Noise and information in Econophysics

Rosario Nunzio Mantegna
Department of Physics and Chemistry, Palermo University, Italy
Department of Computer Science, UCL, London, UK
Complexity Science Hub, Vienna, Austria

Outline

- A few concepts in finance
- Aggregation and filtering of information in complex systems
- Random matrix theory and proximity based networks
- Trading decisions in the presence of exhogenous and endogenous market signals
- Conclusions

Econophysics is a hybrid discipline

Financial markets

A financial market can be considered as a 'model' complex system.

In a financial market there are many heterogeneous agents interacting to perform the collective task of finding the best price for financial assets.

Market efficiency definition

"A capital market is said to be efficient if it fully and correctly reflects all relevant information in determining security prices. Formally, the market is said to be efficient with respect to some information set .... if security prices would be unaffected by revealing that information to all participants. Moreover, efficiency with respect to an information set .... implies that it is impossible to make economic profits by trading on the basis of that information set"¶

¶ B. Malkiel, Efficient market hypothesis, New Palgrave Dictionary of Money and Finance, Macmillan, London 1992

Information aggregation in markets

The first observation that markets can act as institutions performing information aggregation is due to Fredrich van Hayek. In his 1945 work on "The use of knowledge in society" he wrote:

"The mere fact that there is one price for any commodity – or rather that local prices are connected in a manner determined by the cost of transport, etc. – brings about the solution which (it is just conceptually possible) might have been arrived at by one single mind possessing all the information which is in fact dispersed among all the people involved in the process."

Hayek FA. The Use of Knowledge in Society. American Economic Review. 1945;35(4):519-30.

Arbitrage opportunity

One of the main paradigms used for the modeling of price dynamics in a financial market is the concept of absence of **arbitrage opportunity.**

An *arbitrage opportunity* is present in a market when an economic actor can devise a trading strategy which is able to provide her or him a financial gain continuously and without risk.

An example

St.Louis

Miami

At a given time 1 kg of wheat costs 1.30 USD in St. Louis and 1.45 USD in Miami.

The cost of transporting and storing 1 kg of wheat from St. Louis to Miami is 0.05 USD

By buying 10,000 kg in St. Louis and selling them immediately after in Miami it is possible to make a risk-free profit

10000 (1.45-1.30-0.05)=1000 USD

Exploited arbitrage opportunities

If this action is repeated this implies that the demand of wheat (and therefore its price) in St. Louis increases whereas the supply in Miami increases and its price decreases.

In an efficient market, the exploiting of an arbitrage opportunity implies its disappearance in a (usually) short time period.

The absence of arbitrage opportunity implies that the price of a financial asset is hard to predict or **unpredictable.**

Geometric Brownian motion

In 1965, Paul Samuelson[‡] formalized the assumption that properly anticipated stock prices should behave as random processes.

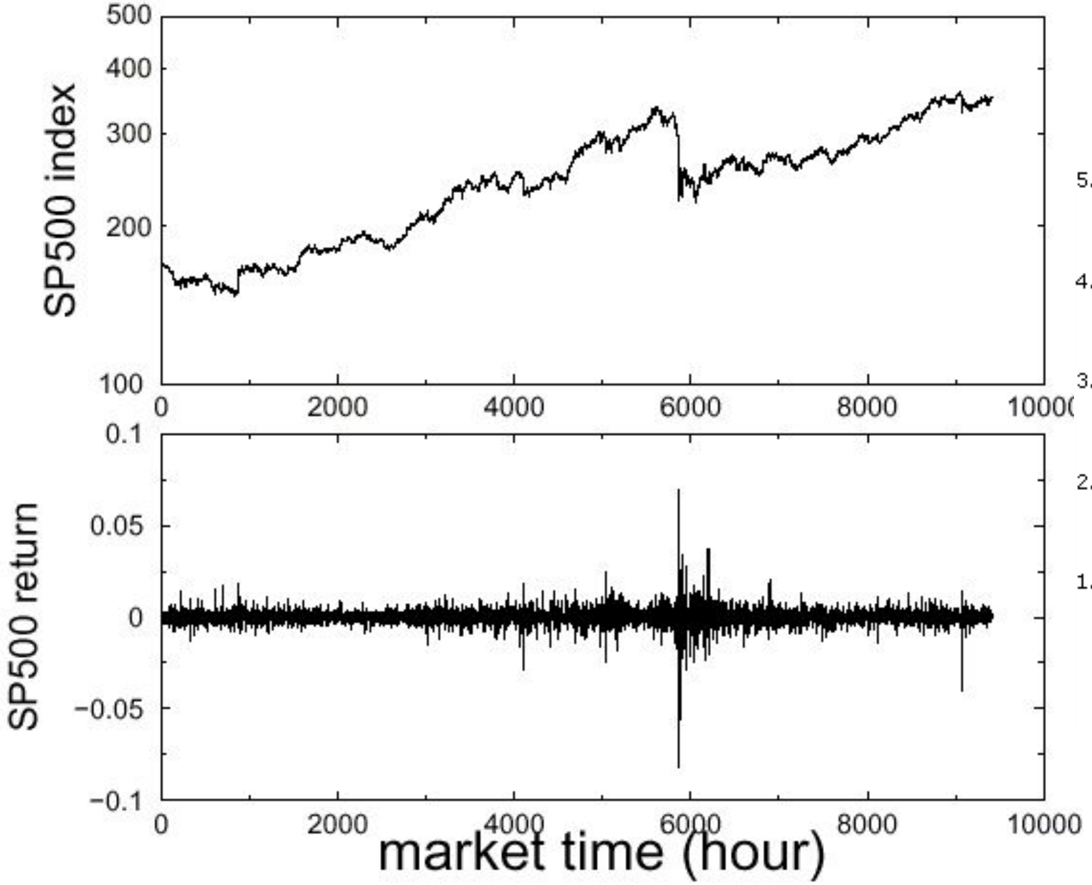
The geometric Brownian motion is considered the simplest random process describing the price dynamics of a financial asset.

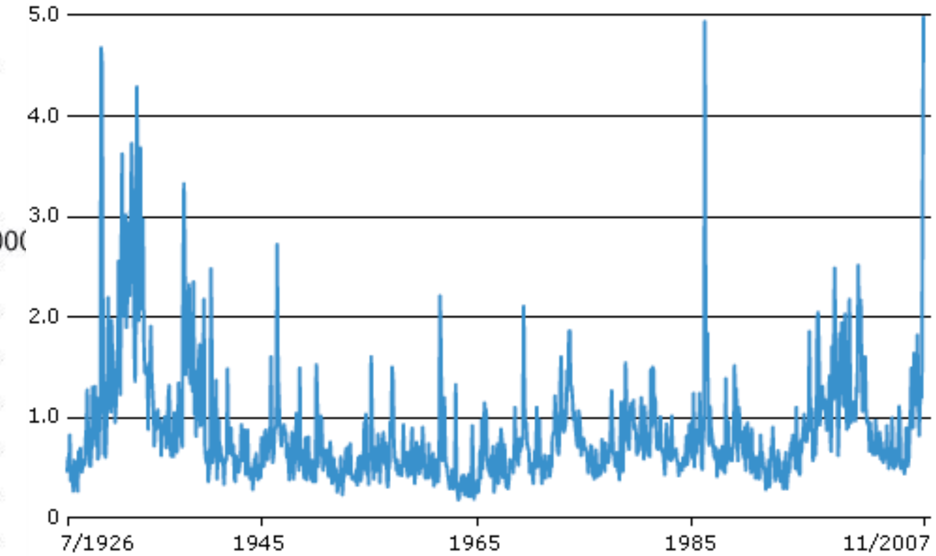
$$dP(t) = \mu P(t)dt + \sigma P(t)dW$$

An idealized model of stock market where the stock price dynamics is described by a geometric Brownian motion exists and provides the theoretical foundation for quantitative finance.

[‡]P. Samuelson, Industrial Man. Review, **6**, 41-45 (1965)

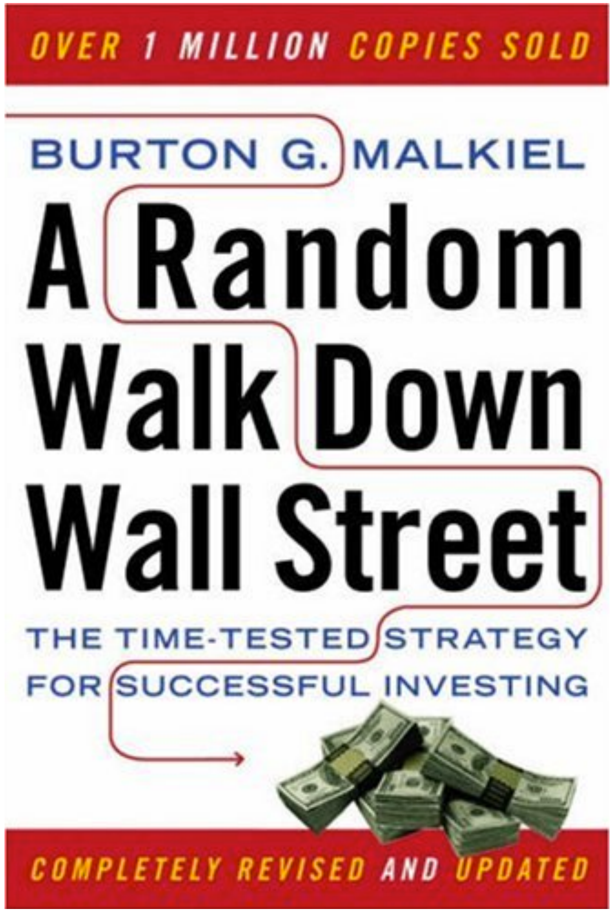
The price (and return) dynamics of a financial asset





daily volatility

1/1984 1/1989


OVER 1 MILLION COPIES SOLD
BURTON G. MALKIEL
A Random Walk Down Wall Street
THE TIME-TESTED STRATEGY FOR SUCCESSFUL INVESTING
COMPLETELY REVISED AND UPDATED


A Non-Random
Walk Down Wall Street
ANDREW W. LO & A. CRAIG MACKINLAY

Pourquoi donner un Nobel à trois économistes aux théories opposées ?

