Kiel

Working Papers

Kiel Institute
for the World Economy

Phase Transition in the S&P Stock Market

by

Matthias Raddant
Friedrich Wagner

No. 1846 | June 2013

Web: www.ifw-kiel.de

Kiel Institute for the World Economy, Hindenburgufer 66, 24105 Kiel, Germany

Kiel Working Paper No. 1846 | June 2013

Phase Transition in the S&P Stock Market

Matthias Raddant and Friedrich Wagner

Abstract:

We analyze the stock prices of the S&P market from 1987 to 2012 with the covariance matrix of the firm returns determined in time windows of several years. The eigenvector belonging to the leading eigenvalue (market) exhibits in its long term time dependence a phase transition with an order parameter which can be interpreted within an agent model. From 1995 to 2005 the market is in an ordered state and after 2005 in a disordered state. We show that the influence of stocks on the market is changing and that this influence can be explained by trading volume and the stocks' beta.

Keywords: stock price correlations, CAPM, S&P500

JEL classification: G11, G12

Matthias Raddant

Kiel Institute for the World Economy
Hindenburgufer 66
24105 Kiel, Germany
Email: matthias.raddant@ifw-kiel

Friedrich Wagner

Christian-Albrechts-Universität zu Kiel
Institut für Theoretische Physik
und Astrophysik
Leibnizstraße 15
24098 Kiel, Germany
Email: wagner@theo-physik.uni-kiel.de

Phase Transition in the S&P Stock Market

Matthias Raddant *
Institut für Volkswirtschaftslehre, Universität Kiel
and Institute for the World Economy, IfW

Friedrich Wagner
Institut für Theoretische Physik, Universität Kiel

June 18, 2013

Abstract

We analyze the stock prices of the S&P market from 1987 until 2012 with the covariance matrix of the firm returns determined in time windows of several years. The eigenvector belonging to the leading eigenvalue (market) exhibits in its long term time dependence a phase transition with an order parameter which can be interpreted within an agent-based model. From 1995 to 2005 the market is in an ordered state and after 2005 in a disordered state.

1 Introduction

In this paper we analyze the structure of the U.S. stock market. We show that the influence of stocks on the market is changing and that this influence can be explained by trading volume and the stocks' beta.

The analysis of the structure of stock markets is dominated by two research approaches. The first one tries to explain the differences between the rates of return of stocks and relates to the seminal work by Sharp and Lintner [1, 2] and the CAPM model. The second point of view is that of the

*Institute for the World Economy, Hindenburgufer 66, 24105 Kiel, Germany, matthias.raddant@ifw-kiel.de

investor, hence, the choice of a portfolio and the analysis of risk. Both are related by the need to evaluate the comovement of stocks with each other and some index or market proxy.

The original version of the CAPM is in fact a one factor model which postulates that the returns of the stocks r should be governed by the market return r_M and only differ by the idiosyncratic β_i of the stock i, such that

$$r_i(t) = \alpha_i(t) + \beta_i r_M(t) + \epsilon_i(t). \tag{1}$$

Hence, stocks differ by the amount of volatility with respect to the market, and economic rational necessitates that higher stock volatility is compensated by higher absolute returns. Empirical tests of this model had rather mixed results and have let to the conclusion that beta values are not constant but time-varying [3]. The Fama-French model [4] extends this approach to a three-factor model incoporating size and book-to-market equity. Several other extension of the original models have been suggested, mostly building on some kind of conditional CAPM, where the entire model follows a first-order auto-regressive process [5]. The reasons for the change of the betas are manifold. They could change due to microeconomic factors, changes in the business environment, macroeconomic factors or due to changes of expectations [6]. [7, 8] also note that the non-normality of stock returns and especially conditional skewness can lead to distorted estimations of the CAPM.

In order to manage the risk of a portfolio one can derive optimal portfolio weights from the spectral decomposition of the covariance matrix of stock returns. Many studies show that the non-normality of stock returns can lead to an under-estimation of risk. Different approaches exist which try take the non-normality into account. A common way to describe the properties of stock comovements is to look at the eigenvalue spectrum of the correlation matrix. Random matrix theory suggests that a market that behaves like a one-factor model should result in one dominant eigenvalue. Both, the non-normality in the data and any other factors will result in deviations from this simplified model, see [9, 10].

Approaches which utilize the spectral properties of correlations matrices have their limits once the number of variables becomes large in relation to the number of observations. Networks approaches, which derive dependency networks from the correlations matrix can be useful, as long as one does not need explicit portfolio weights for each single stock [11, 12]. A related approach is to try to identify different states of the stock market, either by an analysis of the correlation matrix [13] or transaction volumes [14]. Recent studies show that the correlation structure in stock markets are rather volatile, and partly mirror economic and political changes [15]. [16] for example shows that a structural break seems to happen in the U.S. market around 2001. This strand of literature is also related to approaches from econometrics. Beile and Candelon [17] for example argue that correlations increase in times of crisis, which has profound implication for portfolio choice and hedging of risk. Other studies like [18] analyze if correlations in and between markets have increased due to more openness and tighter economic relations between countries.

Since financial markets tend to react very fast on any changes in the economy but also inhibit a lot of noise we found that a look at longer time horizons is a worthwhile contribution to the field, since many of the above mentioned studies look at time horizons of months or a few years. We found that the S&P 500 contains around 170 stocks with a history of price quotes of 25 years (the number drops rapidly with much more than this 25 years). We analyze the long-run development of the stocks influence upon the market. We derive both, a market index and the stocks influence, from the spectral decomposition of the covariance matrix. We show that for most of the period under consideration the market was in a ordered state, characterized by a disproportionate influence of stocks from the IT sector. While some changes in the market seem to happen in 1995, the collapse of his regime starts with the burst of the dot.com bubble. A disordered state is found around 2005, we will show that from here the market develops into a new (although weaker) ordered state where the driving sector is the financial industry.

