Kiel
Working Papers

Kiel Institute
for the World Economy

Transitions in the Stock Markets of the US, UK, and Germany

by Matthias Raddant and
Friedrich Wagner

No. 1979 | December 2014

Web: www.ifw-kiel.de

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

Kiel Working Paper No. 1979 | December 2014

Transitions in the Stock Markets of the US, UK, and Germany

Matthias Raddant and Friedrich Wagner

Abstract:

In this paper we analyze transitions in the stock markets of the US, the UK, and Germany. For all this markets we find that while the markets were focused on stocks from the IT and technology sector around the year 2000, this focus has vanished and the markets have mostly moved towards a focus on stocks from the financial sector. This development is paralleled by changes in the returns distributions and the tail exponent. We show that we can extend the concept of beta values to systematically describe a risk measure for stocks from different sectors of the economy. This slowly varying sector specific risk measure describes ordered states in the market and identifies sectors which show concentration of market risk.

Keywords: stock price correlations - financial risk - CAPM

JEL classification: G11, G12

Matthias Raddant

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

Friedrich Wagner

Institute of Theoretical Physics
Kiel University
24098 Kiel, Germany
Email: wagner@theo-physik.uni-kiel.de

Transitions in the Stock Markets of the US, UK, and Germany

Matthias Raddant

Institute for the World Economy and Department of Economics, Kiel University
Kiellinie 66, Kiel 24105, Germany, matthias.raddant@ifw-kiel.de,

Friedrich Wagner

Institute of Theoretical Physics, Kiel University
Leibnizstr. 19, Kiel 24098, Germany,
wagner@theo-physik.uni-kiel.de

This version: December 2014

In this paper we analyze transitions in the stock markets of the US, the UK, and Germany. For all this markets we find that while the markets were focused on stocks from the IT and technology sector around the year 2000, this focus has vanished and the markets have mostly moved towards a focus on stocks from the financial sector. This development is paralleled by changes in the returns distributions and the tail exponent. We show that we can extend the concept of beta values to systematically describe a risk measure for stocks from different sectors of the economy. This slowly varying sector specific risk measure describes ordered states in the market and identifies sectors which show concentration of market risk.

INTRODUCTION

In this paper we analyze transition in the stock markets of the US, the UK and Germany. For all this markets we find that while the market was focused on stocks from the IT and technology sector around the year 2000, this focus has vanished and the markets have moved towards a focus on stocks from the financial sector. This development is paralleled by changes in the returns distributions and the tail exponent.

The analysis of the structure of stock markets has long been dominated by the discussion around different versions of a CAPM model [20, 27]. The original version of the CAPM is in fact a one factor model, which postulates that the returns r_i of the stocks should be governed by the market return r_M and only differ by the an idiosyncratic component β_i for each stock i, such that

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

In this setting α can also be interpreted as the risk free rate of interest. Hence, stocks differ by the amount of volatility with respect to the market. When we take the point of view of a risk-avers investor, this also means that stocks with a high volatility in returns should have a greater average returns than stocks with low volatility.

Empirical tests of this model had rather mixed results and have let to the conclusion that beta values are not constant but time-varying, see [8]. The Fama and French [13] model extends this approach to a three-factor model, incorporating firm size and book-to-market ratio. 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 autoregressive process, see [7]. The reasons for the change of the betas are manifold. They could change due to microeconomic factors, the business environment, macroeconomic factors, or due to changes of expectations, see, e.g., [6]. [1] and [14] also note that the non-normality of stock returns and especially conditional skewness can lead to distorted estimations of the CAPM.

An different approach to the analysis of asset returns is to try to identify different states of stock markets, either by an analysis of the correlation matrix like in [23], or by the analysis of transaction volumes as in [25]. The properties of the correlation matrix of asset returns have been analyzed for example by [18, 21, 22, 24]. Recent studies like [17] show that the correlation structure in stock markets are rather volatile, and partly mirror economic and political changes. [16] for example shows that a structural break seems to happen in the US market around 2001. This strand of literature is also related to approaches from econometrics. [4] for example argue that correlations increase in times of crisis, which has profound implication for portfolio choice and hedging of risks.

