KIEL
WORKING
PAPER

Multivariate GARCH for a large number of stocks

Matthias Raddant, Friedrich Wagner

MULTIVARIATE GARCH FOR A LARGE NUMBER OF STOCKS

Matthias Raddant[1,2] and Friedrich Wagner[3]

ABSTRACT

The problems related to the application of multivariate GARCH models to a market with a large number of stocks are solved by restricting the form of the conditional covariance matrix. It contains one component describing the market and a second simple component to account for the remaining contribution to the volatility. This allows the analytical calculation of the inverse covariance matrix. We compare our model with the results of other GARCH models for the daily returns from the S&P500 market. The description of the covariance matrix turns out to be similar to the DCC model but has fewer free parameters and requires less computing time. The model also has the advantage that it contains the calculation of dynamic beta values. As applications we use the daily values of beta coefficients available from the market component to confirm a transition of the market in 2006. Further we discuss properties of the leverage effect.

Keywords: Multivarite GARCH models, CAPM, market risk

JEL classification: C58, C55, G12

Matthias Raddant
Kiel Institute for the World Economy
Kiellinie 66, D-24105 Kiel, Germany
phone: +49-431-8814-276
Fax: +49-431-85853
Email: matthias.raddant@ifw-kiel.de

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

[1]Institute for the World Economy, Kiellinie 66, 24105 Kiel, Germany
[2]Department of Economics, Kiel University, Olshausenstr. 40, 24118 Kiel
[3]Institute of Theoretical Physics, Kiel University, Leibnizstr. 19, 24098 Kiel

1 Introduction

The estimation of co-movement between stocks is an essential problem in the analysis of asset returns, financial integration, and for portfolio management. On the one side, the statistical properties of asset returns necessitate to treat them in a setting where time-varying volatility has to be modeled. This is mostly done within the setting of a GARCH model. Co-movement of asset returns on the other side is often involved with problems where the dimensionality of the problem is large – a feature that is at odds with many multivariate versions of GARCH models.

In the following we present a new multivariate version of a GARCH model that can easily deal with a large number of assets from one market, and we show that the estimation results do not differ much from those of existing models. Hence our model provides a solution for the treatment of asset markets with a large number of constituents.

For an index or a single stock the so-called GARCH(1,1) model (Engle, 1982) has turned out to be very successful despite of its simplicity. The return r_t at time t is written as a product of a factor $\sqrt{h_t}$ and an i.i.d. noise factor. h_t corresponds to the conditional expectation value $E_{t-1}[r_t^2]$. It obeys a linear recursion formula with three parameters related to the unconditional expectation value $\bar{h}$ of r_t^2, a time constant and a shape parameter describing the deviation from a Gaussian pdf for r_t. The model accounts for the stylized facts of returns (volatility clustering, fat tails) and allows to predict future volatility. Values of the parameters can be obtained by maximum likelihood estimates (MLE).

This univariate GARCH model has been generalized for N stocks. The multivariate version of the recursion for h_t becomes a matrix relation for the conditional covariance matrix H_t (see, e.g., Bauwens et al., 2006). These multivariate GARCH models face several difficulties in the case of large N. The first one is related to the increase of the number of free parameters with N. For example in the diagonal vector MGARCH model (Bollerslev et al., 1988) there are $N(N+1)$ GARCH parameters to be determined by MLE. Their number may be reduced to $2N$ assuming factorization as in the BEKK model (Engle and Kroner, 1995) or to two as in the scalar MGARCH (Engle, 2009). However, in addition $N(N+1)/2$ parameters appear in the expectation value $\bar{H}$ of H_t. To avoid these 'nuisance parameters' (Engle et al., 2008) covariance targeting has been proposed (Engle and Mezrich, 1996). In this case the matrix $\bar{H}$ is set to the time expectation value H_T, which creates another problem. When one analyzes H_T for a time period of around $T = 4$ years or longer, one finds that it can be described by one large eigenvalue in the order of N and a bulk of small eigenvalues, which can be described qualitatively by the spectrum originating from a random matrix (Marchenko and Pastur, 1967). The difference between H_T and the true covariance matrix $\bar{H} = \lim_{T\to\infty} H_T$ is in the order of $\sqrt{N/T}$. This makes H_T not very reliable as estimator for $\bar{H}$ for large N. A further problem for multivariate GARCH models is the computing time T_{comp} for large N. In each step of MLE the inverse matrix H_t^{-1} has to be calculated. This implies an increase of T_{comp} with N^3.