Le Monde.fr | 14.10.2013 à 18h23 • Mis à jour le 15.10.2013 à 07h50

October 15, 2013 1:40 pm

The Nobel committee is muddled on the nature of economics

By John Kay

The term 'rationality' lacks relevance in a world characterised by imperfect information

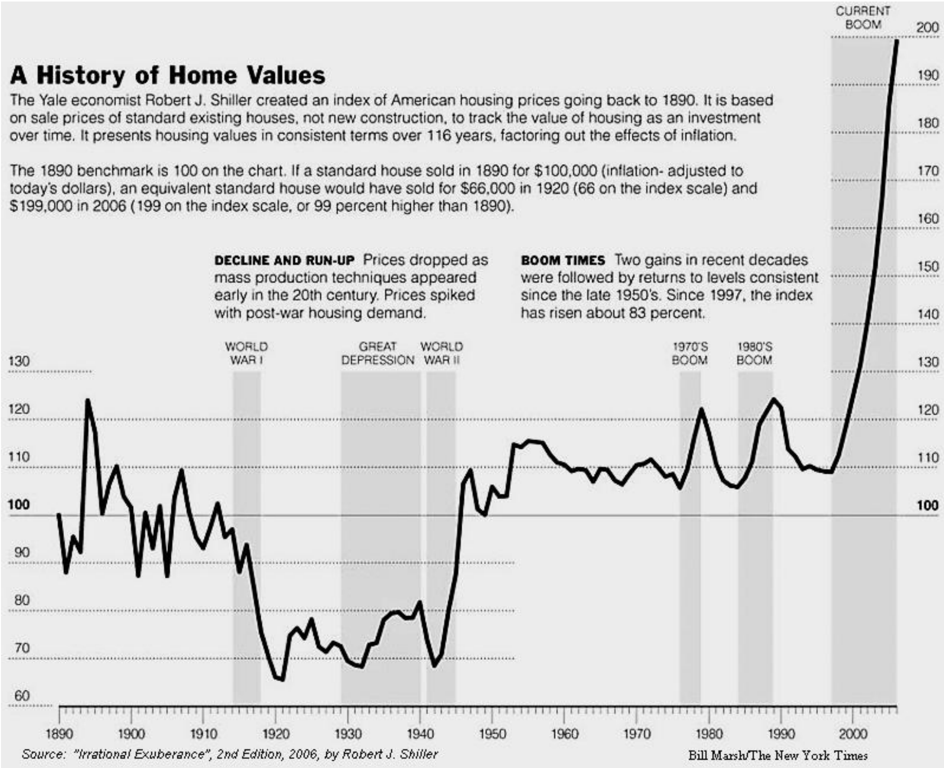
The Royal Swedish Academy of Sciences continues to astonish the public when awarding the Nobel Memorial Prize in Economics. In 2011 it celebrated the success of recent research in promoting macroeconomic stability. This year it pays tribute to the capacity of economists to predict the long-run movement of asset prices.

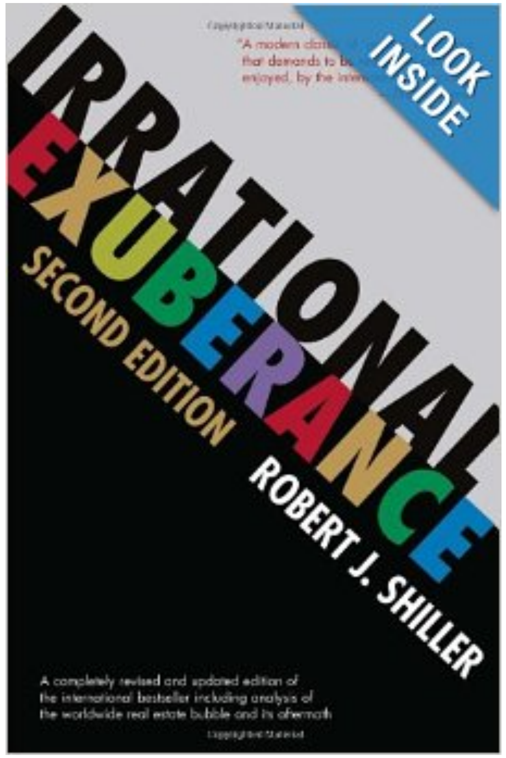
People with knowledge of financial economics may be further surprised that this year Eugene Fama and Robert Shiller are both recipients. Prof Fama made his name by developing the efficient market hypothesis, long the cornerstone of finance theory. Prof Shiller is the most prominent critic of that hypothesis. It is like awarding the physics prize jointly to Ptolemy for his theory that the Earth is the centre of the universe, and to Copernicus for showing it is not.

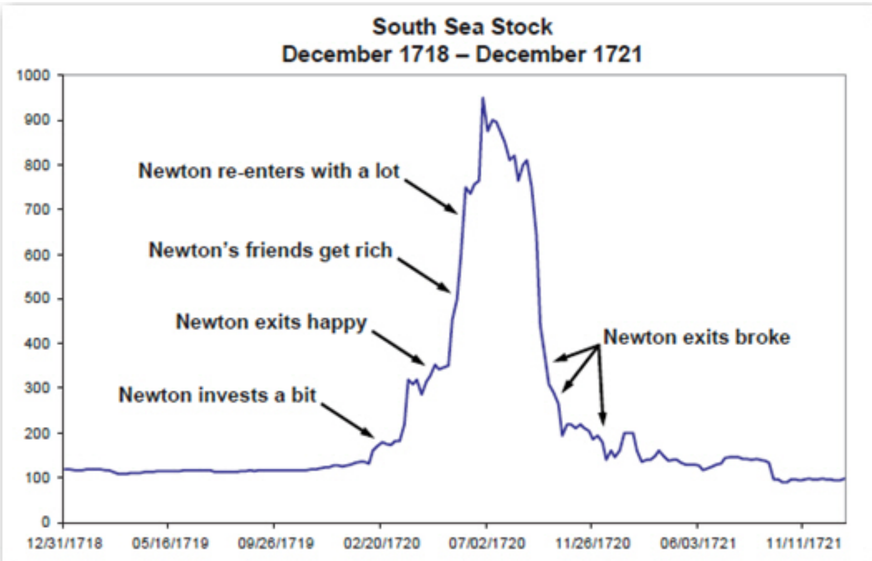
ANALYS

SUDAN

Rage ag
Econom
fears fo
neighbo
Omar a

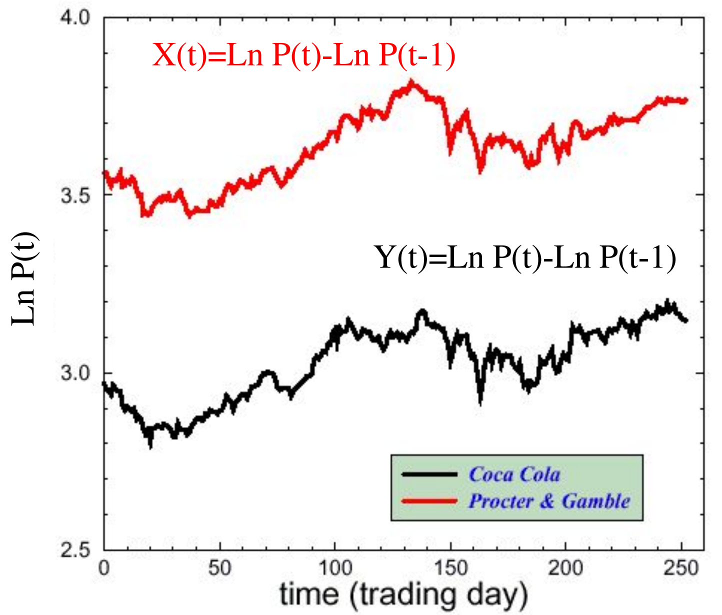

A History of Home Values
The Yale economist Robert J. Shiller created an index of American housing prices going back to 1890. It is based on sale prices of standard existing houses, not new construction, to track the value of housing as an investment over time. It presents housing values in consistent terms over 116 years, factoring out the effects of inflation.
The 1890 benchmark is 100 on the chart. If a standard house sold in 1890 for $100,000 (inflation- adjusted to today's dollars), an equivalent standard house would have sold for $66,000 in 1920 (66 on the index scale) and $199,000 in 2006 (199 on the index scale, or 99 percent higher than 1890).
DECLINE AND RUN-UP Prices dropped as mass production techniques appeared early in the 20th century. Prices spiked with post-war housing demand.
BOOM TIMES Two gains in recent decades were followed by returns to levels consistent since the late 1950's. Since 1997, the index has risen about 83 percent.
CURRENT BOOM
WORLD WAR I
GREAT DEPRESSION
WORLD WAR II
1970'S BOOM
1980'S BOOM
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
Source: "Irrational Exuberance", 2nd Edition, 2006, by Robert J. Shiller
Bill Marsh/The New York Times

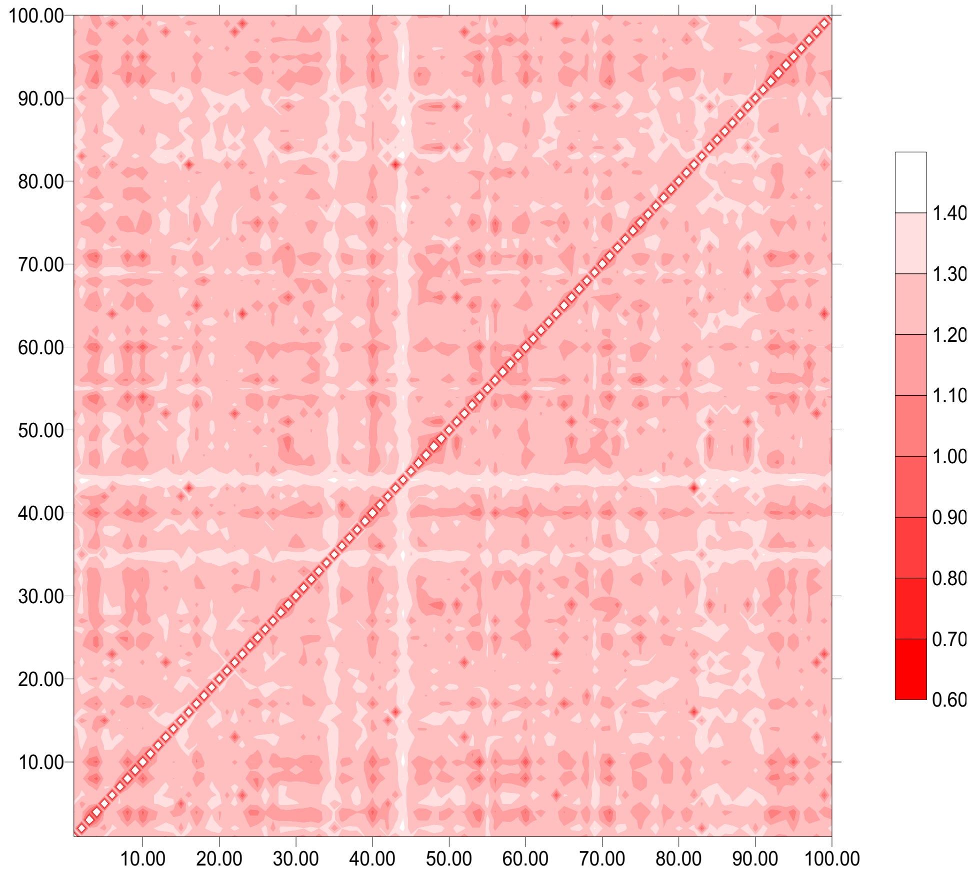

LOOK INSIDE
IRRATIONAL
EXUBERANCE
SECOND EDITION
ROBERT J. SHILLER


South Sea Stock
December 1718 – December 1721
1000
900
800
700
600
500
400
300
200
100
0
Newton re-enters with a lot
Newton's friends get rich
Newton exits happy
Newton invests a bit
Newton exits broke
12/31/1718
05/16/1719
09/26/1719
02/20/1720
07/02/1720
11/26/1720
06/03/1721
11/11/1721