In the analysis of the covariance matrix of asset returns and the determination of beta values, or, more generally speaking, its spectral decomposition, one is faced with the problem of noise due to finite time series length and possibly very large N. If one assumes that the betas are time-varying, then there is a trade-off in the selection of the window size. Shorter time windows might capture the dynamics of the betas, but the estimated betas will have a high level of noise, and vice versa. Also, on the one hand, one can limit the number of stocks in the analysis to improve the accuracy in determining (especially) the non-leading eigenvalues. Our approach, on the other hand, is to allow for a relatively large N, since we are mostly interested in the eigenvector corresponding to the leading eigenvalue. This approach also allows us to apply

	US S&P500	UK FTSE350	Germany CDAX
period	1987–2011	1997–2013	1999–2013
T	6321	4294	3691
N	172	132	122
sector			
Energy	10	5	1
Materials	14	7	8
Industrials	33	30	35
Cons. Discr.	24	26	22
Cons. Staples	22	10	7
Health	18	5	14
Financial	28	37	8
Technology/IT	20	4	21
Telecom.	3	4	3
Utilities	0	4	3

TABLE I. Summary statistics of the data sets

perturbation theory to describe the magnitude of error in the determination of stock betas.

This paper is a continuation of the approach presented in [26], where we showed that the US stock market shows transitions between ordered to unordered times, which has similarities to a phase transition. In this paper we extend the analysis to include also the UK and the German stock market, and we relate the transitions to changes in the return distributions.

We find that transitions took place in all three stock markets, although the German markets seems to be less ordered. Changes in the returns distribution hint at the possibility that changes in the recent years are also influenced by a change in trading behavior, while earlier changes might be more influenced by changing investor behavior.

The paper is organized as follows. In the next section we briefly describe the data sets, before we describe the methodology to analyze the covariance matrix of the stock returns. After this, we show the transitions in the markets based on a risk measure. The last section aims to present possible explanations for the observed changes before we conclude.

MATERIALS AND METHODS

The data sets

For our analysis we use data from Thompson Reuters on the closing price of stocks which were continuously traded with sufficient volume throughout the sample period and had a meaningful market capitalization [31]. For the US we choose stocks which are part of the S&P 500 stock index. For the UK the stocks in our sample are listed in the FTSE350, the German stocks are all part of the CDAX (and are with very few exemptions also listed in the MDAX, SDAX or DAX30). The size of the market in the US allows us to collect a time series corresponding to 25 years of data. For the European markets it is not possible to analyze such a long time horizon, since not enough stocks have been traded for such a time span. Further we collected the sector classification of the firms, we used the GICS classification for the US market and the (for our purposes practically identical) TRBC classification from Thompson Reuters for the European markets. Table I summarizes the statistics and gives an overview over the sector breakdown.

Analysis Method for Correlations

Stock markets can be analyzed by the study of the correlation between the returns of the participating firms. There are N firms indexed by $i = 1, \cdots, N$. We use as returns $r_i(\tau)$ the log of the price ratio between consecutive days τ in a range T_0 in the order of 20 years. The returns are normalized by $\sum_{\tau=1,T_0} \sum_{i=1,N} r_i^2(\tau) = NT_0$. For the covariance matrix C we consider time windows of size T centered at time t. C is given by

$$C_{ij}(t) = \langle r_i r_j \rangle_{T,t} \tag{2}$$

with the abbreviation for the time average

$$\langle A \rangle_{T,t} = \frac{1}{T} \sum_{\tau=t-T/2}^{\tau<t+T/2} A(\tau) \tag{3}$$

for any observable A. When C is derived from the returns of many stocks in a long time windows $T \propto T_0$, one usually observes that the matrix C has one large eigenvalue λ_0 in the order of N with a corresponding eigenvector that we denote β_i. All β_i have the same sign and can be chosen positive. They are normalized by $\sum_i \beta_i^2 = N$. The remaining eigenvalues are in the order of 1, except few outliers. This first eigenvector can, for example within the framework of a principal component analysis, be interpreted as the market. This means that this eigenvector can be interpreted as the weights of the single stocks within the market factor. Hence, a market return r_M can be defined by the the projection of r on β within a time window centered at t

$$r_M(t) = \frac{1}{N} \sum_i \beta_i r_i(\tau) \tag{4}$$

Due to the relation

$$\beta_i = \frac{< r_i r_M >_{T_0,t}}{< r_M^2 >_{T_0,t}} \tag{5}$$

the components of the leading eigenvector are β-coefficients in a CAPM approach (leaving out the risk-free interest rate). With $T = T_0$ we would have only one vector β_i centered at time $(T_0/2)$. A time dependence of β can be achieved by using a moderate time window T (in the order of years). A useful criterion for determining the window size is to increase T until all the betas are positive for all t. To derive meaningful betas we assume that the return follows a stochastic volatility model (see, e.g., [2, 28]): The returns are the product of a noise factor and a slowly varying stochastic volatility factor. The latter should be considered as constant over the window size T. Then eq. (2) corresponds to an average over the noise with a statistical error depending on the properties of r_i.

