Clustering dynamics through an emerging market crash in the global crisis 2007-2009

Diane Wilcox[1] and Tim Gebbie[1,2]

[1]School of Computational and Applied Mathematics,
University of Witwatersrand, South Africa,

[2]QT Capital Management

CCSS, Zurich, June 2009

Outline

Outline

A model for correlations in stock markets

Noh ansatz (2000)

A market is composed of several groups, within which the stock price fluctuations are correlated. This connects the spectral properties or price fluctuations and the structure of correlations:

$$X_i = g_{s_i}\eta_{s_i} + \sqrt{1 - g_{s_i}}\epsilon_i.$$

1. s_i denotes the cluster to which the i-th object belongs, given feature X_i
2. η_{s_i} denotes the synchronous variation of the cluster s_i
3. ϵ_i denotes the random noisy part of the features of object i
4. g_{s_i} denotes the similarity of objects within cluster s_i
5. $g_s = 1$ if all objects are identical and $g_s = 0$ if all objects are different.

A model for correlations in stock markets

Noh ansatz (2000)

A market is composed of several groups, within which the stock price fluctuations are correlated. This connects the spectral properties or price fluctuations and the structure of correlations:

$$X_i = g_{s_i}\eta_{s_i} + \sqrt{1 - g_{s_i}}\epsilon_i.$$

1. s_i denotes the cluster to which the i-th object belongs, given feature X_i
2. η_{s_i} denotes the synchronous variation of the cluster s_i
3. ϵ_i denotes the random noisy part of the features of object i
4. g_{s_i} denotes the similarity of objects within cluster s_i
5. $g_s = 1$ if all objects are identical and $g_s = 0$ if all objects are different.

A model for correlations in stock markets

Noh ansatz (2000)

A market is composed of several groups, within which the stock price fluctuations are correlated. This connects the spectral properties or price fluctuations and the structure of correlations:

$$X_i = g_{s_i}\eta_{s_i} + \sqrt{1 - g_{s_i}}\epsilon_i.$$

1. s_i denotes the cluster to which the i-th object belongs, given feature X_i
2. η_{s_i} denotes the synchronous variation of the cluster s_i
3. ϵ_i denotes the random noisy part of the features of object i
4. g_{s_i} denotes the similarity of objects within cluster s_i
5. $g_s = 1$ if all objects are identical and $g_s = 0$ if all objects are different.

A model for correlations in stock markets

Noh ansatz (2000)

A market is composed of several groups, within which the stock price fluctuations are correlated. This connects the spectral properties or price fluctuations and the structure of correlations:

$$X_i = g_{s_i}\eta_{s_i} + \sqrt{1 - g_{s_i}}\epsilon_i .$$

1. s_i denotes the cluster to which the i-th object belongs, given feature X_i
2. η_{s_i} denotes the synchronous variation of the cluster s_i
3. ϵ_i denotes the random noisy part of the features of object i
4. g_{s_i} denotes the similarity of objects within cluster s_i
5. $g_s = 1$ if all objects are identical and $g_s = 0$ if all objects are different.

A model for correlations in stock markets

Noh ansatz (2000)

A market is composed of several groups, within which the stock price fluctuations are correlated. This connects the spectral properties or price fluctuations and the structure of correlations:

$$X_i = g_{s_i}\eta_{s_i} + \sqrt{1 - g_{s_i}}\epsilon_i.$$

1. s_i denotes the cluster to which the i-th object belongs, given feature X_i
2. η_{s_i} denotes the synchronous variation of the cluster s_i
3. ϵ_i denotes the random noisy part of the features of object i
4. g_{s_i} denotes the similarity of objects within cluster s_i
5. $g_s = 1$ if all objects are identical and $g_s = 0$ if all objects are different.