Bitcoin Price
Bitcoin.com
$18000.00
$16000.00
$14000.00
$12000.00
$11157.56
$10000.00
$8000.00
$6000.00
$4000.00
$2000.00
Oct
2017
Apr
Jul
Oct
2018

Information is present in the multivariate time evolution of asset returns





$$\rho_{X,Y} = corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

$$r = \frac{1}{(n-1)} \sum \frac{(X-\mu_X)(Y-\mu_Y)}{\sigma_X \sigma_Y}$$

$$d_{XY} = \sqrt{2(1-\rho_{XY})}$$

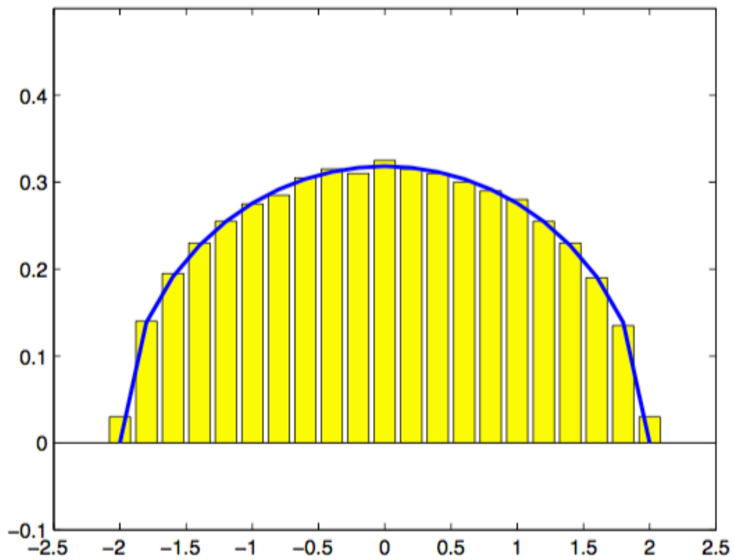
The Semi-circle law in Random Matrix Theory

In 1955, Wigner showed that the limiting eigenvalue distribution of random symmetric $n \times n$ matrices $X = (A + A^T)/2$ where $A = G(n, n)$,

follows a semi-circle distribution, which is given by

$$\rho\left(\tilde{\lambda}\right) = \frac{1}{2\pi}\sqrt{4 - \tilde{\lambda}^2}$$

When properly normalized $\tilde{\lambda} = \lambda / \sqrt{n/2}$, the curve looks like a semi-circle of radius 2



Semi-circle law with one 1000 × 1000 matrix.

E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. Annals of Mathematics, 62:548–564, 1955.

Random Matrix Theory for correlation of random Gaussian variables

Suppose that N elements are described by T records, i.e. we have a random matrix with NxT records. Each record is a realization of independent Gaussian random variables with zero mean and variance σ^2.

In the limit $T \to \infty$, the correlation matrix of this set of N variables is the identity matrix, which has associated a spectrum of eigenvalues which is composed of N delta functions at $\lambda=1$

In the limit $T, N \to \infty$ with a fixed ratio $Q=T/N \geq 1$ random matrix theory provides the asymptotic probability density of the eigenvalues.

The convergence to the asymptotic pdf is already observed for relatively low values of N and T.

The asymptotic PDF of eigenvalues

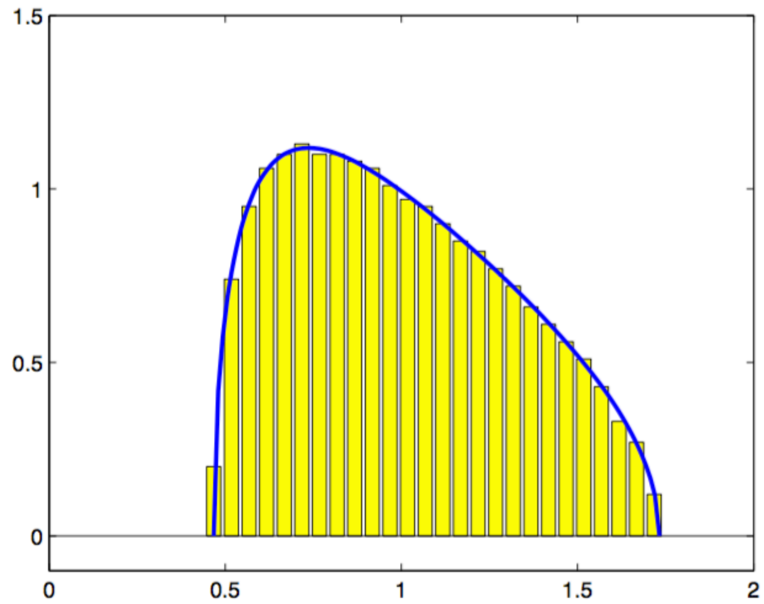
The asymptotic probability density function is

$$\rho(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{\lambda}$$

where

$$\lambda_{\min}^{\max} = \sigma^2\left(1 + 1/Q \pm 2\sqrt{1/Q}\right)$$

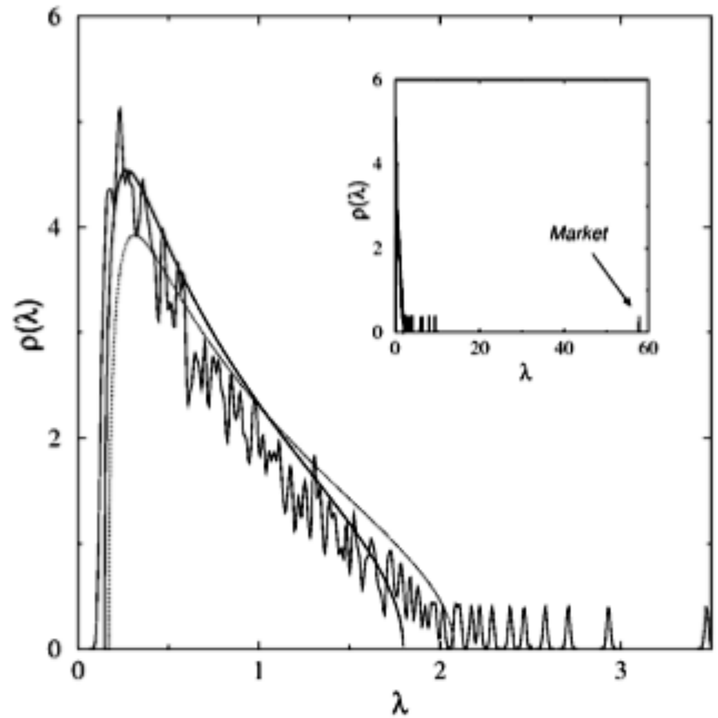
with Q=T/N=1/q

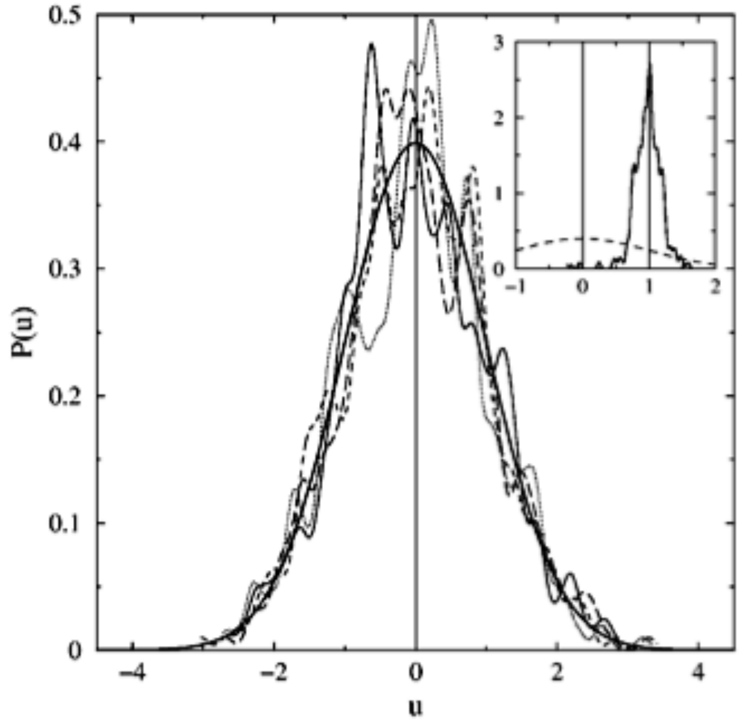


Marcenko-Pastur Law(q=0.1) with a NxT 2000 × 20000 matrix X.

Marčenko, V.A. and Pastur, L.A., 1967. Distribution of eigenvalues for some sets of random matrices. Mathematics of the USSR-Sbornik, 1(4), p.457.

Random Matrix Theory of financial correlation





$$\rho(\lambda) = \frac{T}{2\pi\,\lambda\sigma^2}\sqrt{(\lambda_{\max}-\lambda)(\lambda-\lambda_{\min})}, \quad \text{where} \quad \lambda_{\min}^{\max} = \sigma^2\left(1+1/Q \pm 2\sqrt{1/Q}\right), \quad \sigma^2 = 1-\frac{\lambda_1}{N}, \text{ and } Q = \frac{T}{N}.$$

L. Laloux, P. Cizeau, J.-P. Bouchaud & M. Potters, *Phys. Rev. Lett.* **83**, 1468 (1999).
V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley, *Phys. Rev. Lett.* **83**, 1471 (1999).

Characteristics of the eigenvalue spectrum

The spectrum of a typical portfolio of financial assets (e.g. stocks) can be divided in three classes of eigenvalues:

1) The largest eigenvalue describes the common behavior of stocks (what is called "the market"). It is incompatible with the random matrix theory of random variables.

2) A fraction of 5% of the eigenvalues is also incompatible with the random matrix theory because eigenvalues fall outside $]\lambda_{min}, \lambda_{max}[$

The low part of the spectrum

3) The remaining eigenvalues assume values between λ_{min}, λ_{max} and therefore one cannot say whether the eigenspace, which is corresponding to these eigenvalues, contains information or not.