The paper is organized in the following way. In section 2 we describe the subset of stocks in the S&P market used in this analysis. Section 3 contains the definition of market indices derived from the covariance matrix. We explain why we prefer the latter over the usually applied correlation matrix and that the average return r_{av} and market return $r_M(t)$ may be exact for large number of stocks, shown in section 4. Section 5 contains our results for phase transition and section 6 some conclusions.

2 Materials and Methods

The most important criterion for data selection consists in the length of the time series T of stock prices. Since the amount of listed firms changes in in time, only 289 firms in the S&P 500 remain, given our our time window from January 1987 to December 2011. In the present work we study the correlation matrix of firm returns. To a large part this matrix is a random matrix, where the errors of the quantities of interest is in the order of $\sqrt{N/T}$ [19]. Therefore one can in fact afford a reduction of N by the following criteria: We start out with the 500 stocks which are listed as part of the S&P500 index at the end of 2011. We drop all those which were not trading since January 1987. We then filter for illiquid stocks; we define stocks as illiquid of their price does not change for more than 7% of the trading days. We validate this selection by checking the daily trading volumes. Further we delete single stocks which price does not move for at least 10 days in a row (e.g. due to suspended trading).

Our final set of data comprises the stock prices of $N = 171$ firms at $T = 6312$ trading days in the time window 1987-2011. As a sign of the different sizes of the firms we will later also consider the yearly trading volume of the firms in this period. A disadvantage of our selection consists in the loss of the meaning of the index. Since this refers to to a changing set of 500 firms and may not be representative for our subset. This selection also lead to some form of selection bias, since failed enterprises are excluded from the analysis.

A frequently used tool to analyze financial markets consists in the study of the correlation matrix between the stock returns of a market. This matrix can be used in two ways. Its observation needs a certain time window t_W. For small window sizes (10-20 days in case of daily returns) the matrix is dominated by noise and a principal component analysis does not make any sense. In the first class of studies [20, 16] the noise is reduced by averaging the correlation matrix over the stocks. This means to replace the volatility of the average return r_{av} by the sum of firm returns.

On the other side, when choosing t_W in the order of few years the decomposition into eigenvectors may be meaningful. The correlation matrix possesses one large eigenvalue in the order of the number N of stocks [21]. The corresponding eigenvector can be used as a description of the market [21, 22]. The remaining eigenvalues are qualitatively similar to those of a random matrix with a Marĉenko-Pastur [19] spectrum. Nevertheless with the assumption of a specific model information can be extracted from this part of spectrum [10]. In general only eigenvalues separated by more than $\sqrt{N/t_W}$ from other values have a model independent meaning [23]. Therefore we concentrate in this paper on the long term time behavior of eigenvector of the market eigenvalue.

A daily market return $r_M(t)$ can be obtained by the scalar product of the stock returns with the eigenvector determined in an appropriate window. In this picture the eigenvector denoted by β_i describes the β coefficients relative to the market, as needed for a CAPM portfolio [24]. This interpretation is supported by the empirical result [21]; the β_i are positive and and distributed around one for $t_w > 4$y.

In an alternative description of the market the average return r_{av} may be used [20, 16]. In first question we investigate how much r_{av}, r_M and the index return differ from each other. Secondly we will investigate the time dependence of β_i to find evidence for a phase transition.

Transitions in physics are characterized by an order parameter m which vanishes in the disordered and is non zero in the disordered phase. m could be related to macroscopic or microscopic properties of the system. In general m is discontinuous at the critical point (first order). In special cases the transition is of continuous order with continuous m. Corresponding models of statistical physics near the critical region have been applied to financial markets [25, 26, 27]. They offer an explanation of the stylized facts [28] of the return. However, due to the universality the relation of the model parameters to economical quantities remains obscure. The models require fine tuning of the parameters to maintain the system close to the critical region. Since the system stays always in the disordered phase neither a micro- nor macroeconomical order parameter can directly observed.
In this study we look for a first order transition based on low risk aversion ($\beta_i > 1$) and the traded volume with a macro economical order parameter m.

3 Correlation Matrix and Pseudo Indices

3.1 Properties of the correlation matrix

The daily stock stock prices $S_i(t)$ for stock $i = 1, \ldots, N$ at day t may be converted into returns $r_i(t)$ by

$$r_i(t) = r_N \ln(S_i(t+1)/S_i(t)) \tag{2}$$

We use a normalization factor r_N given by $\sum_{i,t} r_i^2(t) = NT$ in order to avoid small numbers. A covariance matrix C at time τ can be constructed by selecting a time window of size t_w

$$C_{ij}(\tau) = \frac{1}{t_w} \sum_t I(t, \tau, t_w) r_i(t) r_j(t) \tag{3}$$

where I projects on the window $|t - \tau| < t_w/2$. For a statistical meaningful long-run C we found empirically that the window size t_w ought to be larger than 3 years. To compensate for the loss of time resolution we use overlapping bins by choosing time steps of $\Delta\tau$ in τ less than t_w. In almost all previous investigations (as discussed in the introduction) the Pearson correlation matrix has been used. This differs from equation 3.1 by subtracting from C the product of means of r_i and r_j and normalizing C by the square root of variances of r_i and r_j. Here mean and variance refer to the window. Using Pearson correlations consists in a loss of information, especially the time dependent variation of the volatility of the stocks due the volatility clusters. To summarize, the Pearson correlations would ignore the different volatilities of the stocks.