A model for correlations in stock markets

Noh ansatz (2000)

A market is composed of several groups, within which the stock price fluctuations are correlated. This connects the spectral properties or price fluctuations and the structure of correlations:

$$X_i = g_{s_i}\eta_{s_i} + \sqrt{1 - g_{s_i}}\epsilon_i.$$

1. s_i denotes the cluster to which the i-th object belongs, given feature X_i
2. η_{s_i} denotes the synchronous variation of the cluster s_i
3. ϵ_i denotes the random noisy part of the features of object i
4. g_{s_i} denotes the similarity of objects within cluster s_i
5. $g_s = 1$ if all objects are identical and $g_s = 0$ if all objects are different.

A model for correlations in stock markets

Noh ansatz (2000)

A market is composed of several groups, within which the stock price fluctuations are correlated. This connects the spectral properties or price fluctuations and the structure of correlations:

$$X_i = g_{s_i}\eta_{s_i} + \sqrt{1 - g_{s_i}}\epsilon_i.$$

1. s_i denotes the cluster to which the i-th object belongs, given feature X_i
2. η_{s_i} denotes the synchronous variation of the cluster s_i
3. ϵ_i denotes the random noisy part of the features of object i
4. g_{s_i} denotes the similarity of objects within cluster s_i
5. $g_s = 1$ if all objects are identical and $g_s = 0$ if all objects are different.

Correlations and Emergence

Giada-Marsili log-liklihood function (2001)

$\{\mathcal{G}, \mathcal{S}\}$ denotes the parametrization $(\{g_s\}, \{s_i\})$ and the model assumes η_s and ϵ_i are zero-mean, unit variance processes. The probability density $P(X_i|\mathcal{G}, \mathcal{S})$ of observing the data X_i can be computed and from this, the likelihood of the parameters given the data: $P(\mathcal{G}, \mathcal{S}|X_i)$ can be deduced. This probability is only dependent on the Pearson correlation ρ_{ij} with the resulting likelihood function:

$$\mathcal{L}_c(\mathcal{S}) = \frac{1}{2} \sum_{s, n_s > 1} \left[\log \frac{n_s}{c_s} + (n_s - 1) \log \frac{n_s^2 - n_s}{n_s^2 - c_s} \right],$$

where n_s is the number of objects in cluster s and c_s is internal correlation of the s^{th} cluster,

$$c_s = \sum_{i=1, j=1}^{N} \rho_{ij} \delta_{s_i, s} \delta_{s_j, s}.$$

$\mathcal{L}_c(\mathcal{S})$ is maximal when

$$g_s^* = \sqrt{\frac{c_s - n_s}{n_s^2 - n_s}}.$$

The recursive merging algorithm

1. Begin with N clusters composed of singletons (for any s, $n_s = 1$).
2. At each step, merge the 2 clusters which result in the cost function $\mathcal{L}_c$ of the configuration $\mathcal{S}$ being minimised
3. Repeat step 2 until the single-cluster configuration is reached.

Simulating annealing approach

1. Sweep the lattice, generating new configurations by small random permutations of clusters
2. Accept a new cluster configuration (merge clusters) using Metropolis conditions
3. Repeat steps 1 and 2, decreasing 'temperature' in the Metropolis conditions, until likelihood is maximized

[KGV 1983, RGF 1990, GM 2001]

Novel genetic algorithm approach for $\mathcal{L}_c$

Each configuration is represented by a chromosome. The genes are the integer valued cluster indices and the genome length is the number of objects. The mutation rate is scaled by the current generation relative to the terminating generation.

1. *Mutation*: The number of clusters is randomly changed and objects are randomly re-assigned to new feasible clusters.
2. *Cross-over*: For a fixed number of clusters, generate random but feasible cross-overs.
3. Iterate 1 and 2 until fitness function reaches the convergence criterion or the terminating generation is reached.

Remarks: Implemented using a modification of the Matlab GA Toolbox.