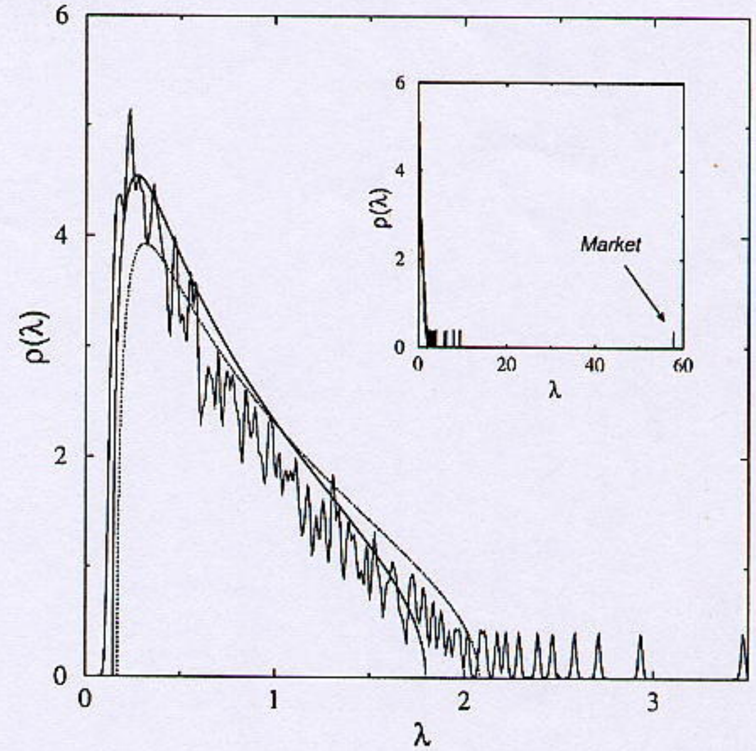
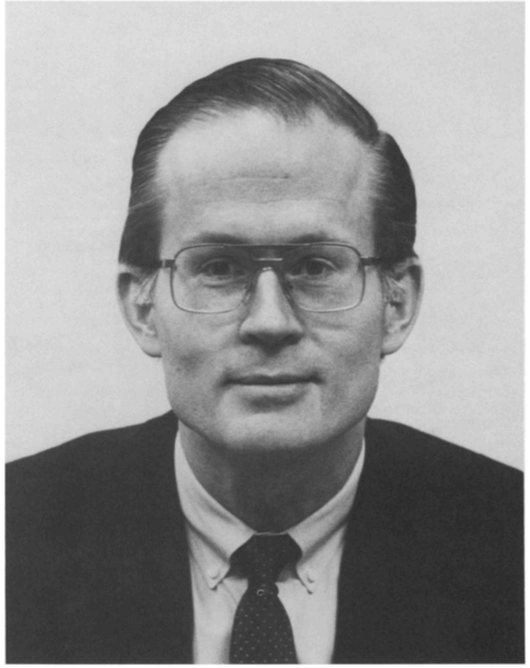


FIG. 1. Smoothed density of the eigenvalues of **C**, where the correlation matrix **C** is extracted from $N = 406$ assets of the S&P 500 during the years 1991–1996. For comparison we have plotted the density Eq. (3) for $Q = 3.22$ and $\sigma^2 = 0.85$: this is the theoretical value obtained assuming that the matrix is purely random except for its highest eigenvalue (dotted line). A better fit can be obtained with a smaller value of $\sigma^2 = 0.74$ (solid line), corresponding to 74% of the total variance. Inset: Same plot, but including the highest eigenvalue corresponding to the market, which is found to be 25 times greater than λ_{max}.



THE JOURNAL OF FINANCE • VOL. XLI, NO. 3 • JULY 1986

Noise

FISCHER BLACK*

ABSTRACT

The effects of noise on the world, and on our views of the world, are profound. Noise in the sense of a large number of small events is often a causal factor much more powerful than a small number of large events can be. Noise makes trading in financial markets possible, and thus allows us to observe prices for financial assets. Noise causes markets to be somewhat inefficient, but often prevents us from taking advantage of inefficiencies. Noise in the form of uncertainty about future tastes and technology by sector causes business cycles, and makes them highly resistant to improvement through government intervention. Noise in the form of expectations that need not follow rational rules causes inflation to be what it is, at least in the absence of a gold standard or fixed exchange rates. Noise in the form of uncertainty about what relative prices would be with other exchange rates makes us think incorrectly that changes in exchange rates or inflation rates cause changes in trade or investment flows or economic activity. Most generally, noise makes it very difficult to test either practical or academic theories about the way that financial or economic markets work. We are forced to act largely in the dark.

In system where information is densely packed into sets of records (e.g. time series) we have noise and information simultaneously present.

When information is providing marginal advantage the discrimination between noise and information is not straightforward (all easy detections have been ruled out by the action of arbitrage opportunities).

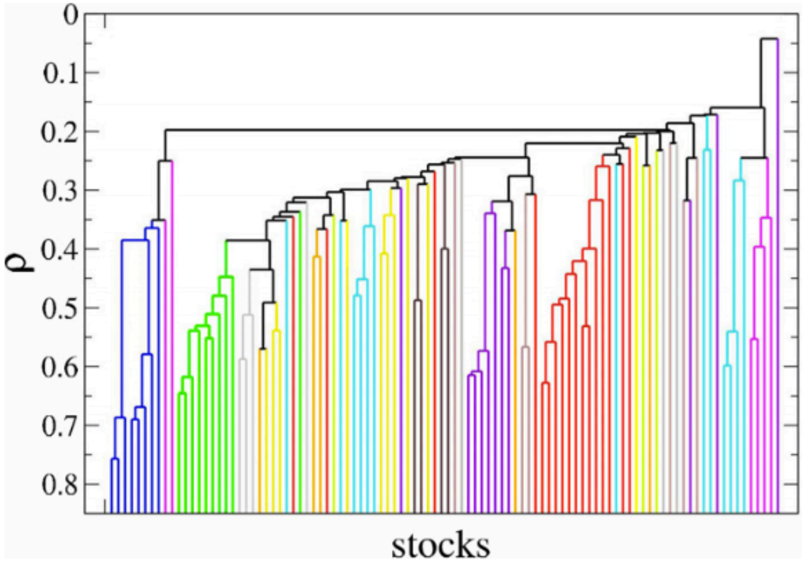
Random matrix theory suggests that some information is quite distinct from noise but there is also information hardly distinguishable from noise and noise hardly distinguishable from information.

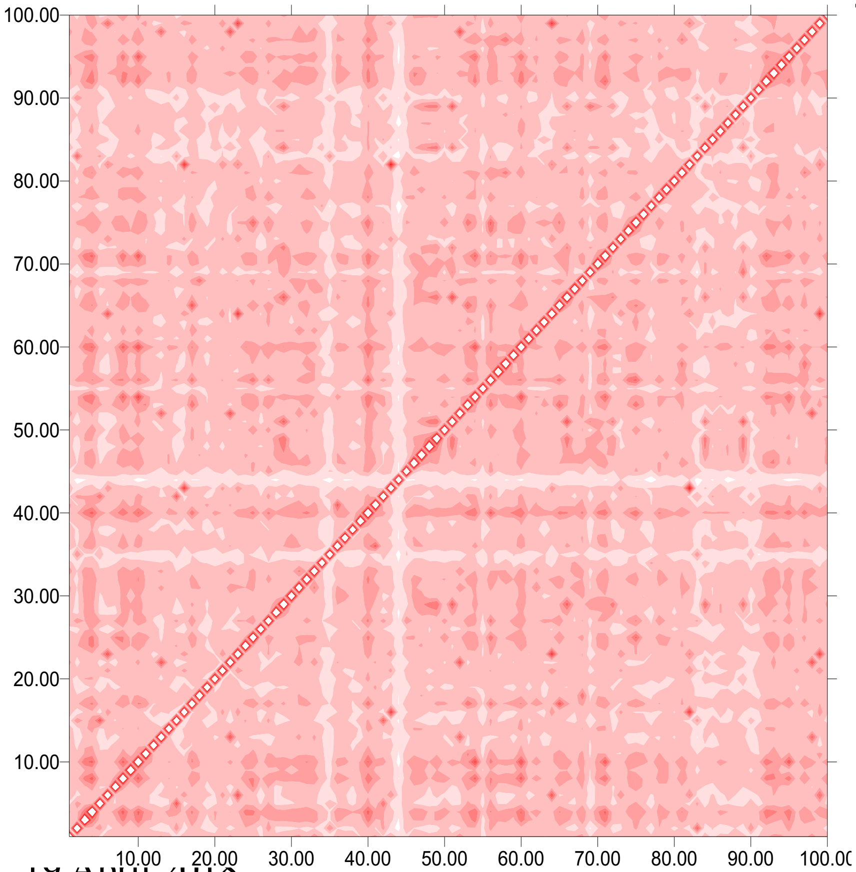
Bouchaud and collaborators have coined the sentence that the correlation matrix is “noise dressed”. My view is that some information is “dressed” as “noise” it is, i.e. it is hardly distinguishable from noise.

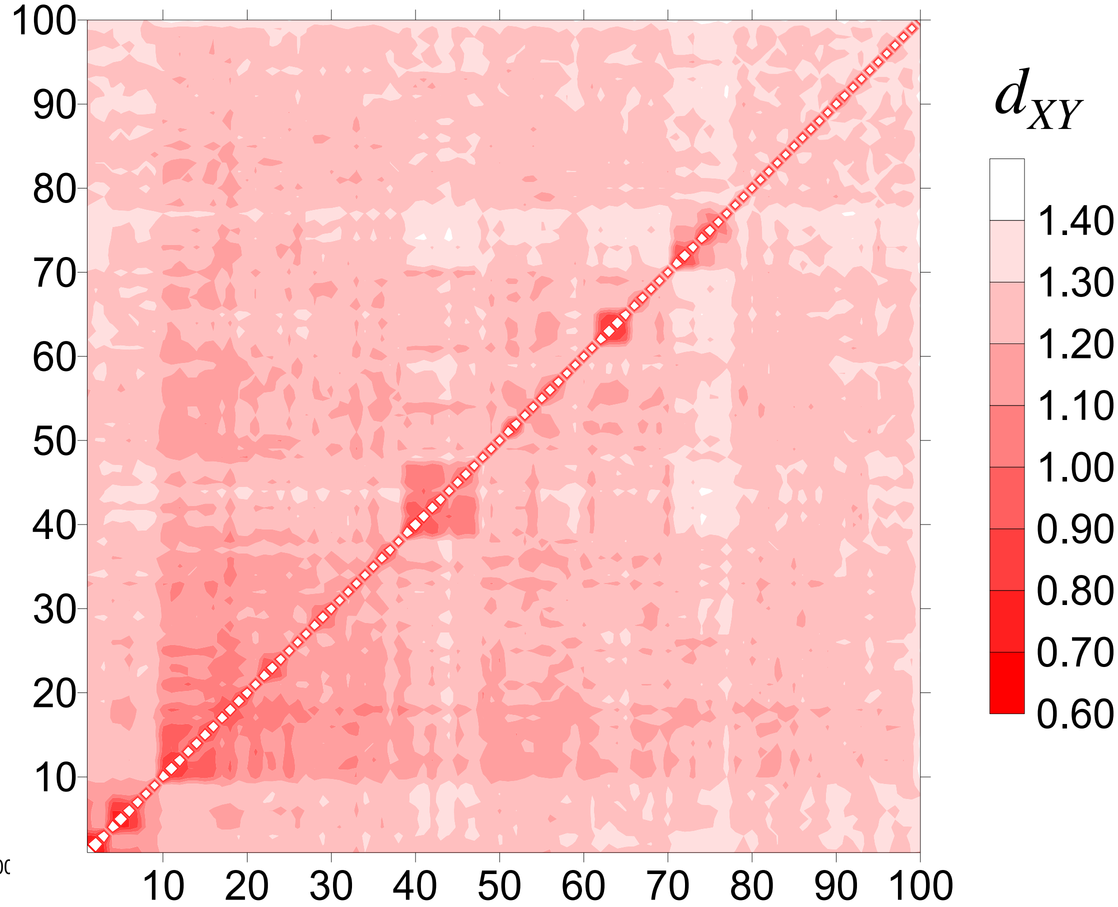
Hierarchical clustering is another method detecting information which is present in the multivariate time evolution of asset returns