Several methods have been proposed to solve these problems. In the

CCC (Bollerslev, 1986) or the OGARCH model (Alexander, 2001) a restricted time dependence of H_t is used. CCC uses a constant correlation matrix and OGARCH time-independent eigenvectors of H_t. In both cases the problem reduces to N univariate GARCH estimates of the diagonal elements of H_t (CCC) or the time-dependent eigenvalues (OGARCH). The necessary matrix operation in the latter can be calculated outside MLE. Constant correlations in CCC or a constant leading eigenvector in OGARCH disagree with the observed change in the latter (Raddant and Wagner, 2017). Time-dependent correlations can be obtained in the DCC model (Engle, 2002) using a scalar MGARCH model. The problem with H_t^{-1} allows only very modest N. Several extensions especially of the DCC model have been proposed, for example by Franses and Hafner (2009). In the modification proposed by Engle et al. (2008) only correlations, resp. covariances between pairs of stocks are estimated. Using all pairs leads to $T_{comp} \propto N^2$. By selecting only N pairs one achieves $T_{comp} \propto N$. Aielli (2013) however shows that certain inconsistencies can also arise in this case and proposes a different modification.

In our solution of the problems we investigate a model that requires computing time for powers of H_t^α in the order of N and involves only well defined parts of H_T. This is achieved by using a restricted form of H_t with one large time dependent eigenvalue $N v_0(t)$ and $N-1$ degenerate time dependent smaller eigenvalues $v_1(t)$. The parameter problem is solved by applying the diagonal vector MGARCH model not to H_t, but rather to H_t projected on its eigenvector space.

A benefit of this approach is the possibility to interpret the eigenvector belonging to the large eigenvalue as beta coefficients in the CAPM sense relative to the market. Many approaches to describe beta coefficients in a way that is consistent with varying covariances have been discussed in the literature, starting with Bollerslev et al. (1988) and more recently Engle (2014). Hence, here we show that it is also possible to combine the determination of beta values with a restricted multivariate GARCH model.

The paper is organized in the following way. The derivation of the model is given in section 2. We obtain two recursions for the eigenvalues. In addition there is a GARCH type recursion for the leading eigenvector $\beta(t)$. MLE fits using our model are applied to the daily returns of 356 stocks from the S&P market in the years 1995-2013 in section 3. In section 4 we compare our model with other GARCH models and data. In section 5 we utilize the fact that our model generates daily beta values and H_t for two applications: First we analyze how the H_ts have developed over time and we show transitions in the market. As a second application we investigate the leverage effect, especially we look for any correlation with $\beta(t)$. The last section contains some conclusions.

2 Multivariate GARCH with restricted covariance matrix

In the univariate GARCH(1,1) model the returns r_t $(t = 1 \ldots T)$ of a single stock or an index are written as

$$r_t = \sqrt{h_t}\, \eta_t \tag{1}$$

with η_t a white noise with mean zero and variance one. h_t is the conditional expectation value of r_t^2 or volatility factor. It obeys the recursion

$$h_{t+1} = \omega + \alpha r_t^2 + b\, h_t. \tag{2}$$

We prefer a slightly different parametrization making the mean reverting property of equation (2) more explicit by setting

$$b = 1 - \alpha - \gamma \quad \text{and} \quad \omega = \gamma\, \bar{h} \tag{3}$$

with $\bar{h}$ the expectation value of r_t^2 or h_t. Equation (2) is changed to

$$h_{t+1} = h_t\ + \alpha(r_t^2 - h_t)\ + \gamma(\bar{h} - h_t). \tag{4}$$

For N stocks of a financial market we have returns r_{ti} with $i = 1 \ldots N$ normalized to $\sum_{t,i} r_{ti}^2 = TN$. The relation (1) between noise and return is generalized in matrix notation[1] to