The matrix C can be written in a spectral decomposition as

$$C_{ij}(\tau) = \sum_{\nu=0}^{N-1} e_i^\nu(\tau)\, e_j^\nu(\tau) \lambda_\nu(\tau) \tag{4}$$

where λ_ν denote the eigenvalues and e_i^ν the eigenvectors of C. For sufficiently large t_w we have the following properties of C. Within our normalization of returns it possesses one large eigenvalue λ_0 in the order of $N/3$ and a few medium size eigenvalues. The remaining small ones are similar to a random matrix with a Marĉenko-Pastur spectrum [19]. The large eigenvalue and its eigenvector e_i^0 can be interpreted as a description of the market [21]. It can be used to define a market return $r_M(t)$

$$r_M(t) = \frac{1}{\sqrt{N}} \sum_i I(t, \tau, \Delta\tau) e_i^0(\tau) r_i(t) \tag{5}$$

In equation 5 we use the step size $\Delta\tau$ instead of t_w to avoid ambiguities in case of overlapping bins ($\Delta\tau < t_w$). For the CAPM model [24] one needs a so-called β_i-coefficient which describes how close a stock follows the market described by a reference return $\bar{r}$.

$$\beta_i = \frac{E[r_i \, \bar{r}]}{E[\bar{r}^2]} \tag{6}$$

We will use our market return r_M as for $\bar{r}$ instead of the common choice of the index, because the latter may not be representative for our data selection. Taking the time averages in equation 6 and using $\bar{r} = r_M$ from equation 4 we obtain the β_i coefficients

$$\beta_i(\tau) = \sqrt{N} e_i^0(\tau) \tag{7}$$

After resolving the sign ambiguity in e_i^ν we find the surprising fact that for $t_w > 4$ years all e_i^0 are positive. Due to the normalization $\sum_i (e_i^0)^2 = 1$ the $\beta_i(\tau)$ are distributed around a mean close to 1. Firms with large β_i follow the market more than others and also they influence the market more than firms with small β_i. For this reason we will refer to stocks with $\beta_i > \beta_c$ as the *market leaders*.

3.2 The market

A simple way of defining a market would be the average return

$$r_{av}(t) = \frac{1}{N} \sum_i r_i(t) \tag{8}$$

r_{av}^2 averaged over time corresponds to a correlation matrix averaged over the stocks which has also its defenders in the literature [20, 16]. From the returns defined by equations (5) and (8) we can calculate logarithms of pseudo indices L_{av} or L_M by the recursion

$$L_{M,av}(t+1) = L_{M,av}(t) + \frac{r_{M,av}(t)}{r_N} \tag{9}$$

The integration constant in equation (9) is fixed by the normalization

$$\sum_t L_{M,av}(t) = 0 \tag{10}$$

We compare these pseudo indices in figure 1, where they are plotted as index $S_0(t)$, written in the form $L_0(t) = \ln S_0(t) - l_0$ where l_0 ensures equation (10) also for S_0. As expected the market index L_M calculated with a time window of $t_w = 4$ years agrees with L_0 only qualitatively. The average pseudo index $L_{av}(t)$ is very similar to the market index. For easier comparison we plot the one year average of $L_{av}(t)$ (dashed line) with the averaged $L_M(t)$ (solid line). Especially in the years 2001-2008 the index exhibits considerably larger average returns than the pseudo indices. This effect interpreted as phase transition in to a stiff market has been detected by Kenett et al. [16], where correlations with the index have been subtracted from the stock correlations. An increase of the former correlation leads to smaller subtracted correlations after 2001. L_M and L_{av} are very close. This is somewhat surprising since we will see that $e_i^{(0)}(\tau)$ can change with time.

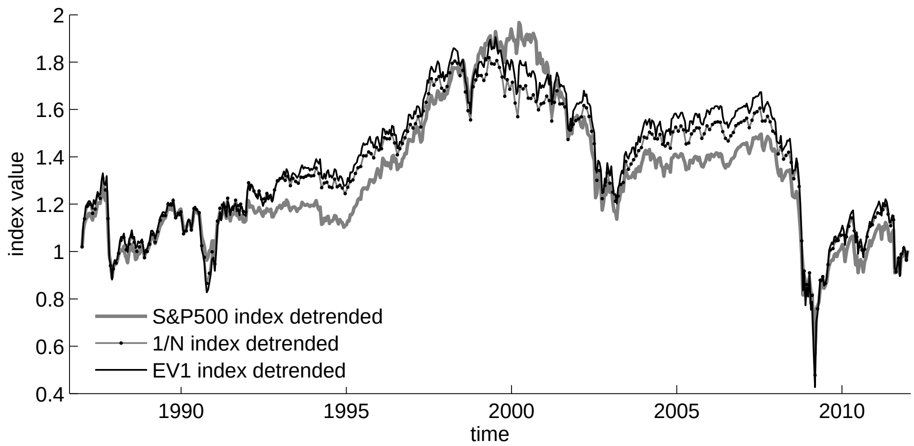

Figure 1: *S&P500 index (grey), and pseudo indices.* L_{av} *(solid) and* L_M *(dotted) are very similar. They show changing deviations from the real S&P index.*