$$\rho_{XY} = corr(X, Y)$$

$$d_{XY} = \sqrt{2(1 - \rho_{XY})}$$

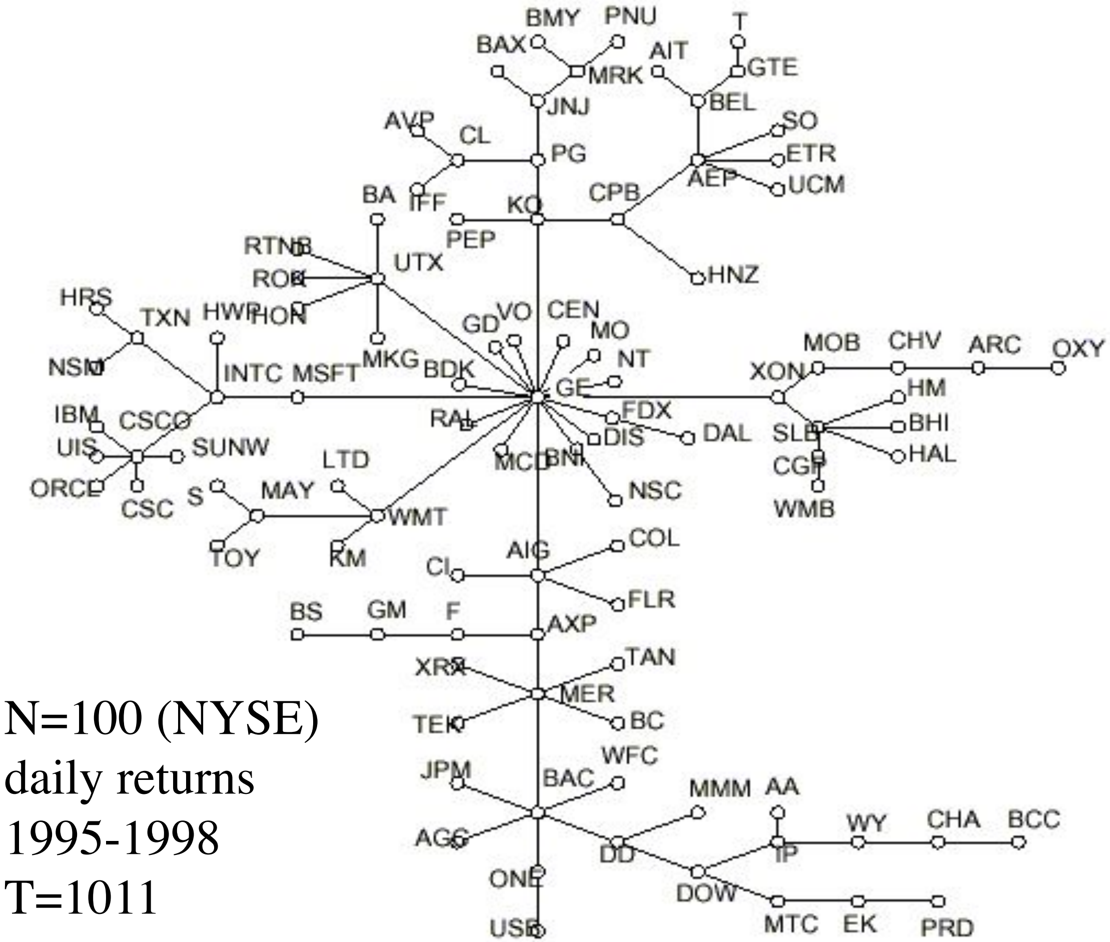




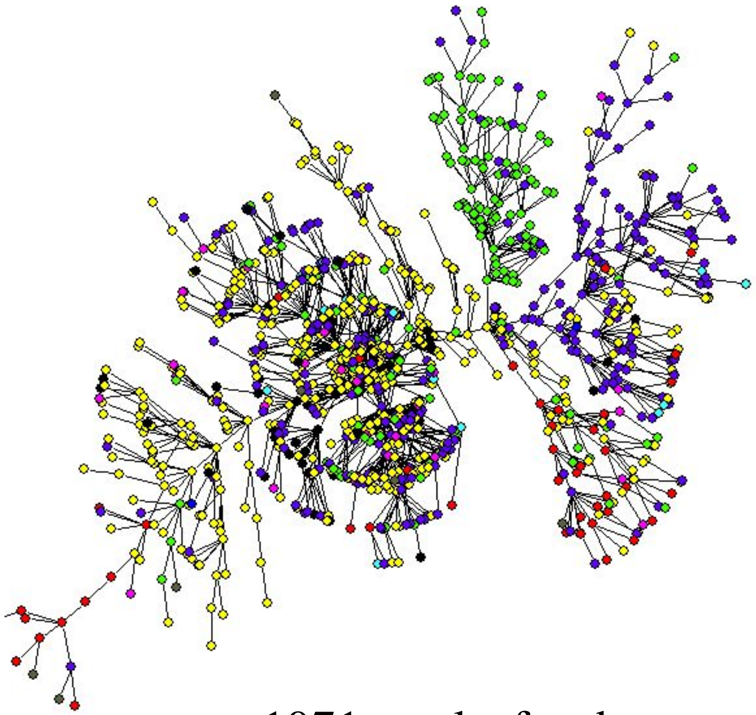


Another popular method of information filtering

Proximity based networks: Minimum spanning tree



N=100 (NYSE)
daily returns
1995-1998
T=1011



1071 stocks for the period 1987–1998
T=3030 trading days .

Mantegna, R.N., 1999. Hierarchical structure in financial markets. The European Physical Journal B-11(1), pp.193-197.
Bonanno, G., Caldarelli, G., Lillo, F. and Mantegna, R.N., 2003. Topology of correlation-based minimal spanning trees in real and model markets. Physical Review E, 68(4), p.046130.

A data driven study: The NCSD database.

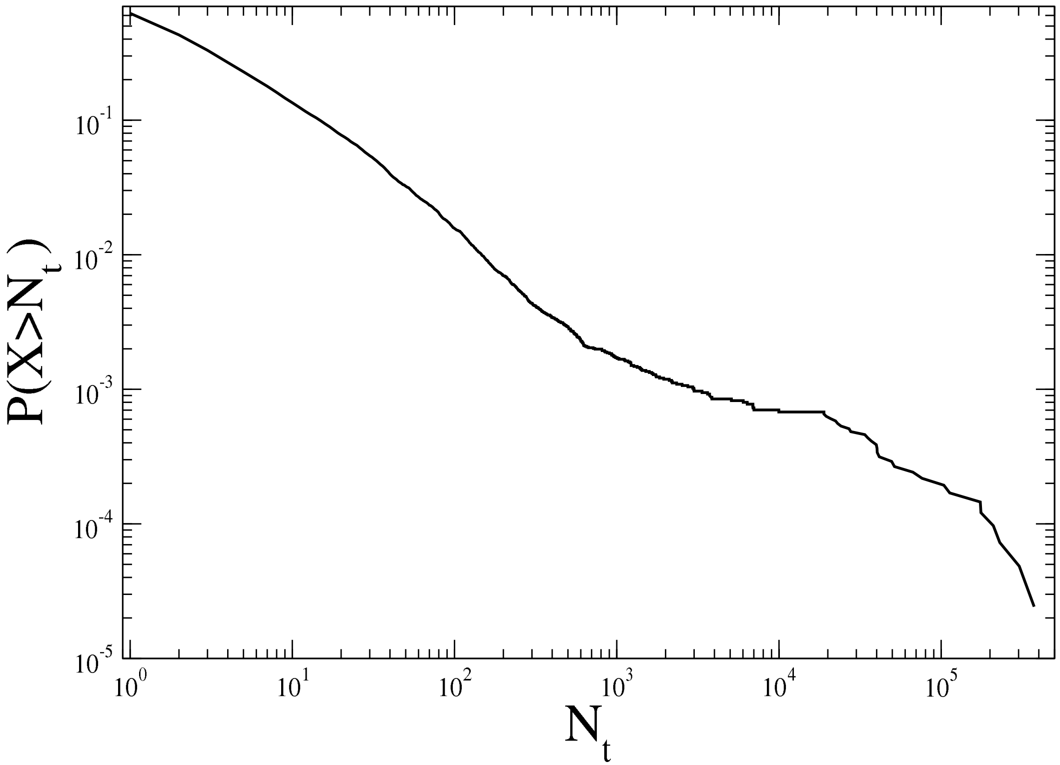
From January 1st, 1995 the Nordic Center Securities Data (NCSD) (today Euroclear) collects a database recording the daily ownership of financial portfolios and trading records (until 2009) of different categories of Finnish investors (companies, governmental investors, foreign institutions investors, non profit organizations, financial institutions and households).

Identity of the investors is coded for privacy reasons.

This database has been extensively investigated at the aggregated level of categories of investors in the financial literature by Grinblatt and Keloharju (2000, 2001,2009).

The set of investors is quite heterogeneous in terms of volume, frequency of trading, portfolio nature, wealth, etc .

Ex: Cumulative probability density $P(X > N_t)$ of the number of transactions N_t performed by 41250 Finnish legal entities trading the Nokia stock (typically but not exclusively at the Nordic Stock Exchange) during 2003.



The flux of financial news about Nokia: a proxy from Thomson Reuters

THOMSON REUTERS

Contact Us

Search

HOME | PRODUCTS & SERVICES | NEWS & IDEAS | INVESTOR RELATIONS | ABOUT US | CAREERS

FINANCIAL HOME > FINANCIAL PRODUCTS & SERVICES > PRODUCTS A-Z > Thomson Reuters NewsScope Application License

FINANCIAL

FINANCIAL PRODUCTS & SERVICES

- Commodities
- Corporate Services
- Deal Making
- Equities & Derivatives
- Fixed Income
- Foreign Exchange & Money Markets
- Growth Markets
- Hedge Fund Solutions
- Independent Advisors
- Individual Traders/Investors
- Investment Management
- Islamic Finance
- Post Trade Services
- Pricing, Reference Data and Derived Analytics

PRODUCTS A-Z

THOMSON REUTERS NEWSSCOPE APPLICATION LICENSE

Trading professionals trust Reuters news for its speed and accuracy. Now you can leverage the power of our global news to fuel your mission-critical applications and algorithms in real-time. Seize opportunities, exploit market inefficiencies and manage event risk with world-renowned news from Reuters.