$$r_t = H_t^{1/2} \cdot \eta_t \tag{5}$$

The matrix H_t corresponds to the conditional expectation value of the covariance matrix of r_{ti}. As generalization of the recursion (4) the vector MGARCH model (Bollerslev et al., 1988) has been proposed. In its diagonal form (Engle, 2009) the recursion for H_t is written as

$$(H_{t+1})_{ij} = (H_t)_{ij} + \gamma_{ij}(\bar{H} - H_t)_{ij} + \alpha_{ij}(r_t \cdot r_t' - H_t)_{ij} \tag{6}$$

Three problems restrict applications of (6) to only very modest N. The first is the necessity to calculate $H_t^{-1/2}$ or H_t^{-1} for the likelihood which can be very time consuming for large N. The second problem is related to the expectation value $\bar{H}$. Considered as parameter this amounts to $N(N+1)/2$ parameters leading again to long computing times. The alternative of determining $\bar{H}$ from the time average of $r_t \cdot r_t'$ suffers from the uncertainty of order $\sqrt{N/T}$ as discussed in section 1. The third problem consists in the number $N(N+1)$ of GARCH parameters α and γ. Even when we assume factorization of α and γ as in the BEKK model (Engle and Kroner, 1995), the number of parameters is too large for an application for example to the S&P500 market.

In our approach we solve the first two problems by restricting the form of H_t. The last problem is solved by applying the recursion of the diagonal vector

---

[1] We use a' for the transpose of vector a. $a' \cdot b$ denotes the scalar and $a \cdot b'$ the tensor product. $M \cdot b$ denotes a matrix multiplication. r_t describes a vector (of stock returns) unless it is indexed otherwise.

MGARCH model not to H_t itself, but rather to H_t projected on the eigenvector space of H_t.

For the restriction we start from the spectral decomposition of a general H_t

$$H_t = W' \cdot \Lambda(t) \cdot W(t) \tag{7}$$

with $\Lambda(t)$ the diagonal matrix of the eigenvalues $\lambda_\mu(t)$ $\mu = 1, \ldots, N$ and $W_{\mu i}$ the matrix of eigenvectors. Motivated by the observation that $r_t \cdot r_t'$ averaged over few years (Raddant and Wagner, 2017) has within our normalization one large eigenvalue λ_0 of order N and $N - 1$ eigenvalues of order 1, we keep the large eigenvalue in H_t and approximate the remaining by a degenerate spectrum $\lambda_1(t)$. Using the normalization $\lambda_0(t) = Nv_0(t)$ and $\lambda_1(t) = Nv_1(t)/(N-1)$ we obtain the following decomposition for H_t

$$H_t = N \left(v_0(t) \, P_0(t) \; + \; \frac{v_1(t)}{N-1} \, P_1(t) \right) \tag{8}$$

with the projection matrices $(P_0)_{ij} = W_{0i}W_{0j}$ and $P_1(t) = 1 - P_0(t)$. v_0 describes the market volatility and v_1 the non-market volatility. The projectors P_ν are orthogonal and idempotent. This allows analytical computation of functions $f(H_t)$ by

$$f(H_t) = f\big(\lambda_0(t)\big) \, P_0(t) \; + \; f\big(\lambda_1(t)\big) \, P_1(t) \tag{9}$$

thereby solving the first problem. Using P_ν is preferable, since they are uniquely determined, whereas W are not, especially in the case of degeneracy.

Another advantage is a possible economical interpretation of P_0. The component of the large eigenvalue can be interpreted as a market component (Laloux et al., 1999). One can define a market return by projecting r_t on the λ_0 component

$$r_{Mt} = \frac{1}{\sqrt{N}} \sum_i W_{0i} r_{ti}. \tag{10}$$

The beta coefficients in a CAPM approach (Sharpe, 1964; Lintner, 1965) relative to r_{Mt} are given by

$$\beta_i(t) = \frac{E_{t-1}(r_{ti}r_{Mt})}{E_{t-1}(r_{Mt}^2)} \tag{11}$$

With equation (10) the conditional expectations can be evaluated and leads to time dependent $\beta(t)$

$$\beta_i(t) = \sqrt{N} \, W_{0i} \tag{12}$$

with the normalization $\sum_i \beta_i^2 = N$.

