 0 \leq q_k \leq 1$  であり、  $p_{jj} = 0, \sum_{j=1}^M \sum_{k=1}^M p_{jk} = 1, \sum_{k=1}^M q_k = 1$  が成立する。

更に、通貨ペアと通貨の注文回数/約定回数の偏りを規格化した平均情報量により定義する。

$$H_{cp} = - \frac{\sum_{j=1}^M \sum_{k=1}^M p_{jk} \log p_{jk}}{\log M(M-1)}, \quad (13)$$

$$H_c = - \frac{\sum_{k=1}^M q_k \log q_k}{\log M}. \quad (14)$$

定義により、  $0 \leq H_{cp} \leq 1, 0 \leq H_c \leq 1$  である。もし、単一の通貨ペアの取引あるいは通貨の取引が支配的である場合、  $H_{cp}$  と  $H_c$  は 0 に近い値をとる。反対に、複数の通貨ペアあるいは通貨の取引が等しい頻度で認められる場合、  $H_{cp}$  と  $H_c$  は 1 に近い値をとる。また、観測期間  $T_1$  と  $T_2$  の類似度は Jensen-Shannon divergence を用いて

$$JS_{cp}(T_1, T_2) = JS(p_{jk}(T_1), p_{jk}(T_2)), \quad (15)$$

$$(16)$$

$$JS_c(T_1, T_2) = JS(q_k(T_1), q_k(T_2)). \quad (17)$$

により定量化されうる。

## 6 外国為替市場における実証分析

本分析では、ICAP EBS Platform により収集された、Data Mine Level 1.0 を用いて分析を行った。ICAP

EBS Platform は世界中の約 800 のディーリンググループで約 2,000 名のトレーダに利用される電子ブローキングシステムである [4].

2008 年 6 月から 2009 年 7 月までの 13 ヶ月間のデータを用いて分析を行った. 分析期間中に 30 通貨と 5 種類の貴金属および 1 種類のバスケット取引<sup>1</sup>を含むことが認められた.

図. 2 に  $\Delta = 1$  [min] としたときの, USD/JPY に対する注文回数/約定回数の時系列を示す. 地球の自転に運動して日中の活動時間帯が変化することに伴う, 市場参加者の入れ替わりのため 24 時間の周期で注文回数/約定回数ともに周期的な変動パターンを示す. 更に, 取引全体では, 図. 3 のような注文回数/約定回数の空間的な変動が確認される.

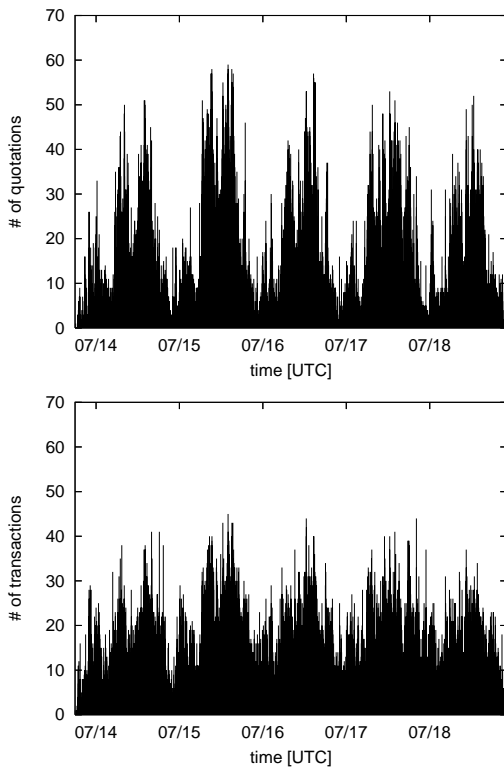


図 2: 2008 年 7 月 14 日から 7 月 20 日 (UTC) の USD/JPY の注文回数 (上) と約定回数 (下) の時系列.  $\Delta = 1$  [min] とした.

<sup>1</sup>観測期間中に以下の通貨を含む注文/約定の存在を確認した. AUD (Australia Dollar), CHF (Swiss Fran), EUR (EU Euro), GBP (British Sterling), NZD (New Zealand Dollar), USD (USA Dollar), JPY (Japanese Yen), CZK (Czech Koruna), DKK (Danish Krone), HUF (Hungarian Forint), ISK (Iceland Krona), NOK (Norwegian Krone), PLN (Poland Zloty), SEK (Swedish Krona), SKK (Slovak Koruna), ZAR (South African Rand), CAD (Canadian Dollar), HKD (Hong-Kong Dollar), MXN (Mexico Peso), RUB (Russian Ruble), SDG (Singapore Dollar), TRY (Turkish Lira), THB (Thai Baht), RON (Romania New Lei), BKT (USD+EUR/RUB Basket Trade), ILS (Israel New Shekels), XAG (Silver), XAU (Gold), SAU (Gold), XPD (Paradium), XPT (Platinum).