WITH THOMSON REUTERS NEWSSCOPE APPLICATION LICENSE, YOU CAN:

- **Leverage the power of our global news coverage** to fuel your applications in real time.
- **Seize opportunities** , exploit market inefficiencies and manage event risk.
- **React to breaking developments at lightning speed** – able to process more news, faster and more consistently than human operators, the next generation of news-powered algorithms is key to managing your investment portfolio.

WHY CHOOSE THOMSON REUTERS NEWSSCOPE APPLICATION LICENSE?

- **The most widely read news** – Reuters news is seen by over one billion people every day and is on the screens of over 370,000 financial and media professionals worldwide. We deliver over 500,000 alerts and over two million unique stories a year, consistently breaking political, company and economic news ahead of our

CONTACT US

Email, Phone

THOMSON REUTERS EIKON

THOMSON REUTERS

CLICK TO SEE THE DEMO

SHARE

We consider headline news in English language associated with one of the following Reuters Instrument Codes (RICs)

NOK.W, NOK1V.HE,
NOK1V.AS, NOKN.MX, NOKA.BA, NOKy.BE, NOK.MW,
NOKy.F, NOK1VEUR.VIp, NOK1VEUR.Ip, NOK1VEUR.STp,
NOKS.HA, NOKS.H, NOKS.DE, NOKy.D, NOKAc.BA,
NOKy.MU, NOKy.DE, NOK1VM0110.HE, NOK1VEUR.PZ,
NOK, NOK1VEUR.DEp, NOKI.ST, NOKS.BE, NOKS.F,
NOK.DF, NOK.N, NOKAd.BA, NOK.P, NOK.C,
NOK1VEUR.MIp, 0HAF.L, NOK1VEUR.PAp, NOK1V.MI,
NOKS.D, NOKS.MU

The set comprises 11 484 unique headlines.

2003-01-10,12:48:07.604,"20030110124749nL1011908","**HEADLINE**","nL1011908","2003-01-10 12:47:49","2003-01-10 12:48:07","Finnish HEX, Nokia edge up on back of German data","   HELSINKI, Jan 10 (Reuters) - Finnish stocks turned north at midsession on Friday with Nokia<NOK1V.HE> erasing earlier losses.

Nokia <NOK1V.HE>, the world's largest mobile phone maker and

Forestries were followed Nokia in the top three traded reporting season that kicks off with Nokia on January 23.

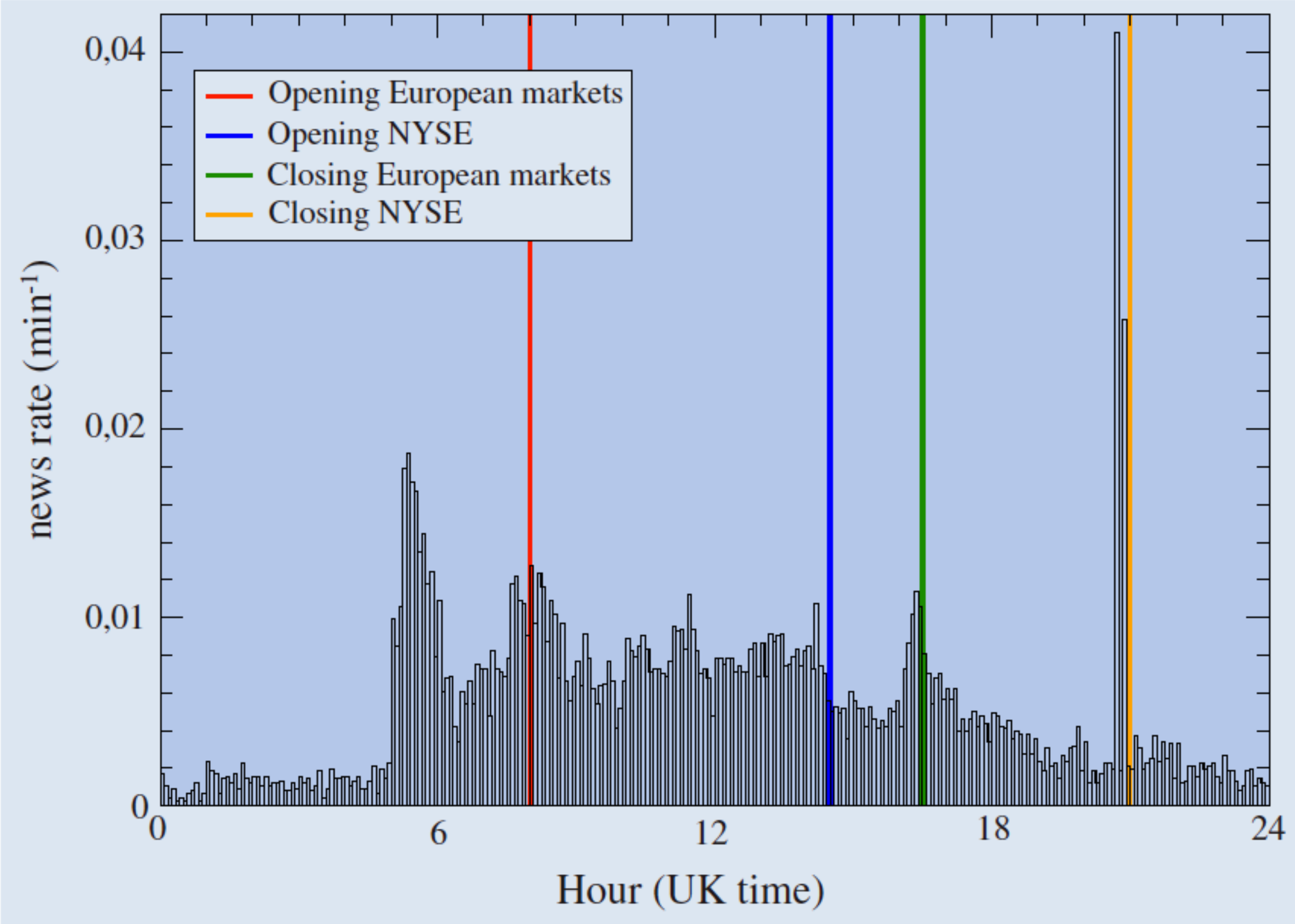


Figure 1. Average daily pattern of the arrival rate of news on the Nokia company. The rate is measured in number of headlines per minute. The vertical lines indicate the time of opening and closing of European and New York Stock Exchange market. Data are adjusted for the daylight saving time.

We consider the trading decisions of single investors trading Nokia stocks from Jan 2nd 2003 to Dec 30th 2008.

Category	# ids	N	V
Companies	8,396	1,009,226	4,825
Financial	392	4,079,174	21,402
Governamental	124	39,278	1,985
Non profit	922	21,778	248
Households	129,952	1,555,096	1,993
Foreign	1,405	789,552	7,685
Total	141,190	7,494,104	38,138

We define three categorical variables classifying the trading action of the investors. They are:

$$\frac{V_B - V_S}{V_B + V_S} > +\vartheta$$ **Buy state**

$$\frac{V_B - V_S}{V_B + V_S} < -\vartheta$$ **Sell state**

$$-\vartheta < \frac{V_B - V_S}{V_B + V_S} < +\vartheta \quad \text{with } V_B > 0 \text{ and } V_S > 0$$ **BuySell state**

with $\vartheta = 0.25$ and $\vartheta = 0.01$

Variables investigated

$$N^K(t) = N_B^K(t) + N_S^K(t) + N_{BS}^K(t)$$

$$\Delta N_A^K(t) = N_B^K(t) - N_S^K(t)$$

$$\Delta N_R^K(t) = \frac{N_B^K(t) - N_S^K}{N^K(t)}$$

Variables about market trading activity

$$Ret(t) = \log P(t) - \log P(t-1)$$

$$Vol(t) = 2\frac{P_{max}(t) - P_{min}(t)}{P_{max}(t) + P_{min}(t)}$$

Price dynamics market indicators

H(*t*)= # of headlines of news about Nokia

Variables about news

We also perform sentiment analysis on the text of the headlines (also stories have been analyzed).

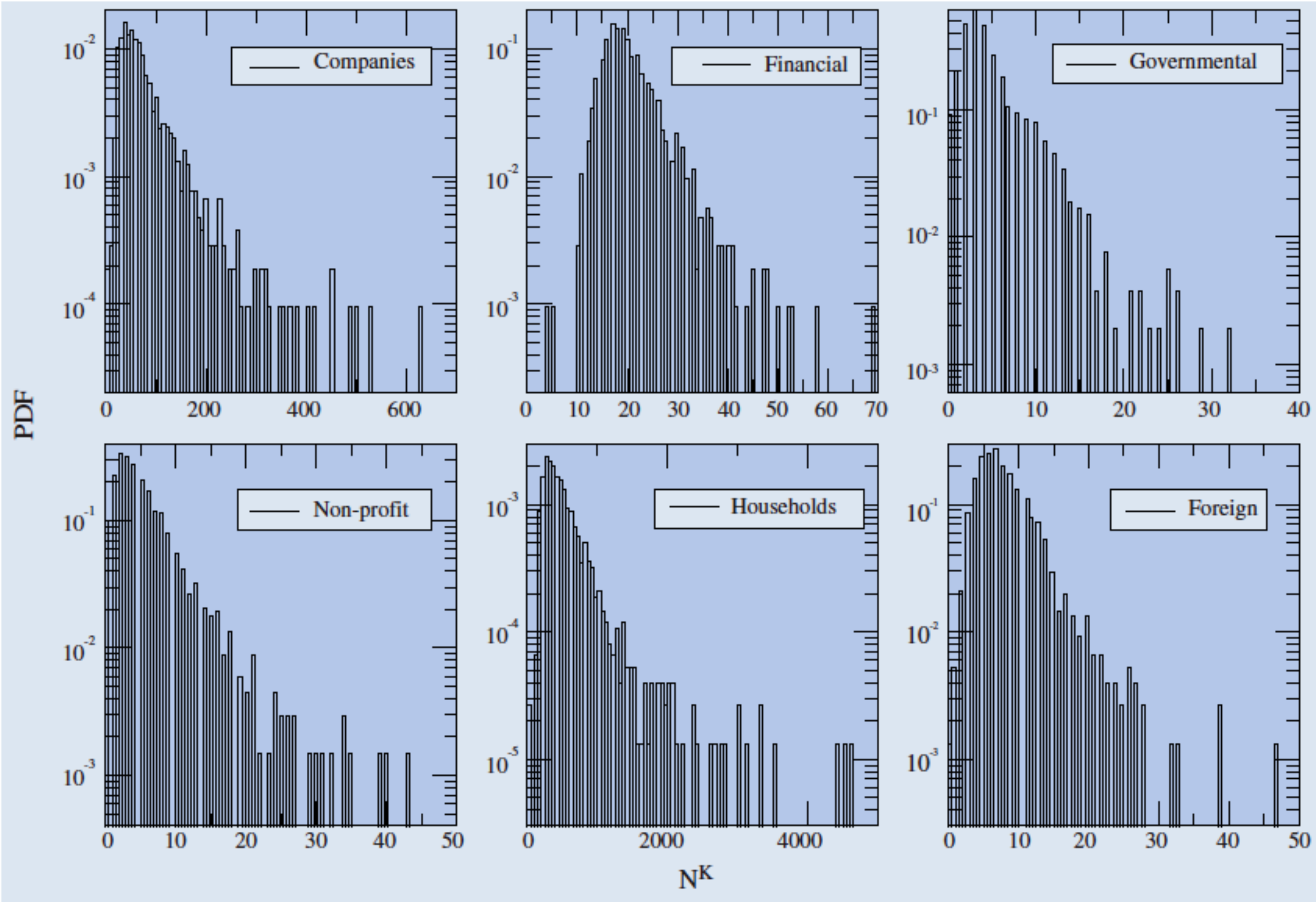
The sentiment analysis was performed by using General Inquire categories from the Harvard psychosocial dictionary

http://www.wjh.harvard.edu/~inquirer/

$$S_A(t) = G(t) - B(t)$$

$$S_R(t) = \frac{G(t) - B(t)}{G(t) + B(t)}$$

The tail of the probability density function of market activity N^K is approximately exponential for all categories of investors.