This is similar to the conditional betas that Engle (2014) derives for the GARCH model

$$\beta_i(t) = N \sum_j (H_t^{-1})_{ij} \, E_{t-1}(r_{tj}r_{Mt}) \tag{13}$$

We apply an additional factor N due to the normalization of r_{Mt}. Inserting equation (10) into equation (13) reproduces our result from (12). In term of β the restricted form of H reads

$$H_t = \left(v_0(t) - \frac{1}{N-1}v_1(t) \right) \beta(t) \cdot \beta'(t) \; + \; \frac{N}{N-1}v_1(t) \, 1 \tag{14}$$

5

This expression is similar to the latent factor model (Diebold and Nerlove, 1989) except that the term proportional to the unit matrix is replaced by a diagonal matrix in the latter.

The restricted form (8) solves also the second problem. $\bar{H}$ involves only the leading eigenvalue $N\bar{v}_0$, its time averaged eigenvector $\bar{\beta}$ and its trace $N(\bar{v}_0+\bar{v}_1)$. As shown in Raddant and Wagner (2017) these can be estimated by covariance targeting with an error of order $1/\sqrt{T}$.

To reduce the number of GARCH parameters in the recursion we apply the vector MGARCH model to H projected on eigenvector space of H_t. Using $(W(t)H_tW(t)' = \Lambda$ we write[2]

$$\left(W(t)H_{t+1}W(t)'\right)_{\mu\mu'} = \Lambda_{\mu\mu'} + a_{\mu\mu'}\left(W(t)\cdot(r_t\cdot r_t')\cdot W(t)' - \Lambda\right)_{\mu\mu'}$$
$$+ g_{\mu\mu'}\left(W(t)\bar{H}W(t)' - \Lambda\right)_{\mu\mu'} \qquad (15)$$

In this equation only the transformed matrix of H_t is a diagonal matrix, whereas those of $H_{t+1}, r_t\cdot r_t'$ and $\bar{H}$ are not. Without any restriction on H_t we have the same number of parameters as in (6). However, the parameter a and g have to be compatible with a restricted H_t. For example equation (15) reproduces the OGARCH model for time independent W. Then all matrices in equation (15) are diagonal and the off diagonal elements of a and g have to vanish. As a consequence equation (15) describes N univariate GARCH(1,1) models depending on the projected data $(W \cdot r_t)^2$. For our ansatz (8) a and g have to respect the degeneracy of H_t. This leads to the conditions

$$a_{00} = \alpha_{00} \quad a_{0\mu} = \alpha_{01} \quad a_{\mu\mu'} = \alpha_{11} \quad \text{for } \mu, \mu' \neq 0 \qquad (16)$$

and an analogue relation between g and γ. With $W \cdot W' = 1$ we can obtain from equation (15) the recursion for H. With the help of the identity

$$\sum_{\mu,\mu'=0,N-1} a_{\mu\mu'}W_{\mu i}W_{\mu k}W_{\mu'l}W_{\mu'j} = \sum_{\nu,\nu'=0,1}\alpha_{\nu\nu'}(P_\nu)_{ik}(P_{\nu'})_{lj} \qquad (17)$$

we can express our final recursion in terms of the projectors

$$H_{t+1} = H_t + \sum_{\nu,\nu'}P_\nu(t)\left(\alpha_{\nu\nu'}(r_t\cdot r_t' - H_t) + \gamma_{\nu\nu'}(\bar{H} - H_t)\right)P_{\nu'}(t) \qquad (18)$$

The restricted form of H_t reduces the number parameters in a maximum likelihood estimate from $N(N+1)$ to manageable number of six. With $\alpha_{\nu\nu'} = \alpha$ and $\gamma_{\nu\nu'} = \gamma$ we recover the scalar MGARCH or BEKK model and for $\alpha_{01} = \gamma_{01} = 0$ a two component OGARCH model with strictly positive H_t.