As first example we consider $r_i(\tau) = \gamma\eta_{i\tau}$ with an i.i.d. Gaussian noise η. For a finite T we obtain a Marcenko-Pastur spectrum [22] spread over an interval $\gamma^2 1 \pm \sqrt{q})^2$ (instead of the degenerate eigenvalue γ^2). For $N \sim 400$, which would be the number of only the most important stocks within most stock markets, a time window of only a few years would lead to a prohibitive large uncertainty. However, this model cannot account for the occurrence of one large eigenvalue.

This can be reproduced by the second example with $r_i(\tau) = \gamma_i\eta_\tau$. In this model all stocks follow the market described by Gaussian noise. For $T \to \infty$ the covariance matrix C has one eigenvalue $\lambda_0 = \sum_i \gamma_i^2$ with eigenvector $\beta_i \propto \gamma_i$ and $N-1$ zero eigenvalues. At finite T the eigenvectors and the zero eigenvalues are unchanged. λ_0 is multiplied with a χ^2 distributed number with mean 1 and variance $2/T$. To describe the observed spectrum of small eigenvalues we consider a second process that leads to an additional additive component C_{1ij} in C. We assume market dominance in the sense that γ^2 is of order N and $(\gamma, C_1^k\gamma){=}A_k\gamma^2$ with constants A_k is of order 1. Perturbation theory for large N, see the appendix, shows that C_1 does not change λ_0 and β_i up to $1/N$ contributions. The remaining eigenvalues are strongly dependent on the noise. Only their sum is given by trace(C)-λ_0. $r_A^2(t)$ =trace$(C)/N$ also has a small error. r_A^2 corresponds to r^2 averaged over i and τ in the window.

$$r_A^2(t) = \frac{1}{N}\sum_i \langle r_i^2\rangle_{T,t} \tag{6}$$

λ_0 determines the size of the market return $< r_M^2 >$ via

$$\langle r_M^2\rangle_{T,t} = \frac{\lambda_0(t)}{N} \tag{7}$$

Equation (5) can also be utilized for any other kind of asset return. For example we can add the returns of the gold price to r_i in order to see whether this follows the market.

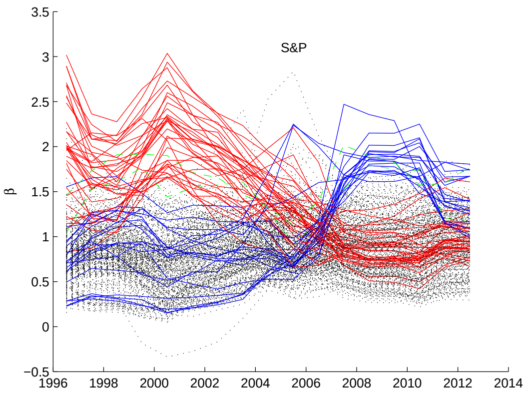

FIG. 1. Time dependence of β_i for 356 stocks of the S&P market. The 35 stocks with largest β in 1998-2002 are shown in red, the 20 largest in 2007-2010 in blue.

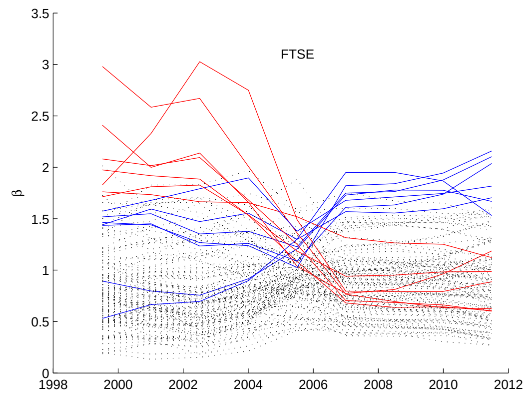

FIG. 2. Time dependence of β_i of the british FTSE market. The 7 stocks with largest β in 1998-2002 are shown in red, the 7 largest in 2007-2010 in blue.