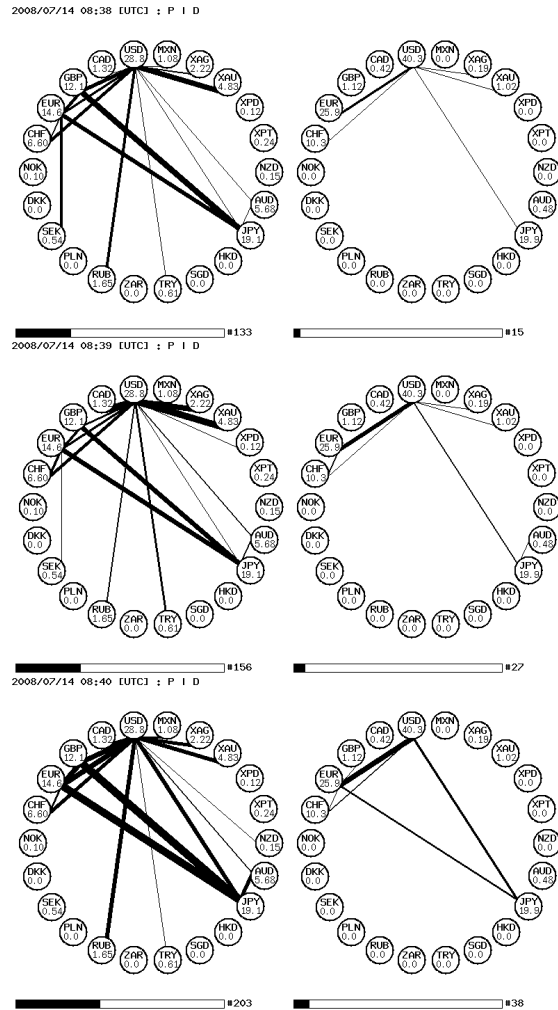


図 3: 2008 年 7 月 14 日 8:38-8:40(UTC) の注文回数 (左) と約定回数 (右) の変化のスナップショット. 通貨間を結ぶ線分の太さが注文回数/約定回数に対応している.

$T\Delta = 10,080$  [min] とし, 月曜日から日曜日までの 1 週間にわたる注文回数と約定回数から計算される通貨ペア/通貨の相対頻度  $p_{jk}/q_j$  から (13) 式と (14) 式を用いて, 規格化平均情報量の時間変化を計算した. 図. 4 に示すように, 2008 年 7 月 14 日から 20 日の週に, 注文回数から計算される相対頻度の規格化平均情報量が上昇し, その後, 高い値が持続することが確認される. 一方, 約定回数から計算される平均情報量の変化には, ほとんど違いは認められない. 更に, 各週ごとの差異を見るために, それぞれの週の相対頻度間の類似度を 0 耐性のある Jensen-Shannon ダイバージェンスを用いて計算した. 図 5 は通貨と通貨ペアの類似度を濃淡により図示したものである. この図から, 2008 年 7 月 14 日から 20 日の週にそれ以前と異なるパターンに突然変化し,

その後その傾向は 2009 年 2 月 16 日から 22 日の週末で続いていたことが読みとられる。

注文行動は市場参加者の意思の結果であるので、分析結果からこの時期に市場参加者の行動パターンに変化が生じたと考えることが可能である。しなしながら、約定回数の相対頻度に変化が見られないことから、約定の発生回数に注文行動の変化の影響はあまり生じていないと推察される。これは、約定発生は注文の適合が生じうる偶然性の影響が強いため、市場参加者の行動を直接反映していないことが原因かもしれない。

また、約定の発生頻度に大きな変化が生じていないことから、資金流通パターンに大きな変化は生じていないと予想される。このことを結論付けるためには、約定時の為替交換レートと約定枚数に着目した、取引金額に関する詳細な分析を行う必要がある。更に、そのような分析を通じて、世界規模での資金流量の様子をとらえることができるようになるかもしれない。