In the appendix we derive from (18) two recursions for the market volatility $v_0(t)$ and the non-market volatility $v_1(t)$, as well a recursion describing the time dependence of $\beta(t)$. Since the result given in equations (31) and (33) is rather complicated we illustrate them by quoting only the limit of large N and small deviations of $\beta(t)$ from its equilibrium value $\bar{\beta}$:

$$v_\nu(t+1) = v_\nu(t) + \gamma_{\nu\nu}\left(\bar{v}_\nu - v_\nu(t)\right) + \alpha_{\nu\nu}\left(\rho_\nu^2(t) - v_\nu(t)\right) \qquad (19)$$

---

[2]In this notation we use μ ranging from $0\ldots N-1$ to index all eigenmodes. $\nu = 0$, 1 indexes the two modes in the restricted form.

Equation (19) corresponds to two univariate GARCH recursions depending on the combinations

$$\rho_0^2(t) = r_{Mt}^2 \quad \text{and} \quad \rho_1^2(t) = \frac{1}{N}\sum_i r_{ti}^2 \ - \ r_{Mt}^2 \tag{20}$$

of the observed returns. For β we obtain a recursion for each component β_i:

$$\beta_i(t+1) = \beta_i(t) \ + \ \frac{\alpha_{10}\, r_{Mt}}{v_0(t+1)}\big(r_{ti} - r_{Mt}\beta_i(t)\big) \ + \ \frac{\gamma_{10}\, \bar{v}_0}{v_0(t+1)}\big(\bar{\beta}_i - \beta_i(t)\big) \tag{21}$$

$\beta(t)$ changes only for non zero α_{01}, γ_{01}. Its changes depend on the difference between stock returns r_{ti} and market return r_{Mt}, and the deviation from the equilibrium values $\bar{\beta}_i$.

3   Estimation of the parameters with S&P data

In this section we describe the maximum likelihood estimation with data of daily returns of 356 stocks from the S&P market in the years 1995-2013. 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.[3]

We calculate the log likelihood L of our model as described in equations (31) and (33) of the appendix. Initial values for ν and β have been determined from a time average of the covariance matrix of the first 4 years. In a first fit we use only two parameters with $\alpha_{\nu\nu'} = \alpha_{00}$ and $\gamma_{\nu\nu'} = \gamma_{00}$. The noise is assumed to be Gaussian.

The resulting log likelihood (divided by T) and the parameter values are given in table 1. When we invert equation (5) we can obtain the so-called de-garched returns $\eta_G(t)$.

$$\eta_G(t) = H_t^{-1/2} \cdot r_t \tag{22}$$

If the GARCH model represents the data exactly, $\eta_G(t)$ should be again Gaussian distributed. This is not the case, as the pdf for all $\eta_G(t)_i$ given in figure 1 shows. The pdf for $\eta_G(t)$ is well described by a Student's t-distribution with a tail index of $\nu = 3.32$.

Also it has been observed from a non-parametric moment analysis that stocks require a noise different from that of indices (Wagner et al., 2010). This motivates to repeat the estimate with t-distributed noise. The estimated value of $\nu = 3.25$ agrees with the value obtained from figure 1 (lies within the errors). The resulting L (2nd line in table 1) corresponds to an astronomical increase of probability. Even the pessimistic evaluation using the probability change per t given by $exp(-\Delta L/T)$ is highly significant. This sensitivity to the noise is not observed in application of univariate GARCH models to indices. In the case of

---

[3]We excluded stocks which price did not change for more then 8 % of the trading days, or which were exempt from trading or for which no trading was recorded for more than 10 days in a row. We manually deleted 15 stocks which price movements at some point showed similarities to penny stocks and/or which market capitalization was very low.