Outline

After the dot-com bubble

America's current economic woes - from the collapse in share prices to the surge in bankruptcies - can be traced back to the biggest credit boom in its financial history... Without easy credit the stockmarket bubble could not have been sustained for so long, nor would its bursting have had such serious consequences. And unless central bankers learn their lesson, it will happen again. **Economist Sept 2002**

After the dot-com bubble

America's current economic woes - from the collapse in share prices to the surge in bankruptcies - can be traced back to the biggest credit boom in its financial history... Without easy credit the stockmarket bubble could not have been sustained for so long, nor would its bursting have had such serious consequences. And unless central bankers learn their lesson, it will happen again. **Economist Sept 2002**

CDO Market

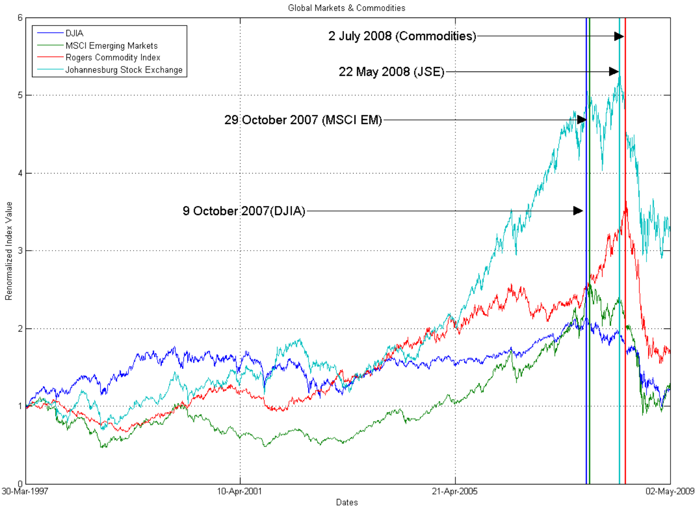
in billions of USD

2004	157.4
2005	271.8
2006	520.6
2007	481.6
2008	61.1
2009	TARP

Source:

Securities Industry and Financial Markets Association (SIFMA)

Global Financial Market Crash



Outline

Johannesburg Stock Exchange

- Started trading in 1887, electronic trading and settlement since 1997.
- 18^{th} Largest stock exchange with a capitalisation of R4.9 trillion (450 billion Euro)
- 405 Listed companies (cf. 3008 listed on LSE) and 1600 listed securities.
- 10 Largest companies constitute 60% of JSE.
- Resource stocks constitute 50% of stocks by market capitalization.
- 4 Big banks in SA (excl. SARB)
- SA has GDP of 277 Billion Euros, ranked 32 (2008, www.imf.org.za)

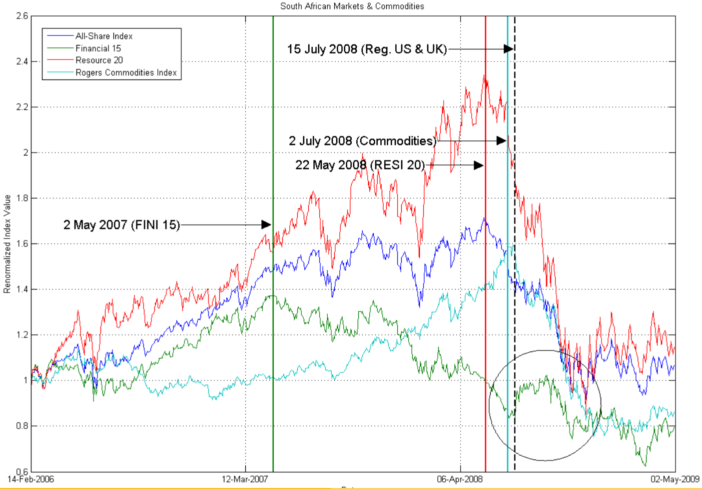
Is B^3E^2 the South African subprime?