For the squared difference Δ^2 of r_M and r_{av} given by

$$\Delta^2(\tau) = \frac{\sum_t I(t,\tau,t_w)(r_M(t) - r_{av}(t))^2}{\sum_t I(t,\tau,t_w) r_M^2(t)} \tag{11}$$

we derive in appendix A the inequality

$$\Delta^2(\tau) \leq (1-\bar{\beta})(\frac{2}{\lambda_0} \cdot trace(C) \ -1-\bar{\beta}) \tag{12}$$

with $\bar{\beta}$ the mean of β_i. The average correlation C_{av}

$$C_{av} = \frac{1}{N(N-1)} \sum_{i \neq j} C_{ij} \tag{13}$$

can be expressed by Δ^2 and the properties of the market component (see appendix A) up to terms of order $1/N$

$$C_{av} = < r_M^2 > (\Delta^2 + 2\bar{\beta} - 1) - \frac{1}{N} \tag{14}$$

In the next section we discuss that for large markets ($N \to \infty$) both empirically and in context of a model $\bar{\beta}$ approaches one and therefore Δ^2 vanishes. In this limit C_{av} corresponds to the volatility $< r_M^2 >$ of the market.

4 Dependence on the Market Size

The qualitative behavior of the correlation matrix suggests a decomposition of the returns r_i according to a stochastic volatility model [29]. $r_i(t)$ is the sum of the two products, of noise η and the market and of noise and the remaining contribution. The coefficients β, the coupling γ_M to the market and the ideosyncratic couplings γ_i are assumed to be constant in each window.

$$r_i(t) = \beta_i^0 \ \gamma_M \cdot \eta_m + \gamma_i \cdot \eta_i \tag{15}$$

The independent noise factors η have mean 0 and variance 1. The coefficients β_i^0 are normalized to $\sum_i (\beta_i^0)^2 = N$. For $t_w \to \infty$ C can be determined from the expectation values

$$C_{ij} = \beta_i^0 \beta_j^0 \gamma_M^2 + \delta_{ij} \ \gamma_i^2 \tag{16}$$

For large N one can solve the eigenvalue problem for C by a $1/N$ expansion (see appendix B). C has one large eigenvalue

$$\lambda_0 = N\gamma_M^2 + \langle \gamma^2 \rangle_\beta + (\langle \gamma^4 \rangle_\beta - \langle \gamma^2 \rangle_\beta^2)/N \tag{17}$$

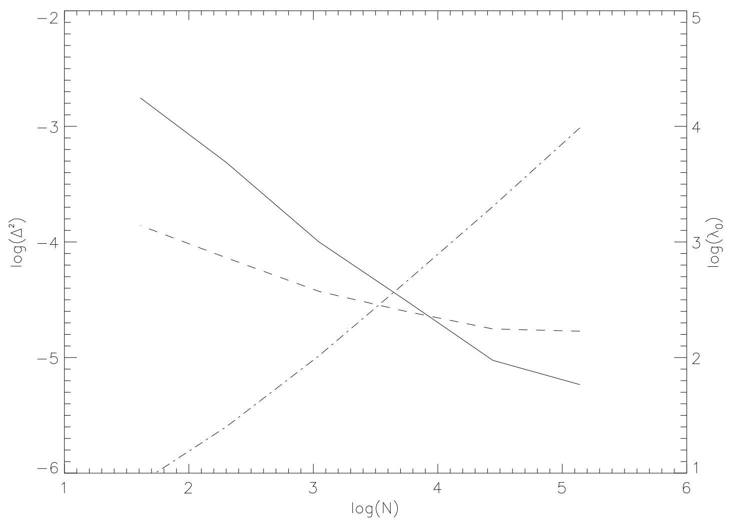

Figure 2: *The solid line shows the squared difference between market and average return as* $\ln(\Delta^2(N))$ *as function of* $\ln N$ *(left hand scale). The dashed dotted line gives* $\ln(\lambda_0)$ *as function of* $\ln N$ *(right hand scale) and the dashed line the variance of the distribution of* β_i *on a logarithmic scale.*

with an eigenvector corresponding to

$$\beta_i = \left[1 + (\gamma_i^2 - \langle \gamma^2 \rangle_\beta)/N\right] \beta_i^0 \tag{18}$$

$\langle a \rangle_\beta$ denotes an average over $a_i(\beta_i^0)^2$.

Neglected terms in equations (17) and (18) are of order $1/N^2$. Empirically the β_i are distributed around a mean $\bar{\beta}$ close to 1 and a variance σ_β decreasing with N. Therefor we can assume the model parameter β_i^0 equal to one. From the inequality (12) we see that the difference between r_M and r_{av} expressed by Δ^2 (see equation (11)) has to vanish in this limit.

In the following we investigate the behavior of the leading eigenvalue $\lambda_0 \propto N$, $\sigma_\beta \propto N^{-\gamma_1}$ and $\Delta^2 \propto N^{-\gamma_2}$ as function of N.

A finite window size t_w leads to deviations of the observed C from the ideal C of equation (16) in the order of $\sqrt{N/t_w}$. To minimize this systematic error we choose the maximum window size $t_w = T$. To have a varying N we adopt the following procedure. Sub-markets are defined by dividing the N_0 stocks into k groups with size $N(k) = N_0/k$ such that each group contains the same fraction of large and small firms as the full set. To improve the statistics we average λ_0, Δ^2 and σ_γ^2 over the groups. The result is presented in figure 2 showing these quantities as a function of $N(k)$ on a log-log scale. λ_0 (dashed dotted line) increases with N^α with a power of α close to 1. $\Delta^2(N)$ (solid line) and $\sigma_\gamma(N)$ (dashed line) exhibit a lesser decrease than the expected $1/N$ behavior. Since they also flatten out at large N, this indicates an influence of the finite observation time window. The observed values of Δ^2 are much smaller than the upper limit from the Schwartz inequality in equation (12). This indicates that the eigenvectors e_i^ν for $\nu > 0$ are almost orthogonal to a constant vector $e_i = 1/\sqrt{N}$ already at finite N, which implies equality of r_M and r_{av}.