N_{par}	L/T	$\alpha_{00}\cdot 10^2$	$\gamma_{00}\cdot 10^2$	$\alpha_{11}\cdot 10^1$	$\gamma_{11}\cdot 10^2$	$\alpha_{10}\cdot 10^2$	$\gamma_{10}\cdot 10^2$
2	-52.2	4.871	0.383	α_{00}	γ_{00}	α_{00}	γ_{00}
		(0.068)	(0.014)				
2	-2.57	3.132	0.284	α_{00}	γ_{00}	α_{00}	γ_{00}
		(0.048)	(0.011)				
4	-0.10	1.592	0.328	2.472	0.766	α_{00}	γ_{00}
		(0.042)	(0.016)	(0.050)	(0.078)		
6	0.00	5.14	4.13	2.487	0.781	1.673	0.298
		(0.16)	(0.31)	(0.041)	(0.065)	(0.044)	(0.014)

Table 1: *Maximum likelihood estimates of the parameters with the S&P market. Column one gives the number of parameters, column two the values of the log likelihood per time relative to the six parameter fit. Errors given in the row below the values correspond to 5% confidence levels. They are taken from the diagonal elements of the inverse Hessian. The results in the two top rows are obtained with Gaussian noise, in rows 3-8 with t-distributed noise.*

the S&P index a t-distributed noise leads to an improvement of $\Delta L/T = 0.02$. The large value of $\nu = 8$ is due to outliers in the return and does not change the GARCH parameters.

Another reason for the need t-distributed noise may be the restricted form of our matrix GARCH model (hereafter called RMG). The parameters α and γ do not depend on individual stocks. They cannot accommodate to the observed very different behaviour of stock returns. To demonstrate the improvement one can compare the Monte Carlo simulated pdf of stocks with the observed pdf. Simulating RMG is not very meaningful since stocks enter only via the mean values $\bar{\beta}$. Instead we use the possibility of GARCH to predict returns from equation (5) for given H_t with repeated noise factors η. These predicted pdf should agree with the observed pdf.

In figure 1 two typical pdf's of returns are compared with the predicted density either using t-distributed (black line) or Gaussian (red line) noise. The latter describes the tails but fails in the transition to the tail, which may be the reason for the difference in probability. In all subsequent estimates we use the value of $\nu = 3.25$. Still the time dependence described by γ_0 is not satisfactory.

The value $\gamma_{00}^{-1} = 339\ days$ exceeds the value needed from the autocorrelation of $|r_{Mt}|$ by an order of magnitude. A four parameter fit with $\alpha_{10} = \alpha_{00}$ and $\gamma_{10} = \gamma_{00}$ improves the likelihood (third line in table 1), but does not solve the problem. Only when all six parameters are varied (fourth line), the time constant for β is much smaller than for $v_{0,1}$ together with an again improved probability. In each case the improvement $\Delta L/T > 0.02$ is much larger than the trivial value inferred from the fit to the index.

Our RMG rests on the simple form of equation (8). This leads for the observed correlation matrix to an eigenvalue spectrum which is the sum of one large eigenvalue of order N and a random matrix component described by a Marchenko and Pastur (1967) spectrum. This describes qualitatively the observations. The RMG is distinguished by a small number of parameters
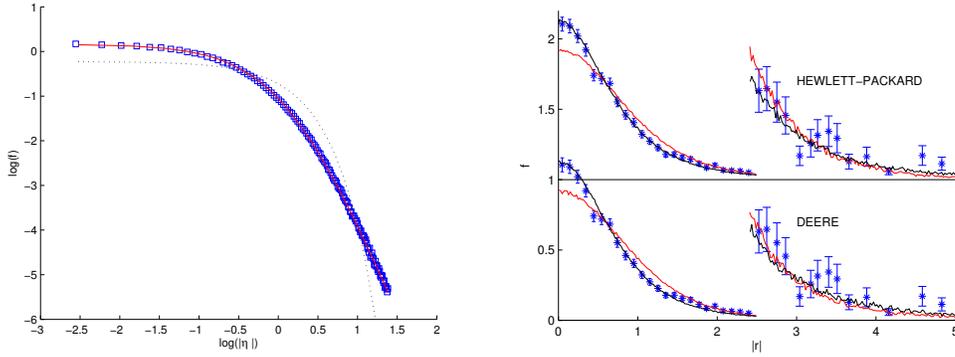
Figure 1: *The left panel shows the pdf of all GARCH filtered returns $\eta_G(t)_i$ obtained from the raw returns by equation (22). The red line corresponds to a fitted t-distribution with $\nu = 3.32$, the dotted line shows a Gaussian distribution. In the right panel two typical pdfs for stocks of the S&P market are compared with the predicted pdf using t-distributed (black) or Gaussian (red) noise. The values (log of the pdf) $|r| > 2.5$ are multiplied by a factor of 10.*
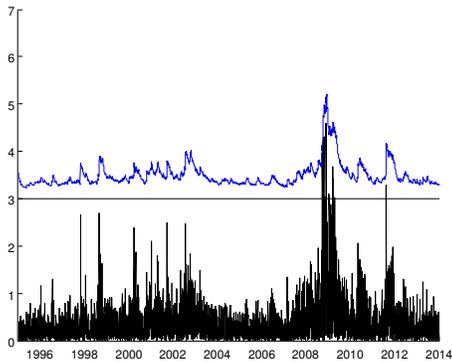