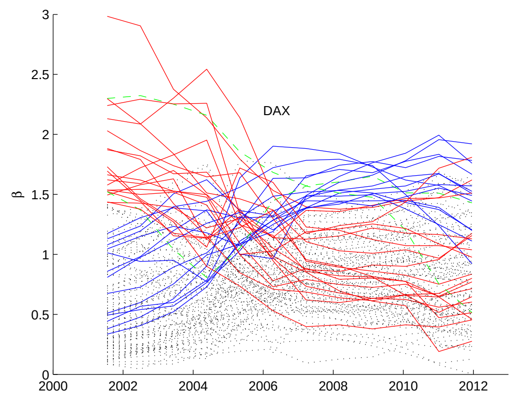

FIG. 3. Time dependence of β_i for 122 stocks of the german DAX market.The 20 stocks with largest β in 1998-2002 are shown in red, the 15 largest in 2007-2010 in blue.

The shape parameter of the returns distribution

In order to analyze changes in the distribution of the stock returns we will estimate the tail parameter of its pdf $f(r)$. We characterize f by a Pareto-Feller distribution [15] with f depending only on r^2 and finite $f(0)$. The two parameters are a scale parameter r_0 and a tail index $\alpha > 2$. It is given by

$$f(r) \propto \left(1 + \frac{r^2}{(\alpha-2)r_0^2}\right)^{-(\alpha+1)/2} \qquad (8)$$

Performing fits with limited statistics α and r_0 are strongly correlated. Therefore we fix r_0 by the condition $r_0^2 = E[r^2]$.

To summarize, for the empirical analysis of C in the next section we make the following assumptions: From the market hypothesis we can establish the leading eigenvector of C as CAPM β-coefficients. By the SVM assumption the time average in eq. (2) corresponds to an average over the noise. Making the market dominance assumption the error on λ_0 and β_i is of the order $1/N, \sqrt{2/T}$.

TRANSITION OF THE MARKETS IN 2006

We apply our approach to 172 stocks from the S&P market 132 stocks of the British FTSE market and to 122 stocks from the German market. To obtain the possible minimum window size T we look at the large eigenvalue $\lambda_0(t)$ and the corresponding eigenvector. As the criterion we use the presence of (only) positive values of $\beta_i(t)$. In this way we find for T a value of roughly 3 years for all markets. For a better visualization of the time variation we use overlapping windows by varying t in steps of years.

The $\beta-$coefficients derived from the first eigenvector of C for the S&P market are shown in figure 1. Except one case around 2001 they are all positive. Some of the stocks exhibit a substantial time variation with a transition around 2006. Stocks with large β during the years 1998-2002 (this time interval is called ITB for IT bubble hereafter) change to small β values around 2006, their values remain low in 2007-2010 (this time interval is called FB for the finance bubble hereafter). Vice versa those stocks with a large β in the finance bubble exhibit small values before 2006. A similar effect occurs also for the FTSE market (shown in figure 2) and the German market (shown in figure 3).

A more detailed characterization can be obtained by considering the sector s out of the GICS/TRBS classification for all firms. An inspection of the firms with large β during the ITB in figure 1 shows that they dominantly belong to the IT/technology sector. Likewise firms with large β during the FB are mostly from the financial sector.

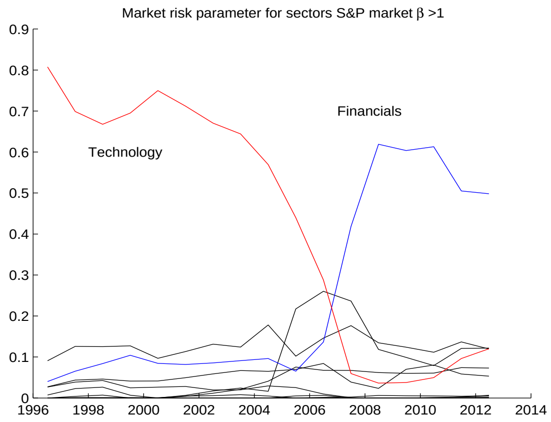

FIG. 4. Time dependence of the risk parameter $R(t, s)$ for the ten sectors of the S&P market.

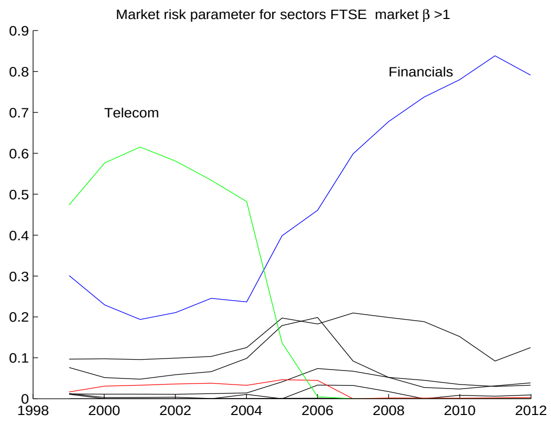

FIG. 5. Time dependence of the risk parameter $R(t, s)$ for the ten sectors of the FTSE market.