South African Market Regulations

- Exchange Control Regulations, 1961, amended 1999, Currency and Exchanges Act, 1933
- The Financial Intelligence Act, 2001 (FICA) *[know you client]*
- The Financial Advisory and Intermediary Services Act, 2002 (FAIS) *[know you banker]*
- Protection of Constitutional Democracy Against Terrorist and Related Activities Act 33, 2004 (POCDATARA)
- The Securities Services Act, 2004 (SSA)
- The National Credit Act, 2005 (NCA)
- (new) The Banks Amendment Bill
- (new) The Companies Bill

Outline

The resources and financials factors in SA



Outline

Observations

- Markets coagulated on important event days.
- JSE often lagged the LSE.
- Periods of fragmentation quickly follow after coagulation.
- JSE has more larger groups, the LSE had more smaller groups.

JSE Market structure from correlations

Minimal Spanning Tree (MST)

Data: Top 40 of JSE and Top 100 of LSE by market capitalization, 1 Jan 09 - 1 Jan 09

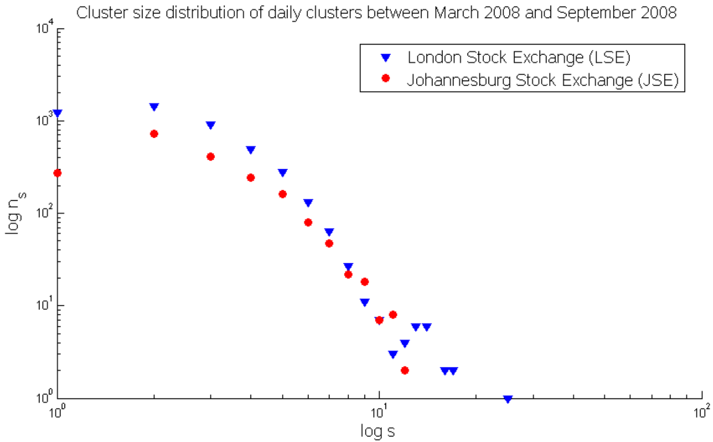
Four movies using an exponentially weighted moving average to estimate the covariance:

- JSE_D_LJSEOVER_SPIN_MARKET.avi
- JSE_D_LJSEOVER_SPIN_NOMARKET.avi
- LSE_D_LTOTMKUK_SPIN_MARKET.avi
- LSE_D_LTOTMKUK_SPIN_NOMARKET.avi

Possible explanations

- Macroeconomic sector effects may explain larger groups on the JSE.
- Pairs-trading is relatively more profitable on the JSE than LSE - this may be linked to an explanation for the dominance of smaller groups on the LSE.
- Any hedging or risk management strategy based on economic sectors would have been problematic.

Ex-market Mode Clustering on LSE & JSE



Outline

Can we anticipate possible crashes?

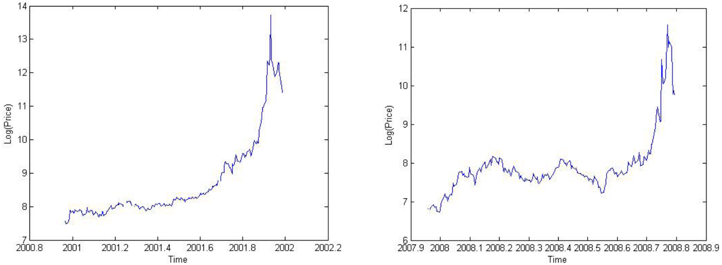


Figure: 20 December 2001 and 20 October 2008 crashes of ZAR/USD.

Comparison of predictions of t_c

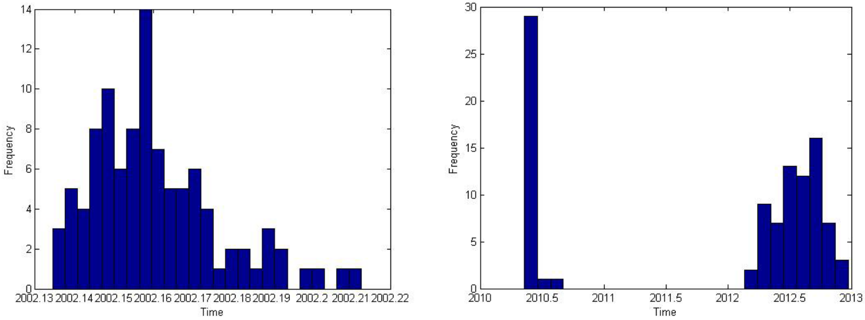


Figure: December 2001 and October 2008 crashes. Histograms of t_c using 95% of the data, 100 itera- tions and forecasting 1 day before the crash.