To summarize, the data support a stochastic volatility model of a sum of market and preferences γ_i for individual stocks where for large N the stock returns couple to the market component in the same way. Therefore r_{av} can be taken as a description of the market. Since it determines the average correlation C_{av} the frequent use of C_{av} in the literature as a proxy for an index is empirically successful. Especially r_{av}^2 consists in a good estimator for the leading eigenvalue $\sqrt{\lambda_0)/N}$, since due to the law of large numbers the statistical error decreases with both $\sqrt{N}$ and $\sqrt{T}$. From the perturbation expansion given in appendix B we see that the leading eigenvalue and its vector can be more accurately determined than the remaining ones.

5 Time Dependence of β_i and Phase Transition

In the previous section the coefficients γ and β have been discussed on a long time scale. At smaller time scales they can be time dependent. To minimize the influence of the noise we produced by a finite observation window t_w, we chose a rather large window of $t_w = 7$ years. This implies that we can only detect long term changes in β_i. We diagonalize C for the S&P data in overlapping steps of 1 year.

The resulting time dependence of $\beta_i(\tau)$ is shown in figure 3 for the 60 firms with the smallest average traded volume. Before 2005 their beta values are relatively constant with values ≤ 1; as expected for small firms having less impact on the market. Around 2005 some firms with previously small β experience a drastic increase and become market leaders (firms with

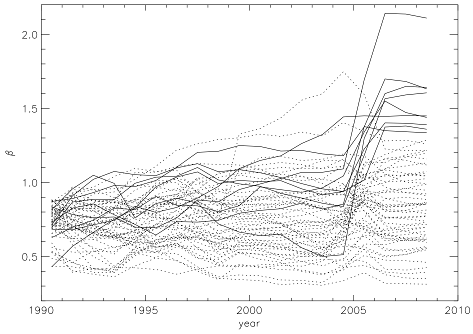

Figure 3: *Time dependence of $\beta_i(\tau)$ for the 60 firms with smallest trading volume. Solid lines indicate market leaders ($\beta_i > 1.3$) in 2008.*

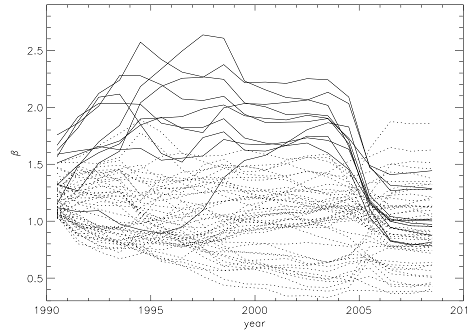

Figure 4: *Time dependence of $\beta_i(\tau)$ for the 60 firms with largest trading volume. Solid lines indicate market leaders ($\beta_i > 1.3$) in 2002.*

Sector name	# in sample	Sector name	# in sample
Consumer Discretionary	24	Consumer Staples	22
Energy	10	Financials	28
Health Care	18	Industrials	33
IT	20	Materials	14
Telecommunication Services	3	Utilities	0

Table 1: *List of* S&P *sectors and frequency in the data set, GICS classification*

$\beta_i(2008) > 1.3$ are indicated by solid lines). A different picture appears if we look at $\beta_i(\tau)$ for firms with large average traded volume shown in figure 4 for the 60 largest firms. Of course this set contains market leaders. Those in 2002 are denoted by solid lines. However, in 2005 they disappear in favor of new firms as in the case of small firms. Therefore in 2005 a reorganization of the market has happened. This depends on the the type of stocks. The 18 market leaders in the set of all firms with $\beta_i > 1.39$ in the year 2002 are listed in table 2 and in table 3 for the year 2008. The 2002 list contains dominantly large firms from the computing sector. After the transition in 2008 the list contains firms of all sizes spread over many sectors.

This behavior of β_i indicates a phase transition. A macroeconomic order parameter that explains this transition should take into account the (large) β_i values, the traded volume (see also [30]) and the sectors s of the firms. For

s we use the GICS classification scheme into $S = 10$ sectors given in table 1. The following function R describes the risk in each sector s due to the β coefficients

$$R(\tau, s) = \sum_{i \epsilon s} \theta(\beta_i - 1.0)\beta_i(\tau)\ v_i(\tau) \tag{19}$$

In the ordered state one specific $R(\tau, s_0)$ is large and the remaining R's are small. In the disordered state $R(\tau, s)$ is independent of s. A macroeconomic order parameter $0 \leq m \leq 1$ can be obtained by normalizing R

$$m(\tau, s_0) = \frac{S}{S-1}\left[\frac{R(\tau, s_0)}{\sum_{s'} R(\tau, s')} - \frac{1}{S}\right] \tag{20}$$

In the ordered state the 'wrong' order parameters m for $s \neq s_0$ are small and negative.

