How High Frequency Trading Affects a Market Index

Dror Y. Kenett[1,2,*], Eshel Ben-Jacob[1], H. Eugene Stanley[2], Gitit gur-Gershgoren[3,4]

[1] *School of Physics and Astronomy,*

The Raymond and Beverly Sackler Faculty of Exact Sciences,

Tel-Aviv University, Tel-Aviv 69978, Israel

[2] *Center for Polymer Studies and Department of Physics,*

Boston University, Boston, MA 02215, USA

[3] *Department of Economic Research,*

Israel Securities Authority, Jerusalem 95464, Israel and

[4]*Faculty of Business Administration,*

Ono Academic College, Kiryat Ono 55000, Israel

Supplementary Information A - Influence Analysis

To further study the relationship between stocks and an index, we use partial correlation influence analysis, introduced via dependency network methodology [1]. By studying variable partial correlations, we determine how each variable influences other system variables. A partial correlation is a parameter that indicates how a third mediating variable affects the correlation between two other variables [1–9]. The first step in dependency network analysis is constructing correlation and partial correlation matrices. The correlations between the time series of all variables in the dataset, are calculated using Pearson's formula [10]

$$C(i,j) = \frac{\langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle}{\sigma_i \sigma_j}. \tag{1}$$

Here X_i and X_j are the time series of variables i and j of length n, σ_i and σ_j are the standard deviation, and $\langle \cdots \rangle$ indicates an average. Note that the correlation for all pairs of variables defines a symmetric correlation matrix in which the (i,j) element is the correlation between variables i and j.

We next use the resulting correlation matrix to compute the partial correlations. The first-order partial correlation coefficient is a statistical measure indicating how a third variable affects the correlation between two other variables. The partial correlation between variable i and k with respect to a third variable j, $PC(i,k|j)$, is defined

$$PC(i,k|j) = \frac{C(i,k) - C(i,j)C(k,j)}{\sqrt{(1 - C^2(i,j))(1 - C^2(k,j))}}, \tag{2}$$

where $C(i,j)$, $C(i,k)$, and $C(j,k)$ are the correlations defined above. The relative effect of variable j on the correlation $C(i,k)$, is given by

$$d(i,k|j) \equiv C(i,k) - PC(i,k|j). \tag{3}$$

This transformation avoids the trivial case where variable j appears to strongly affect correlation $C(i,k)$, primarily because the values of $C(i,j)$, $C(i,k)$, and $C(j,k)$ are small. We note that this quantity can be viewed either as the correlation dependency of $C(i,k)$ on variable j (the term used here) or as the correlation influence of variable j on the correlation $C(i,k)$. We next define the total influence of variable j on variable i, or the dependency $D(i,j)$ of variable i on variable j to be

$$D(i,j) \equiv \frac{1}{N-1} \sum_{k \neq j}^{N-1} d(i,k|j). \tag{4}$$

As defined, $D(i,j)$ is a measure of the average influence of variable j on correlations $C(i,k)$ over all variables k not equal to j. All variable dependencies define a dependency matrix D whose i,j element is the dependency of variable i on variable j. Note that, while the correlation matrix C is a symmetric matrix, because the influence of variable j on variable i is usually not equal to the influence of variable i on variable j, the dependency matrix D is non-symmetrical.

In the final step, we sort the variables according to the system-level influence of each variable on the correlations between all other pairs. The system-level influence of variable j, SLI(j), is defined as the sum of the influence of j on all other variables i not equal to j,

$$\mathrm{SLI}(j) \equiv \sum_{i \neq j}^{N-1} D(i,j). \tag{5}$$

Supplementary Information B - List of Securities

TABLE I: List of asset security number, used in the analysis, for all securities belonging to the TA25 index for the period 2006-2010.

#	**Asset number**	#	Asset Number
1	126011	19	691212
2	224014	20	695437
3	230011	21	736579
4	260018	22	746016
5	268011	23	777037
6	273011	24	798017
7	281014	25	829010
8	304014	26	1081124
9	475020	27	1081165
10	576017	28	1081819
11	585018	29	1083484
12	604611	30	1084128
13	608018	31	1087949
14	611012	32	1092428
15	629014	33	1098474
16	639013	34	1100007
17	649012	35	1101534
18	662577	36	2590248

Acknowledgments.

We would like to thank Yoash Shapira, Idan Michaeli and Dustin Plotnick for all of their help. DYK and EBJ acknowledge support in part by the Tauber Family Foundation and the Maguy-Glass Chair in Physics of Complex Systems at Tel Aviv University. HES and DYK thank the support of the Office of Naval Research (ONR, Grant N00014-09-1-0380, Grant N00014-12-1- 0548), Keck Foundation and the NSF (Grant CMMI 1125290) for support. This work was also supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center (DoI/NBC) contract number

D12PC00285. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.

Author contributions.

DYK, EBJ, HES and GGG designed and performed the analysis, and have equally contributed to the writing of this manuscript.

Additional Information.

The authors declare no competing financial interests.

---

[1] Kenett, D. Y. *et al.* Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PloS one* **5**, e15032 (2010).

[2] Baba, K., Shibata, R. & Sibuya, M. Partial correlation and conditional correlation as measures of conditional independence. *Australian & New Zealand Journal of Statistics* **46**, 657–664 (2004).

[3] Shapira, Y., Kenett, D. Y. & Ben-Jacob, E. The index cohesive effect on stock market correlations. *The European Physical Journal B - Condensed Matter and Complex Systems* **72**, 657–669 (2009).

[4] Kenett, D. Y., Shapira, Y. & Ben-Jacob, E. Rmt assessments of the market latent information embedded in the stocks raw, normalized, and partial correlations. *Journal of Probability and Statistics* **2009** (2010).

[5] Madi, A. *et al.* Analyses of antigen dependency networks unveil immune system reorganization between birth and adulthood. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **21**, 016109–016109 (2011).

[6] Kenett, Y. N., Kenett, D. Y., Ben-Jacob, E. & Faust, M. Global and local features of semantic networks: Evidence from the hebrew mental lexicon. *PloS one* **6**, e23912 (2011).

[7] Madi, A., Bransburg-Zabary, S., Kenett, D. Y., Ben-Jacob, E. & Cohen, I. *Naturally Occuring Antibodies (NAbs)*, chap. The natural autoantibody repertoire in newborns and adults: a current overview (Landes, 2012).

[8] Kenett, D. Y., Preis, T., Gur-Gershgoren, G. & Ben-Jacob, E. Dependency network and node influence: Application to the study of financial markets. *International Journal of Bifurcation and Chaos* **22**, 1250181 (2012).

[9] Kenett, D. Y., Raddant, M., Zatlavi, L., Lux, T. & Ben-Jacob, E. Correlations in the global financial village. *International Journal of Modern Physics Conference Series* **16**, 13–28 (2012).

[10] Pearson, K. Contributions to the mathematical theory of evolution ii: skew variations in homogeneous material. *Phil. Trans. R. Soc. Lond. A* **186**, 343–414 (1895).