(see Honours Project of David Teagle)

Outline

Stable global money supply?

US DEBT! details below

MBS Market

in billions of USD

	Mortgage related	Asset Backed
1996	2 486.1	404.4
2000	3 565.8	1 071.8
2004	5 862.0	1 827.8
2005	7 127.7	1955.2
2006	8 452.8	2130.4
2007	8 931.4	2472.4
2008	8 897.3	2671.8

Source:
Securities Industry and Financial Markets Association (SIFMA)

col2: Includes GNMA, FNMA, and FHLMC mortgage-backed securities and CMOs, and CMBS, and private-label MBS/CMOs

US Money Supply

M1: assets that strictly conform to the definition of money: assets that can be used to pay for a good or service or to repay debt.

M2: M1 + savings deposits, time deposits less than $100,000 and money market deposit accounts for individuals.

M3: M2 + large time deposits, institutional money-market funds, short-term repurchase agreements, along with other larger liquid assets.

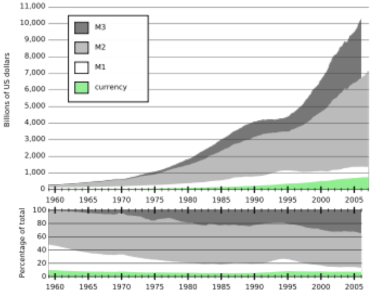


Figure: Wikipedia, see also http://www.federalreserve.gov

US Debt

2008 GDP (in million USD):

Global	USA	SA
60,689,812	14,624,660	277 (ranked 32)

Some 2008 figures:

- US Consumers spend 130% of income
- US current account deficit rises to 7% of GPP (800 billion USD)
- US Gov debt as percentage of GDP betw 61% & 73% (10.7 trillion USD)
  - excl. Fannie Mae and Freddie Mac (MBS debt 8.9 trillion USD)
  - more than 25% of this is foreign held (mostly China and Japan)

Source: www.imf.org

see also: http://www.federalreserve.gov [USA]

Summary

- Clustering was present but not stable.
- There was a drop in the number of small clusters, particularly on the JSE.

Further questions:

- Can we learn more about what the market agents were doing?
- Does the analysis tell us anything interesting for risk management?

Some references

- S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi, Optimization by simulated annealing, Science, Vol. 220, No. 4598, 671-680 (1983)
- RGF K. Rose, E. Gurewitz and G.C. Fox, Statistical mechanics and phase transitions in clustering, *Physical Review Letters,* Vol. 65, No. 8, 945-948 (1990)
- M. Blatt, S. Wiseman and E. Domany, Superparamagnetic clustering of data, *Physical Review Letters,* Vol. 76, No. 18, 3251-3254 (1996)
- J.D. Noh, *Physical Review E*, Vol. 61, 5981 (1) (2000)
- L. Giada and M. Marsili, Data clustering and noise undressing of correlation matrices, *Physical Review E*, Vol. 63, 061101-(1-8) (2001)
- M. Marsili, Dissecting financial markets: Sectors and states, *Quantitative Finance*, Vol. 2, 297-302 (2002)

A. Johansen and D. Sornette, Endogenous versus exogenous crashes in financial markets, arxiv: cond-mat/0210509

M. Marsili, Eroding market stability by proliferation of financial instruments, SSRN http://ssrn.com/abstract=1305174

S. Bornholdt, Expection bubbles in a spin model of markets: Intermittancy from frustration accross scales, *Internation Journal of Modern Physics* C, Vol. 12, No. 5, 667-674 (2001)

S. Bornholdt and F. Wagner, Stability of money: phase transitions in an Ising economy, *Physica A,* **316**, 453-468 (2002)