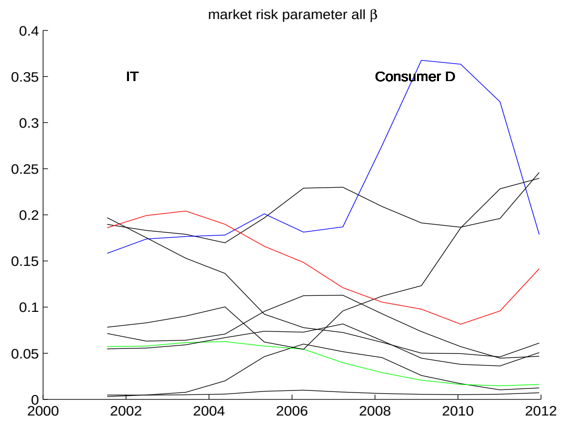

FIG. 6. Time dependence of the risk parameter $R(t, s)$ for the ten sectors of the DAX market.

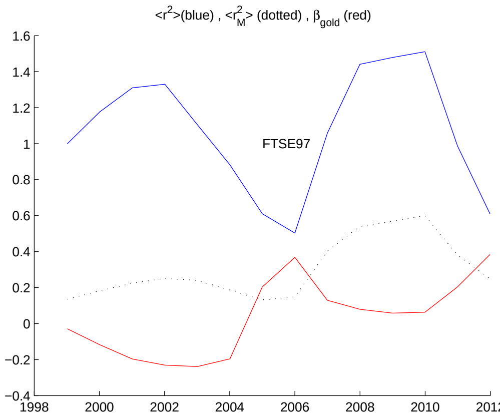

FIG. 7. Stock returns and the gold price: Time dependence of r_A^2 (blue line) for the FTSE market and the correlation β_g (red line).

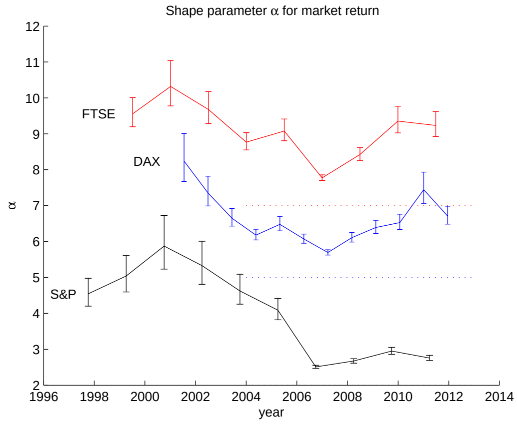

FIG. 8. Time dependence of the shape parameter α. DAX and FTSE are plotted with an offset.

Since a $\beta > 1$ signals a risky investment, we can define a market risk measure $R(t, s)$ for the sectors by multiplying $\beta > 1$ with the number $V(t, i)$ of traded shares in each window

$$R(t, s) = A_S \sum_{i \epsilon s} \theta(\beta_i - 1.0)\beta_i(t)\, V(t, i) \tag{9}$$

The normalization constant A_S is chosen to have $\sum_s R(t, s) = 1$. In figure 4 the risk parameters from eq. (9) is shown as function of time. Only the technology sector (red) before the transition in 2006 and the financial sector (blue) after 2006 exhibit large values of the risk measure. The value of the risk measure is small for all other sectors. Due to the time window of 3 years the time of the transition can be fixed only with an error of 1.5 years. A similar phenomenon is seen for the FTSE market in figure 5 with the difference that during the ITB the value for the telecommunication sector R is large, and the the behavior of the financial sector during the FB is more pronounced.

This may be due to the fact that only 3 small IT firms and a larger number of financial firms are part of our sample from the FTSE. The transition for the S&P appears to be somewhat sharper than for FTSE due to the smaller number N of stocks.

POSSIBLE REASONS FOR THE TRANSITION

The transition we observed poses the following question: Is it due to an external event or can it be explained within a statistical model? Of course the concurrence of one external event does not exclude the latter. There may also be weaker transitions that we cannot resolve within our poor time resolution and limited statistics. However, when one reviews the possible external events around 2006, there seem to be no obvious drastic changes in asset prices nor any political events.