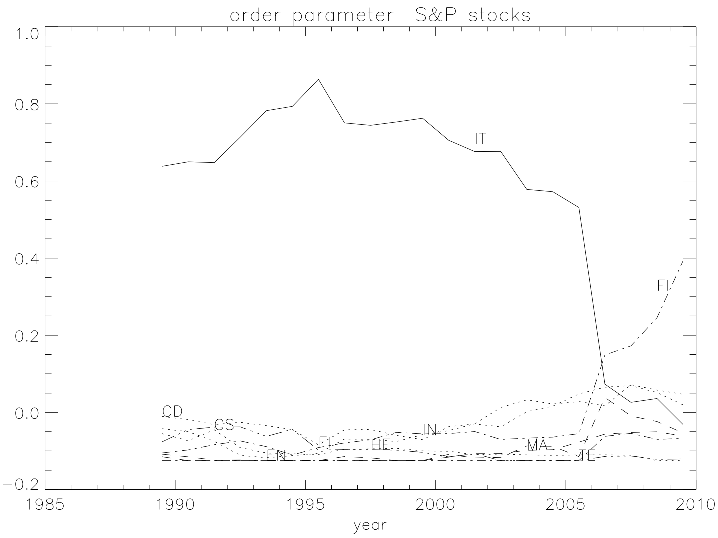

Figure 5: *The order parameter for various sectors. In the years 1992-2005 it is large for the IT sector and small for the remaining. Near 2010 the financial sector may give rise to an ordering,*

In figure 5 we show the order parameters $m(\tau, s)$ in time steps of one year. The $\beta_i(\tau)$ are calculated with a sliding time window of 5 years. Clearly m is large for the IT sector, whereas all others remain small. After 2005 $m(\tau, IT)$

Firm	Sector	β	Vol.	Firm	Sector	β	Vol.
TEXAS I.	IT	2.21	2700	HALLIBURTON	Energy	2.09	2393
ALTRIA	Cons. S.	2.04	7698	BANK OF A.	Finance	1.94	2985
HEWLETT-P.	IT	1.89	2055	MICROSOFT	IT	1.87	18900
APPLIED M.	IT	1.69	10005	AM. EXP.	Finance	1.69	1694
ORACLE	IT	1.66	11688	INTEL	IT	1.66	13160
PFIZER	Health	1.54	3258	ADOBE	IT	1.53	2376
LOWES	Cons. D.	1.51	1955	JOHNSON & J.	Health	1.50	1640
HERSHEY	Cons. S.	1.44	253	MERCK	Health	1.40	1535
APPLE	IT	1.39	3385	INTERN.BUS.	Industry	1.39	2125

Table 2: *List of market leaders with* $\beta(2002) \geq 1.39$. *Vol. gives the annual stock turnover in millions of $.*

Firm	Sector	β	Vol.	Firm	Sector	β	Vol.
HUMANA	Health	2.11	438	INTERN.BUS.	Industry	1.86	2020
WEYERH.	Finance	1.86	1721	CHUBB	Finance	1.80	591
VARIAN	Health	1.64	290	TEXTRON	Industry	1.63	351
CHEVRON	Energy	1.63	1514	DONNELLEY	Industry	1.60	309
A. DATA	Industry	1.55	639	TARGET	Cons. D.	1.51	2050
LOWES	Cons. D.	1.46	2440	C R BARD	Health	1.45	157
AVON	Cons. S.	1.45	728	MICROSOFT	IT	1.44	15697
G. PARTS	Cons. D.	1.44	254	PROG. OHIO	Finance	1.41	1159
CIGNA	Health	1.39	389	WASH. PST.	Cons. D.	1.39	5

Table 3: *List of market leaders with* $\beta(2008) \geq 1.39$. *Vol. gives the annual stock turnover in millions of $.*

decreases. Near the end of our time series an ordering towards the financial sector may be possible.[1]

There are different possibilities to relate this phase transition to agent behavior. A microeconomic order parameter could be obtained by incorporating a behavioral agent-based model for the market of the Kirman type [31]. The sector specific returns [32, 33] could be defined as

$$r_s(t) = \theta_s(\tau) \cdot \eta_s(t) \tag{21}$$

where θ_s is proportional to the ratio of noise traders and fundamentalist

[1]The dominance of the IT sector and the change around 2005 can also be found for smaller window sizes of down to 2-3 years, revealing one sharp peak around 1995/96. For much smaller time windows the influence of single events like the 1987 stock market crash or the 9/11 attacks become rather large and distort long-run trends.

agents that trade stocks in sector s and Gaussian noise η_s. θ_s is (on a longer time scale) time dependent, because the opinion of the agents changes. With suitable choice of the parameters in the asymmetric Kirman model [32] (using the bimodal version) this ratio can occasionally be large, and the system stays in this state for longer times. Such a situation cannot be distinguished empirically from a real phase transition. A microeconomic order parameter related to the herding effect would then be given by

$$m(\tau, s) = \frac{\theta_s(\tau)}{1 + \theta_s(\tau)} \tag{22}$$

An alternative model is the application of a Potts model with S states [34]. The β dependent interaction between agents is attractive if neighboring agents trade in the same sector. For strong enough dependent interaction the system will order with one sector dominating the others.

6 Conclusions

Our analysis of the market indices revealed that for sufficiently large samples of stocks and longer time horizons weighted indices differ only very slightly from any form of market average. This is a result of strong overall stock correlations and relatively stable long-run correlations structures, that we have shown by the analysis of the properties of the correlations matrix.

In our analysis of how the market is influenced by different stocks in the long-run, we have seen that the IT sector has played a dominant for a long time. The time dependence of the CAPM coefficients β exhibit a transition in 2005. This is connected with the disappearance of an macroeconomic order parameter for the IT sector. Despite our poor time resolution due to the window size this phase transition appears to be sharp. The transition lies between the crash of the dotcom bubble and the Lehmann desaster in 2008. In this time period we see no other pronounced effect in the index or in the stock prices.

A possible reason for a sharp transition may be the following: From 1990 to 2002 the stock prices experienced a steady increase. This led investors to buy in the most increasing sector, the IT sector. They minimized the risk by choosing only large firms. Disappointed by the crash of the dotcom bubble in 2001 they changed their investment strategy completely. Investments and trading volume became much more scattered over all segments. Figure 5 shows that later a (weaker) form of ordering took place by focusing on the financial sector.