The basic explanatory variables are correlated

Corr[H, Vol]=0.501

Corr[S_A, Ret]=0.155

Corr[S_R, Ret]=0.118

We therefore also consider the partial correlation among variables

$$\rho\left(x, y \middle| z\right)$$

between variables x and y conditioned on the variable z. This quantity is the Pearson correlation coefficient between the residuals of x and y that are uncorrelated with z.

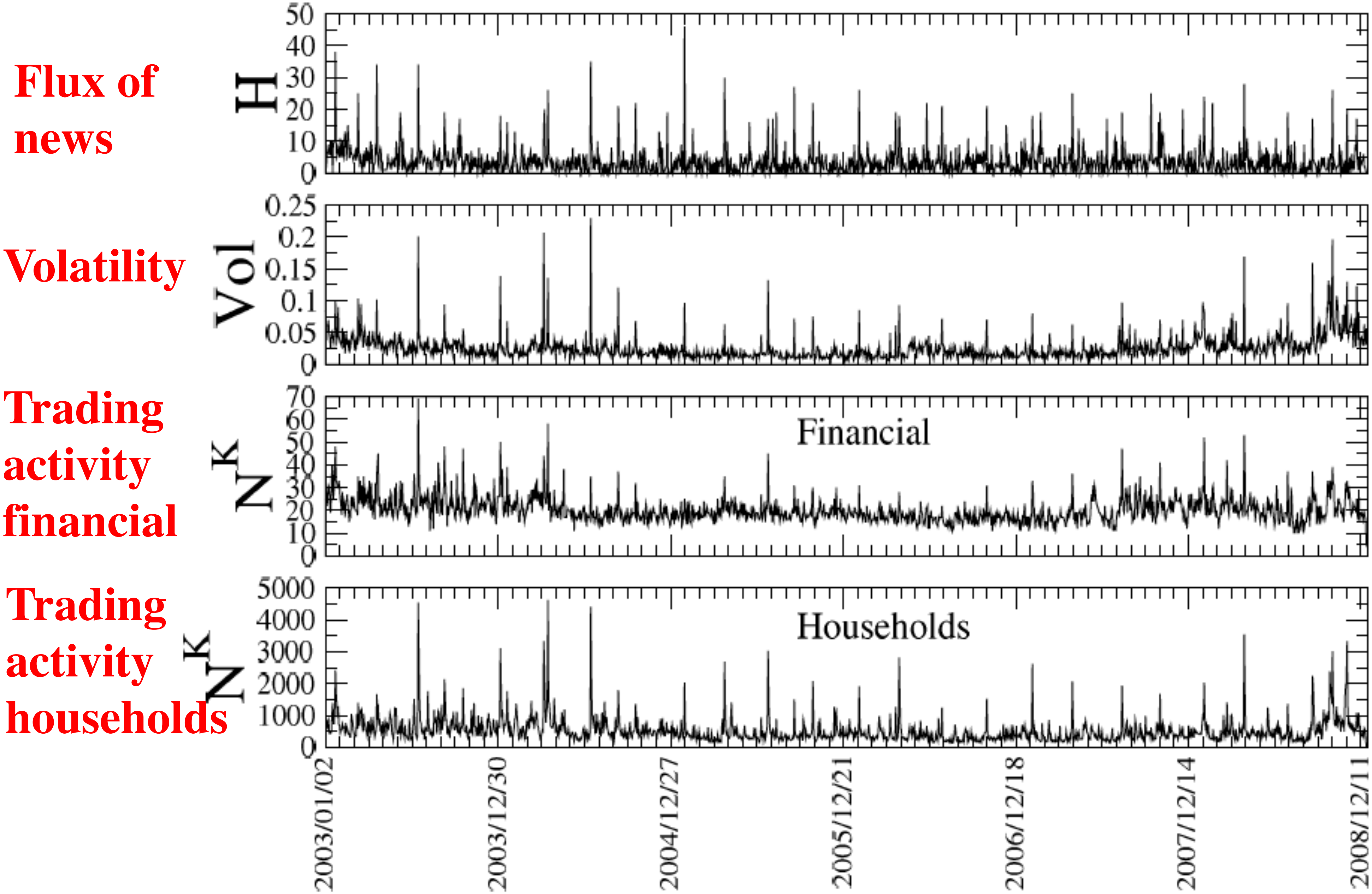


Figure 2. From top to bottom the figure shows the time series of the number of Nokia headlines $H(t)$, the daily volatility $Vol(t)$ of Nokia stock, and the time series of $N^K(t)$ for the category of Financial investors and for the category of Households investors.