One explanation for the transitions that we observe could be the behavior of the investors. One approach to examine this is to compare the stock market with the market for gold.

In figure 7 we show the average squared return r_A^2 as in eq. (6). It exhibits large values during the boom times ITB and FB. We compare this with the correlation β_g of gold price return [32] with the market obtained by inserting r_g into eq. (5). β_g has minima with negative or small values where r_A^2 is maximal and a maximum at the transition time 2006.

The figure suggests that the behavior of these two markets is at least weakly related, but that this relationship depends on the state of the stock market. At times when the stock market is risky the return for gold is decoupled from the stock market, while at calm times, the stock and the gold market are weak substitutes for investors.

A different approach would be to relate the changes to the way investors trade and to a technological change. A plausible explanation could be the increased amount of trading and the increase of high frequency trading (HFT), which became important in the years after 2005, see, e.g., [9, 11, 12].

If HFT corresponds to an external reason for the transition we should see differences in the market behavior before and after 2006, which can be related to HFT. The β coefficients are relatively constant before and after 2006. Therefore we concentrate on the market return or the index. Since both are very similar we prefer the former having less fluctuations. In the left panel of figure 9 we show (as a representative example) the response time and the traded volume at the Eurex exchange. The small response time of fractions of msec at the maximum of the volume indicates a growing dominance of computerized

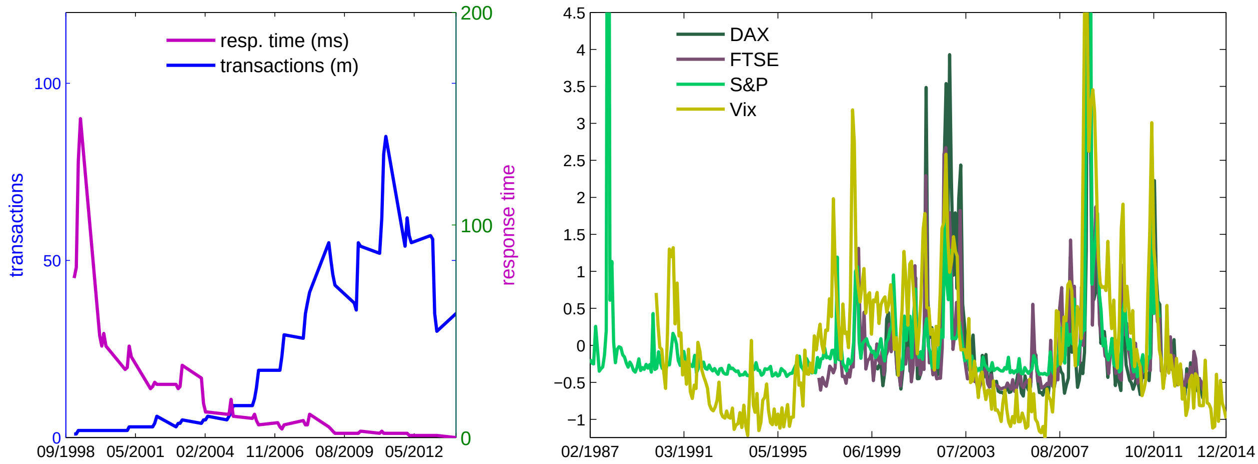

FIG. 9. (a) Time dependence of response time and traded volume at the Eurex exchange, source: Eurex Exchange [12]. (b) Normalized monthly volatility of the S&P, FTSE and DAX indices, normalized VIX index.

high-frequency trading after 2005. The peak in the trading volume corresponds roughly to the second peak in the risk measure for the S&P and FTSE market.

It is however difficult to attribute HFT to the IT bubble that shows as the first peak in our risk measure. In the right panel of figure 9 we show some estimates of the general market volatility. We calculate the monthly volatility (variance of the returns) for the three market indices and also plot the VIX volatility index [33].

If the kind in which trading is done has changed, this could also leave traces in the distribution of returns. The market return can be characterized by a shape parameter α using the Pareto-Feller parametrization from eq. (8).

We obtain α by maximizing the Log-Likelihood L in each window. Errors on α correspond to a change of L by 0.5. In figure 8 we show the time dependence of α for the three markets. Before 2006 one finds values $\alpha \sim 4-5$ with good χ^2 probabilities. For all three markets a drop to values below 3 appears after 2006. This implies divergent kurtosis or skewness. The χ^2 probabilities are worse, but still acceptable on the 5% level. However, the lower probabilities are due to systematic deviations from (8).