Appendix

A Correlation Matrix

Denoting the time average as in equation (3) by $[\,]_{\tau,t_W}$ we get from equations (4) and (5) the average of r_M^2

$$[r_M^2]_{\tau,t_W} = \frac{\lambda_0(\tau)}{N} \tag{23}$$

Similarity we get for $[r_{av}^2]_{\tau,t_W}$ and $[r_M \cdot r_{av}]_{\tau,t_W}$

$$[r_{av}^2]_{\tau,t_W} = \frac{1}{N}\sum_{\mu=0}^{N-1} a_\mu^2(\tau)\lambda_\mu(\tau) \tag{24}$$

with

$$a_\mu(\tau) = \frac{1}{\sqrt{N}}\sum_i e_i^\mu(\tau) \tag{25}$$

a_0 corresponds to the mean value $\bar{\beta}$ of β_i over i.

$$[r_M \cdot r_{av}]_{\tau,t_W} = \frac{\lambda_0(\tau)}{N} \cdot \bar{\beta} \tag{26}$$

Inserting equations (23), (24) and (26) into Δ^2 from (4) we get

$$\Delta^2 = (1-\bar{\beta})^2 + \sum_{\mu>0} \frac{\lambda_\mu(\tau)}{\lambda_0(\tau)} a_\mu^2(\tau) \tag{27}$$

Since e_i^μ and e_i^0 are orthogonal for $\mu > 0$ we can write a_μ as

$$a_\mu(\tau) = \frac{1}{\sqrt{N}}\sum_i e_i^\mu (1 - \sqrt{N} e_i^0) \tag{28}$$

Applying the Schwartz inequality to (28) we get

$$a_\mu^2(\tau) \le 2(1-\bar{\beta}) \tag{29}$$

Together with $\sum_{\mu=0} \lambda_\mu = trace(C)$ this leads to the inequality (12) for Δ^2. Since the average C_{av} is a function $[r_{av}^2]_t$ insertion of Δ^2 into equation (24) leads to equation (14).

B Perturbation Expansion

The matrix C in equation (16) is a sum of two matrices. The first $C^0_{ij} = \beta^0_i \beta^0_j \gamma^2_M$ has one large eigenvalue $E_0 = \gamma^2_M N$ with an eigenvector $f^0_i = \beta^0_i/\sqrt{N}$ and $N-1$ degenerate zero eigenvalues with vectors f^μ_i with $\mu > 0$.These must satisfy only the orthogonality relation

$$(f^0, f^\mu) = 0 \tag{30}$$

with (a, b) denoting the scalar product. To obtain a complete basis we impose on f^μ in the $N-1$ dimensional subspace the following conditions with the second matrix $C^1_{ij} = \delta_{ij}\gamma^2_i$

$$(f^\nu, C^1\ f^\mu) = 0 \quad \text{for } \mu \neq \nu \text{ and } \mu, \nu \neq\ 0 \tag{31}$$

We apply standard second order Rayleigh Schrödinger perturbation theory [35] with C^1 as perturbation. For general matrices C^0 and C^1 using only the spectrum of C^0 and condition (30) we get up to order $1/E^2_0$ for the leading eigenvalue

$$\lambda_0 = E_0 + (f^0, C^1 f^0) + \frac{1}{E_0}[(f^0, (C^1)^2 f^0) - (f^0, C^1 f^0)^2] \tag{32}$$

and its eigenvector

$$e^0_i = f^0_i + \frac{1}{E_0}[(C^1 f^0)_i - (f^0, C^1 f^0) f^0_i] \tag{33}$$

The other eigenvalues require the in general complicated solution of equation (16) for f^μ_j. They are given by

$$\lambda_\nu = (f^\nu, C^1\ f^\nu) - \frac{1}{E_0}(f^0, C^1 f^\nu)^2 \tag{34}$$

Inserting the specific form of C^1 we obtain for λ_0 and β_i with

$$(f^0, C^1 f^0) = \frac{1}{N}\sum_i (\beta^0_i)^2 \gamma^2_i = \langle \gamma^2 \rangle_\beta \tag{35}$$

$$(f^0, (C^1)^2 f^0) = \frac{1}{N}\sum_i (\beta^0_i)^4 \gamma^2_i = \langle \gamma^4 \rangle_\beta \tag{36}$$

$$\lambda_0 = \gamma_M^2 N + \langle\gamma^2\rangle_\beta + \frac{1}{\gamma_M^2 N}[\langle\gamma^4\rangle_\beta - (\langle\gamma^2\rangle_\beta)^2] \tag{37}$$

$$\beta_i = \beta_i^0(1 + \frac{1}{\gamma_M^2 N}(\gamma_i^2 - \langle\gamma^2\rangle) \tag{38}$$

$\langle\ \rangle_\beta$ denotes the average over i weighted with $(\beta_i^0)^2$. Since neglected terms are of order $1/N^2$ these formulae describe λ_0 and β_i fairly accurate already for moderate N. Due to the degeneracy the general formalism [19] for the modification due to noise does not apply. Using the Wishart [36] formula we find the relative error in λ_0 due to a finite observation window T is of order $1/\sqrt{T}$ instead of $\sqrt{N/T}$ expected from [19] . The non-leading eigenvalue will be changed considerably if the spread of γ_i is small [23].

References

[1] J Lintner (1965), The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. Reviewof Economics and Statistics 47, 13–37.