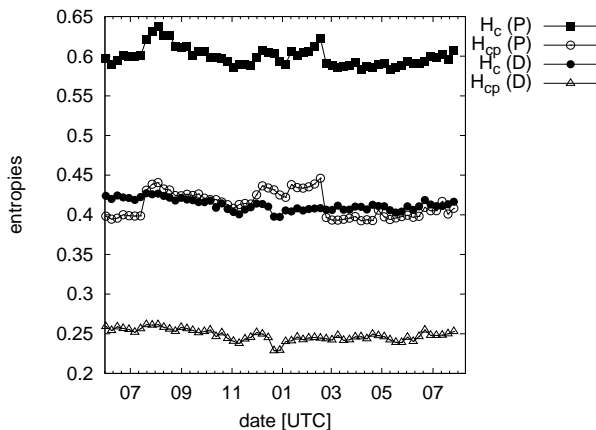


図 4: 注文回数と約定頻度の相対頻度から計算された規格化平均情報量の時間変化。 ( $M = 31$ ) 各相対頻度は 1 週間分の時系列 (10,080 分=7 日) を用いて計算した。

## 7 むすび

金融市場の状態を注文と約定の発生回数から計算される相対頻度を用いて網羅的な観点から計量し、市場状態の変調をとらえるための数理的枠組を提案した。そして、提案手法を用い、近年の外国為替市場の市場状態を定量化し、市場の変化の様子を観測した。本提案手法から 2008 年中盤から 2009 年前半に世界同時に発生した金融危機前後に、外国為替市場の通貨注文パターンと通貨取引パターンに大きな変化が生じていたことが確認された。

今後の課題としては、提案方法を複数の金融市場に適用し、長期間にわたり持続的に実証分析を行っていく必

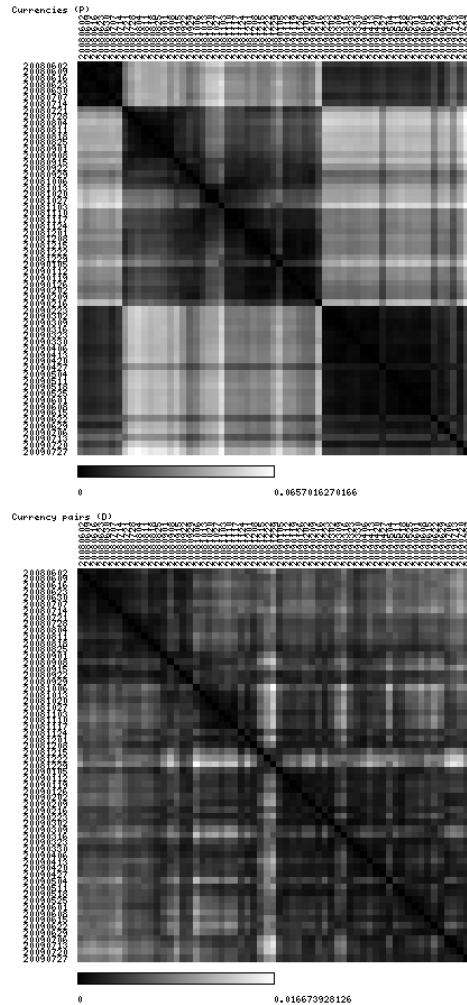


図 5: 注文回数と約定回数の相対頻度から計算された週ごとの通貨ペアの相対頻度の類似度 ( $M = 31$ ). 各相対頻度は 1 週間分の時系列 (10,080 分=7 日) を用いて計算した。

要がある。更に、定量結果と金融実務家の直感的理解との関連をフィールド研究を通じて実証的に調査していく必要がある。

## 謝辞

本研究は文部科学省科学研究費補助金 若手研究 (B) (21760059) の財政的援助を受けて実施された。

## 参考文献

- [1] ロンドン証券取引所:  
<http://www.londonstockexchange.com/>

- [2] ニューヨーク証券取引所: <http://www.nyse.com/>
- [3] CQG: <http://www.cqg.com/>
- [4] ICAP EBS: <http://www.icap.com>
- [5] P. Clark: *Econometrica*, Vol. 41, p. 135 (1973)
- [6] T. Tauchen, and M. Pitts: *Econometric Society*, Vol. 51, p. 485 (1983)
- [7] M. Richardson, and T. Smith: *Journal of Financial and Quantitative Analysis*, Vol. 29, p. 101 (1994)
- [8] J.D. Noh and H. Rieger: *Physical Review Letters*, Vol. 92, pp. 118701-1–118701-4 (2004)
- [9] M.A. de Menezes, A.-L. Barabási: *Physical Review Letters*, Vol. 92, pp. 028701-1–028701-4 (2004)
- [10] M.A. de Menezes, A.-L. Barabási: *Physical Review Letters*, Vol. 93, pp. 068701-1–068701-4 (2004)