In figures 10 to 12 we show some typical pdfs of market return before and after 2006 for our markets. For the S&P market we see a perfect description by eq. (8) at 2000.8 and 2002.3, wheras at 2008.2 and 2009.7 a substantial excess at $r \sim 0$ occurs and the badly described tail extends to much larger value as before. This behavior is expected from HFT. Advocates of HFT [12, 30] claim that HFT leads to a more efficient market with less price changes. Also on a daily scale we expect smaller returns. Critics [5, 19] assert computerized trading increases instabilities. As a consequence the number of large daily returns is increased. Obviously both might be correct.

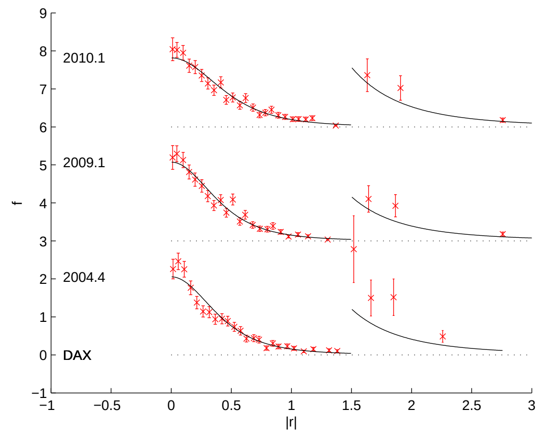

FIG. 10. Pdf of the DAX.

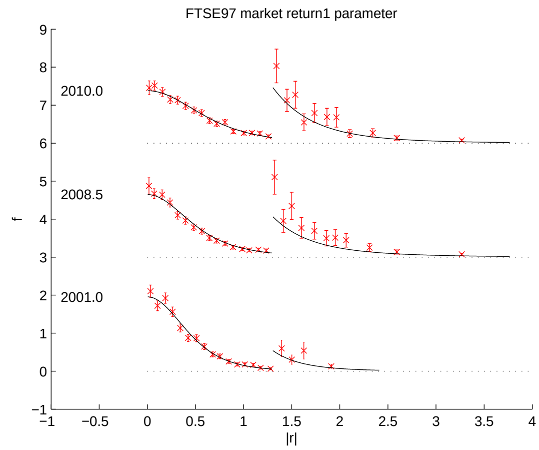

FIG. 11. Pdf of the FTSE.

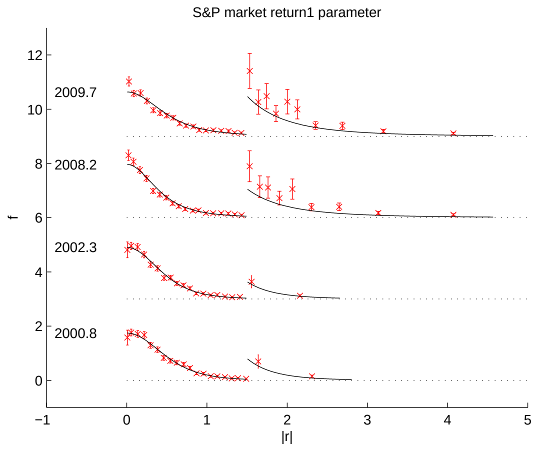

FIG. 12. Pdf of the S&P.

The different behavior of the return cannot be seen from the average r_A^2 in figure 7. In the boom times ITB and FB r_A^2 is large and of equal size. However, during ITB it is caused by medium returns and during FB by a decrease of medium and an increase of large returns. For the European markets the increase of small returns is absent. Only the enlarged tail is observed after 2006. These markets might be less affected by HFT and therefore only the instability effect is seen.

CONCLUSIONS

The literature on regularities in asset returns has for a long time described that the beta values of stocks are time varying. We have shown that we can extend the concept of beta values to systematically describe a risk measure for stocks from different sectors of the economy. These slowly varying sector specific risk measure describes ordered states in the market and identifies sectors which show concentration of market risk.

---

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

[2] Alfarano S, Wagner F, Milakovic M (2008) A nonparametric approach to the noise density in stochastic volatility models. Applied Financial Economics Letters 4(5).

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

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

[5] Bershova N, Rakhlin D (2013) High-frequency trading and long-term investors: a view from the buy-side. Journal of Investment Strategies, 2(2), 25-69.

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

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

[9] Carrion A (2013) Very fast money: high-frequency trading on the NASDAQ. Journal of Financial Markets, 16(4), 680–711.