[2] W Sharpe (1964), Capital Asset Prices: a Theory of Market Equilibrium Under Conditions of Risk. Journal of Finance 19, 425–442.

[3] T Bollerslev, RF Engle, JM Wooldridge (1988), A Capital Asset Pricing Model with Time-Varying Covariances. Journal of Political Economy 96(1), 116–131.

[4] EF Fama, KR French (1992), The Cross-Section of Expected Stock Returns, The Journal of Finance 47(2), 427–465.

[5] JN Bodurtha, NC Mark (1991), Testing the CAPM with Time-Varying Risk and Returns. The Journal of Finance 46(4), 1485–1505.

[6] T Bos, P Newbold (1984), An Empirical Investigation of the Possibility of Stochastic Systemic Risk in the Market Model. The Journal of Business 57(1.1), 35–41.

[7] CJ Adcock, MC Cortez, MJ Rocha Armada, F Silva (2012), Time varying betas and the unconditional distribution of asset returns. Quantitative Finance 12(6), 951–967.

[8] CR Harvey, A Siddique (2000), Conditional Skewness in Asset Pricing Tests. The Journal of Finance LV(3), 1263–1295.

[9] P Citeau, M Potters, JP Bouchaud (2001), Correlation structure of extreme stock returns. Quantitative Finance 1, 217–222.

[10] G Livan, S Alfarano, E Scalas (2011), Fine structure of spectral properties for random correlation matrices, Phys.Rev. E **84** 016113.

[11] DY Kenett, T Preis, G Gur-Gershgoren, E Ben-Jacob (2012), Dependency Network and node influence: Application to the study of financial markets, Int. J. Bifurcation Chaos 22, DOI: 10.1142/S0218127412501817.

[12] M Tumminello, T Aste, T Di Matteo, RN Mantegna (2005), A tool for filtering information in complex systems, Proceedings of the National Academy of Sciences of the United States of America 102(30):10421-10426.

[13] MC Münnix et al. (2012), Identifying States of a Financial Market, Scientific Reports 2(644).

[14] T Preis, JJ Schneider, HE Stanley (2011), Switching processes in financial markets. Proceedings of the National Academy of Sciences doi:10.1073/pnas.1019484108.

[15] DY Kenett DY, M Raddant, T Lux, E Ben-Jacob (2012), Evolvement of Uniformity and Volatility in the Stressed Global Financial Village. PLoS ONE 7(2): e31144.

[16] DY Kenett, Y Shapira, A Madi, S Bransburg-Zabary, G Gur-Gershgoren, E Ben-Jacob (2011), Index Cohesive Force Analysis Reveals That the US Market Became Prone to Systemic Collapses Since 2002. PLoS ONE 6(4): e19378.

[17] M Beine, B Candelon (2011), Liberalization and stock market co-movement between emerging economies. Quantitative Finance 12(2): 299312.

[18] N Ahlgren, J Antell J (2010), Stock market linkages and financial contagion: A cobreaking analysis, The Quarterly Review of Economics and Finance 50, 157–166.

[19] VA Marĉenko, LA Pastur (1967), Distribution of eigenvalues for some sets of random matrices, Math.USSR-Sb **1**(1967)457.

[20] L Borland and Y Hassid (2010), Market Panic on different Time-scales arXiv:010.4917v1.

[21] L Laloux, P Cizeau, J Bouchaud and M Potters (1999), Noise Dressing of Financial Correlation Matrices, Phys.Rev.Lett.**83** 1467.

[22] S Galluccio, JP Bouchaud, M Potters (1998), Rational decisions, random matrices and spin glasses, Physica **A 259**:449–456.

[23] Z Burda, A Görlich, A Jarosz, J Jurkiewicz (2004), Signal and noise in correlation matrix, Physica **A 343**(2004)295.

[24] F Black, MC Jensen, M Scholes (1972), The Capital Asset Pricing Model. Some Empirical Tests in Studies in the Theory of Capital Markets M Jensen ed.:New York, Praeger Publishers.

[25] R Cont, JP Bouchaud (2000), Herd behaviour and aggregate fluctuations in financial markets. Macroeconomic Dynamics **4** 170.

[26] D Stauffer, D Sornette (1999), Self-organized percolation model for stock market fluctuation. Physica **A 271** 496-506.

[27] S Bornholdt (2001), Expectation bubbles in a spin model of markets. Int. Journal of Mod. Physics **C 12** 667-674.

[28] T Lux, M Ausloos (2002), Market fluctuations I: scaling, multi-scaling and their possible origins, In: A Bunde, J Kopp and H Schellhuber(eds.) Theory of Desaster, Springer, Berlin, p.373.

[29] R Cont (2001), Empirical properties of asset returns: stylized facts and statistical issues, Quantitatve Finance **1** 223.

[30] ID Dichev, K Huang, D Zhou (2012), The Dark Side of Trading. Emory Law and Economics Research Paper No. 11-95.

[31] A Kirman (1993), Ants, rationality and recruitment, Quart. Jour.of Economics **108** 137-156.

[32] S Alfarano, T Lux, F Wagner (2005), Estimation of agent-based models: The case of an asymmetric herding model. Comp. Economics **26** 19-49.

[33] F Wagner (2006), Application of Zhangs square root law and herding to financial markets, Physica **A 364** 369-384.

[34] FY Wu (1982), The Potts model, Rev.Mod.Phys.54(1):235268.

[35] See any textbook on quantum mechanics, e,g, A Messiah (1967), Quantum Mechanics Volume I, North Holland.

[36] J Wishart (1928), The generalized product moment distribution in samples from a normal multivariate population, Biometrica **20A**(1928), 32-52.