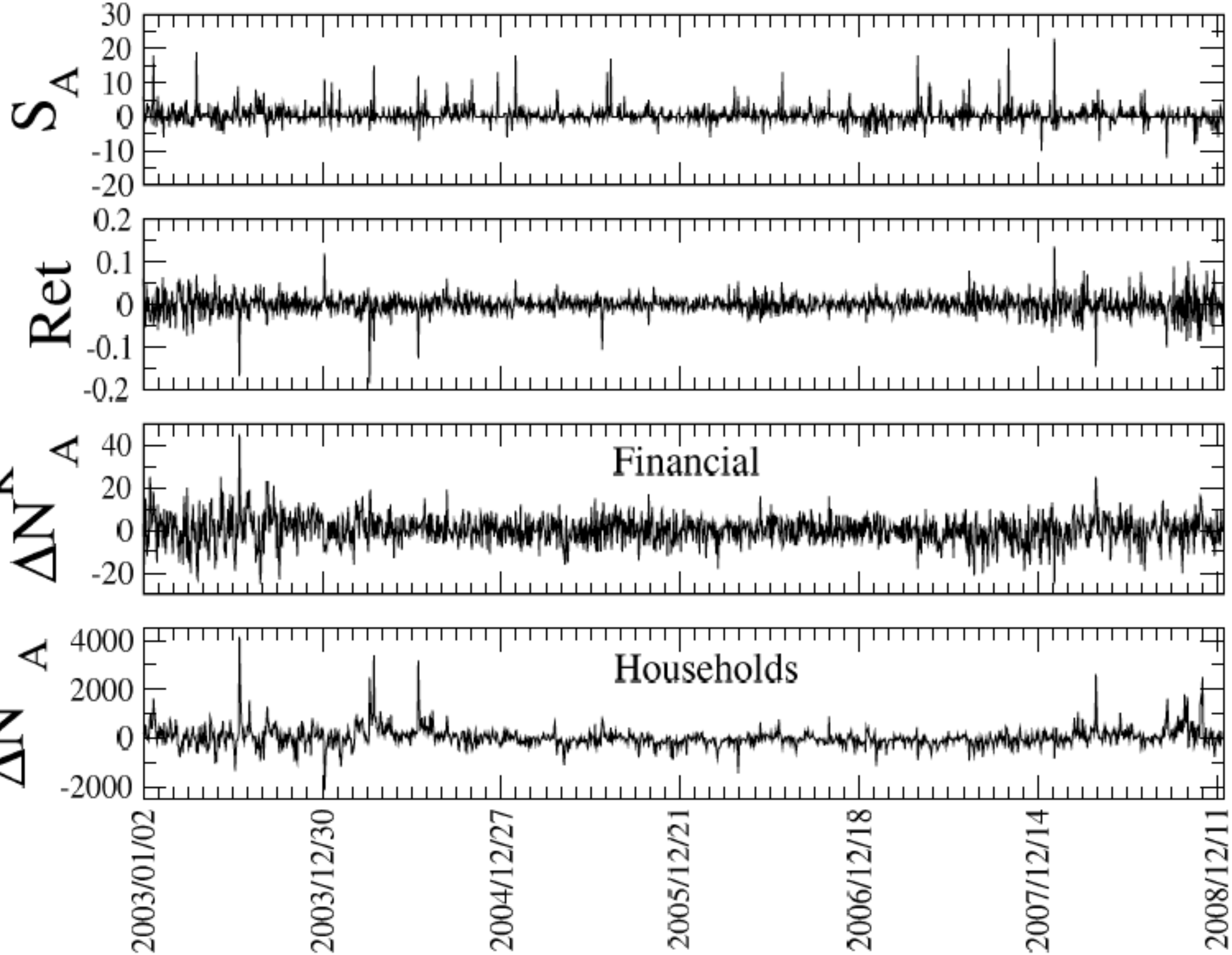
Absolute indicators

Sentiment

Return

Trading unbalance Financial

Trading unbalance Households



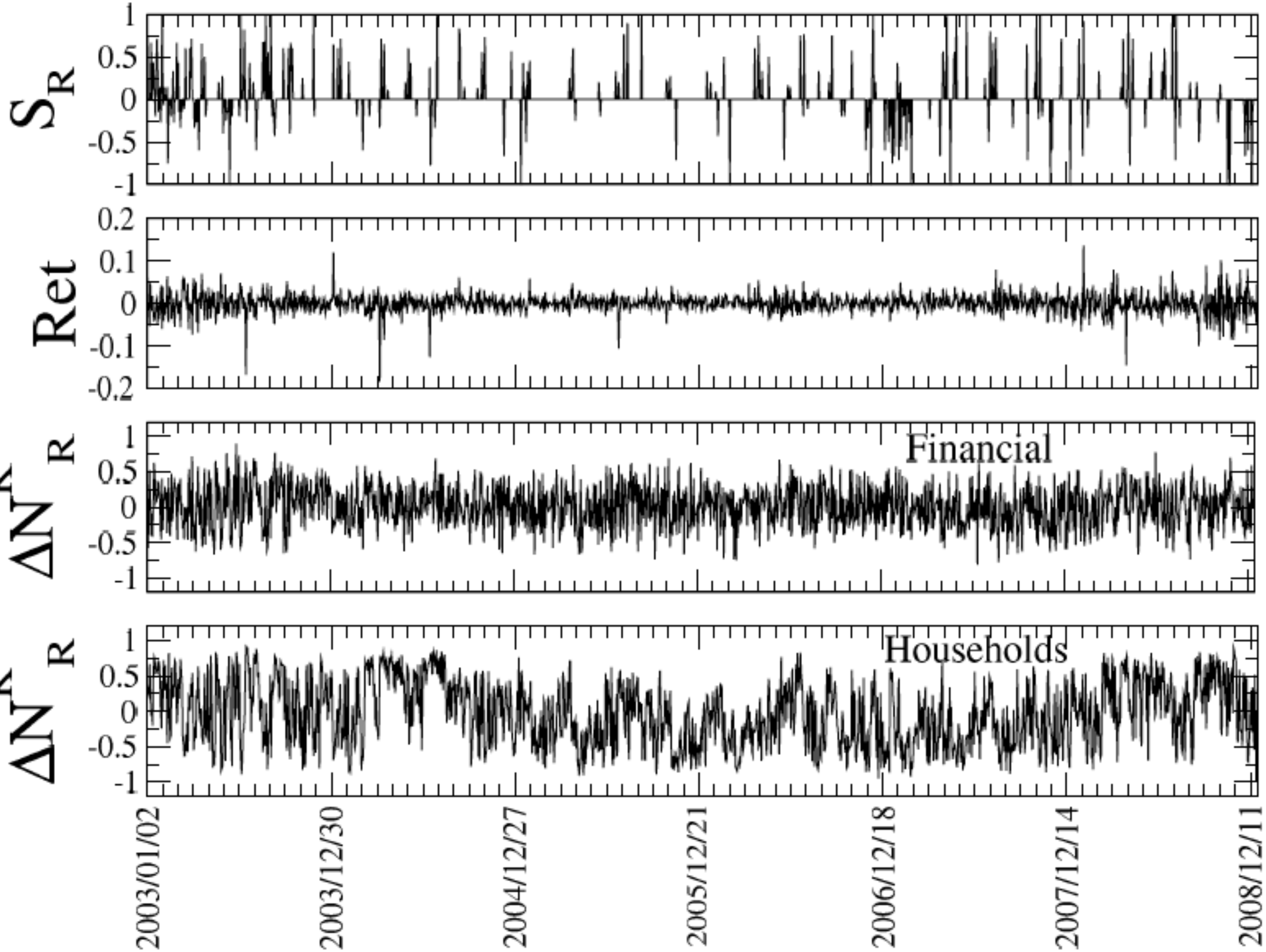
Relative indicators

Sentiment

Return

Trading unbalance Financial

Trading unbalance Households



Impact of news and volatility on market activity (synchronous indicators)

$$\widehat{N}^K(t) = \alpha_H \widehat{H}(t) + \alpha_{Vol} \widehat{Vol}(t) + \epsilon(t)$$

where the new variables are standardized versions of the original ones

Table 3. Summary of the results of the linear regression of equation (8) of the number N^K of trading investors vs. the news intensity signal H and the volatility proxy Vol. The number in parentheses is the 5–95% confidence intervals obtained using bootstrap analysis. The last two columns show the results of the partial correlation analysis.

Investor category	α_H	α_{Vol}	% variance of residual of N^K	$\rho(N^K, H\|Vol)$	$\rho(N^K, Vol\|H)$
Companies	0.271 (0.205,0.335)	0.517 (0.437,0.597)	51.8	0.309	0.534
Financial	0.195 (0.125,0.264)	0.479 (0.407,0.558)	63.8	0.207	0.461
Governmental	0.238 (0.164,0.303)	0.192 (0.119,0.262)	← 86.0	0.215	0.180
Non-profit	0.319 (0.249,0.394)	0.270 (0.199,0.344)	← 73.9	0.305	0.264
Households	0.226 (0.165,0.285)	0.627 (0.554,0.697)	41.4	0.289	0.651
Foreign org.	0.158 (0.094,0.224)	0.442 (0.374,0.517)	70.9	0.160	0.416

Endogenous information drives more the activity of companies, financial, and households investors. The case of governmental and non-profit investors is compatible with the opposite conclusion.

Impact of news and volatility on market activity (overnight indicators)

$$\hat{N}^K(t) = \alpha_H \hat{H}_O(t) + \alpha_{Vol} \hat{Vol}_O(t) + \varepsilon(t)$$

Table 6. Summary of the results of the linear regression of equation (8) of the number N^K of trading investors vs. the news intensity signal $H_O(t)$ of overnight news and the overnight volatility proxy $Vol_O(t) = 2\,|P_{open}(t) - P_{close}(t-1)|/(P_{open}(t) + P_{close}(t-1))$. The number in parentheses is the 5–95% confidence intervals obtained using bootstrap analysis. The last two columns show the results of the partial correlation analysis.

Investor category	α_{H_O}	α_{Vol_O}	% variance of residual of N^K	$\rho(N^K, H_O \| Vol_O)$	$\rho(N^K, Vol_O \| H_O)$
Companies	0.171 (0.109,0.233)	0.193 (0.142,0.242)	93	0.174	0.195
Financial	0.116 (0.061,0.177)	0.224 (0.163,0.288)	93	0.119	0.225
Governmental	0.133 (0.071,0.195)	0.057 (0.007,0.110)	98	0.133	0.057
Non-profit	0.150 (0.090,0.214)	0.077 (0.021,0.140)	97	0.150	0.078
Households	0.199 (0.130,0.268)	0.262 (0.198,0.333)	88	0.207	0.268
Foreign org.	0.127 (0.066,0.189)	0.195 (0.151,0.245)	94	0.129	0.196

The preeminence of volatility impact is still observed for companies, financial, and households investors but it is much weaker than in the synchronous case. The opposite case of governmental and non-profit investors is in the present case more pronounced.

Impact of news sentiment and return on the unbalance of market activity (absolute indicators)

$$\widehat{\Delta N}_A^K(t) = \alpha_{S_A}\widehat{S}_A(t) + \alpha_{Ret}\widehat{Ret}(t) + \epsilon(t)$$

Table 4. Summary of the results of the linear regression of equation (10) of the difference ΔN_A^K between buying and selling investors vs. the absolute sentiment indicator S_A and the stock return Ret. The number in parentheses is the 5–95% confidence intervals obtained using bootstrap analysis. The last two columns show the results of the partial correlation analysis.

Investor category	α_{S_A}	α_{Ret}	% variance of residual of ΔN_A^K	$\rho(N_A^K, S_A \vert Ret)$	$\rho(N_A^K, Ret \vert S_A)$
Companies	0.041 (−0.031,0.102)	−0.653 (−0.749,−0.572)	58.0	0.0528	−0.6463
Financial	0.008 (−0.055,0.069)	−0.576 (−0.651,−0.516)	66.9	0.0095	−0.5709
Governmental	−0.019 (−0.083,0.047)	−0.196 (−0.270,−0.125)	95.9	−0.0196	−0.1940
Non-profit	−0.026 (−0.116,0.056)	−0.220 (−0.303,−0.136)	94.8	−0.0263	−0.2178
Households	0.060 (−0.012,0.131)	−0.655 (−0.745,−0.568)	57.9	0.0781	−0.6474
Foreign org.	−0.006 (−0.072,0.058)	−0.504 (−0.571,−0.438)	74.4	−0.0072	−0.4999

The role of news sentiment is absent within error bars whereas price returns have a large impact. For all investors' categories the unbalance of market activity is anti-correlated with returns (contrarian behavior).

Conclusions

A key aspect of economic and financial systems is that they are heterogeneous open systems processing information.

The knowledge of information is providing marginal advantage and therefore its process of information aggregation is complex.

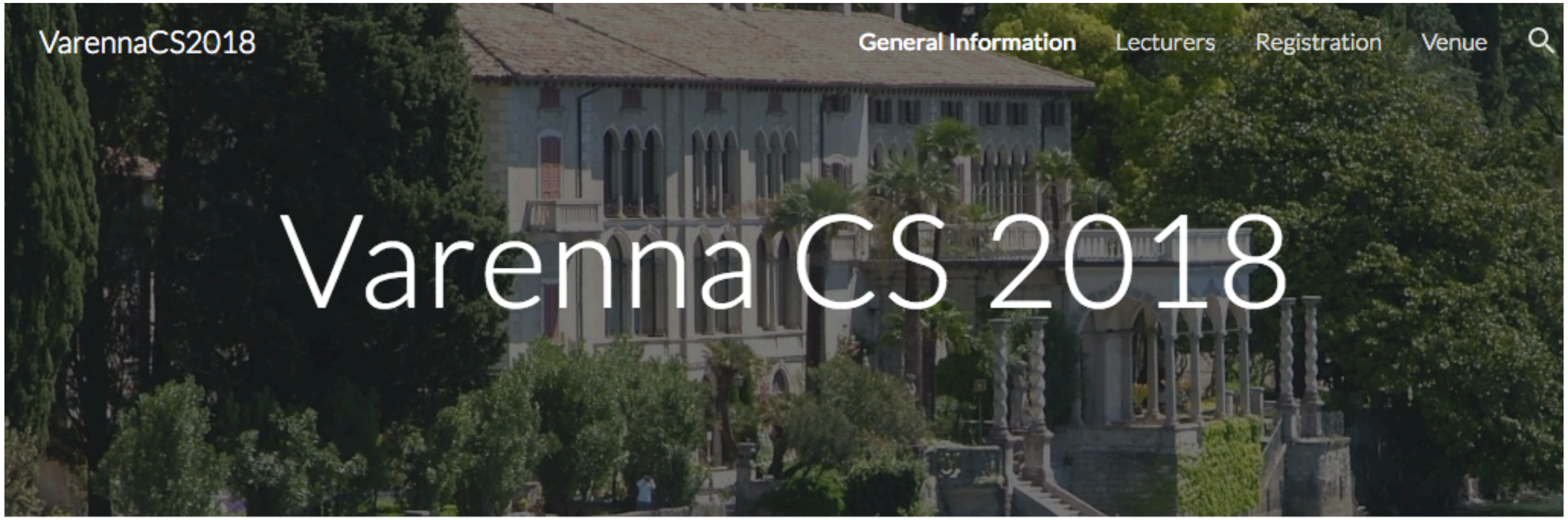
Tools from statistical and theoretical physics and non conventional data mining approaches are suggesting information filtering procedures that are essential for the understanding of these complex systems.

Trading activity of single investors are affected by both exogenous and endogenous news. Different categories of investors show different reactions to them. Discriminating between exogenous and endogenous news is difficult because the two sets are strongly interlinked.

Thank you

Observatory of Complex Systems

2018 Enrico Fermi School of Italian Physical Society



International school on

Computational Social Science and Complex Systems

Opening of the Course: Monday, 16 July 2018 - 9.00 a.m.

Closing of the Course: Saturday, 21 July 2018 - 12.30 p.m

Arrival day: Sunday, 15 July 2018

Topics

- Human behavior from ICT data
- Disease spreading, countermeasures, and large scale ABMs
- Science of success
- Reality mining
- Price formation, market microstructure
- Systemic risk, including cascading failures
- Networks in economics and finance
- Agent based models in economics and finance

1 https://sites.google.com/view/varennacs2018/general-information