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

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

[12] Ebholzer W (2014) Some insights into the details that matter for high-frequency trading! Eurex Exchange.

[13] Fama EF, French KR (1992) The Cross-Section of Expected Stock Returns. The Journal of Finance 47(2), 427–465. The Journal of Finance 46(4), 1485–1505.

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

[15] Kleiber C, Kotz S (2003) Statistical Size Distributions in Economics and Actuarial Sciences, John Wiley, Hoboken, NJ.

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

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

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

[19] Lewis M (2014) Flash Boys, WW Norton & Comp.

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

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

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

[23] Münnix MC, Shimada T, Schäfer R, Leyvraz F, Seligman T, Guhr T, Stanley HE (2012) Identifying States of a Financial Market. Scientific Reports 2(644).

[24] Plerou V et al. (2002) Random matrix approach to cross correlations in financial data, Phys.Rev. E 65, 066126.

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

[26] Raddant M, Wagner F (2013) Phase Transition in the S&P Stock Market, arXiv:1306.2508.

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

[28] Shephard N (2005) Stochastic Volatility Models: Selected Readings. Oxford University Press, Oxford.

[29] Tumminello M, Aste T, Di Matteo T, Mantegna RN (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.

[30] Overdahl J, Opening Statement to High-frequency trading (2012) High Frequency Trading. The Economist, March 2012.

[31] We excluded stocks which price behavior or market capitalization showed similarities to penny stocks, or which were exempt from trading or traded with negligible amounts for more than 10 days, or for which the trading volume was negligible for more than 8% of the total trading days.

[32] For the gold price we use the price quoted at the London Bullion Market, which is the reference gold price for most transactions worldwide.

[33] The VIX is a volatility index that refers to the Chicago Board Options Exchange Market. It describes the implied volatility from S&P500 index options and is thus a measure of expected price changes of this index.

Perturbation Theory

Assume a matrix C can be written as $C = C_0 + C_1$ with a small perturbation C_1. C_0 has one large eigenvalue E_0 and $N-1$ degenerate zero eigenvalues. Due to the degeneracy we can impose for the eigenvectors e_i^μ with $\mu > 0$ of C_0 the conditions

$$(e^\nu, C_1\, e^\mu) = 0 \quad \text{for} \quad \mu, \nu > 0, \nu \neq \mu \tag{10}$$

(a, b) denotes the scalar product. The eigenvalues λ_μ and eigenvectors f_i^μ of C can be expanded in a power serie in C_1/E_0. For $\mu = 0$ we get

$$\lambda_0 = E_0 + (e^0, C_1\, e^0) + \frac{1}{E_0}\left[(e^0, C_1^2\, e^0) - (e^0, C_1\, e^0)^2\right] \tag{11}$$

$$f_i^0 = \left[1 - \frac{1}{E_0}(e^0, C_1\, e^0)\right] e_i^0 + \frac{1}{E_0}(C_1\, e^0)_i \tag{12}$$

The remaining eigenvectors need the solution of condition (10)

$$\lambda_\mu = (e^\mu, C_1\, e^\mu) - \frac{1}{E_0}(e^0, C_1\, e^\mu)^2 \tag{13}$$

$$f_i^\mu = e_i^\mu - \frac{1}{E_0}(e^0, C_1\, e^\mu) e_i^\mu \tag{14}$$

Note this expansion reproduces the exact result for $tr\ C$ and for C_1 proportional to a unit matrix.
With $(C_0)_{ij} = \gamma_i \gamma_j$ we have $E_0 = (\gamma, \gamma) = \gamma^2$ and $e_i^0 = \gamma_i/\sqrt{\gamma^2}$. Inserting $(\gamma, C_1^k \gamma) = A_k\, \gamma^2$ into eqns (11) and (12) we get for λ_0 and $\beta_i = \sqrt{N} f_i^0$

$$\lambda_0 = \gamma^2 + A_1 + \frac{1}{\gamma^2}(A_2 - A_1^2) \tag{15}$$

$$\beta_i = \left[1 - \frac{1}{\gamma^2} A_1\right] \sqrt{\frac{N}{\gamma^2}} \gamma_i + \frac{1}{\gamma^2} a_i \tag{16}$$

with $a^2 = (N/\gamma^2) A_2$. Market dominance implies $E_0 = \gamma^2 \propto N$ and the constants A_k are of order 1. Eqns (15) and (16) show that up to corrections of order $1/N$ the leading eigenvalue λ_0 and its eigenvector β_i do not depend on C_1.