Figure 2: *Market volatility $\sqrt{v_0(t)}$ (top) and market return $|r_{Mt}|$ (bottom). A value of 2 is added to $\sqrt{v_0(t)}$*
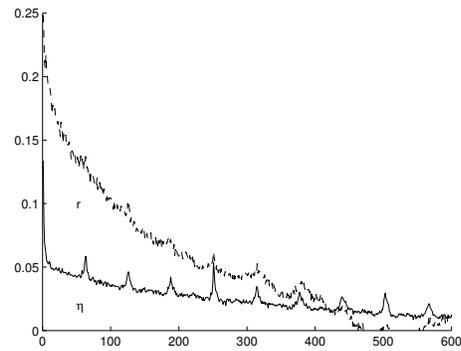
Figure 3: *Autocorrelation for returns $|r_{ti}|$ averaged over all stocks (top) and the averaged GARCH filtered returns $|\eta_G(ti)|$*
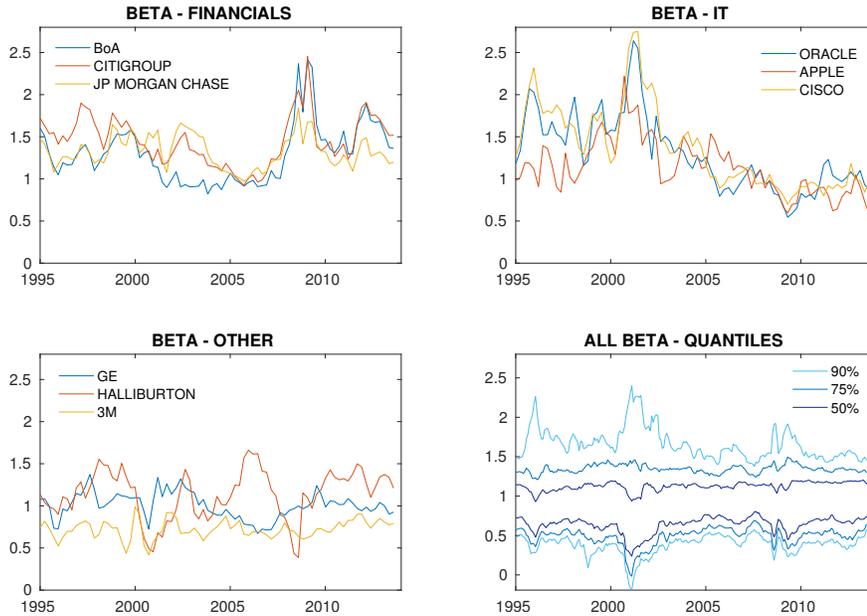
9

Figure 4: *Beta values from the RMG model. The top and the bottom left panel show beta values for different stocks, averaged over a 60-day window. The bottom right panel shows the dynamics of the distribution of beta values by the quantiles, using a 20-day window.*

and analytical calculation of H_t^{-1}, which makes it applicable also for a large number of stocks. It allows separation of a market component with economical meaningful β coefficients and r_t on the same time scale as $v(t)$.

As in any GARCH model one finds volatilities much less noisy than the absolute returns. This is shown in figure 2 where the market volatility $\sqrt{v_0(t)}$ is compared with the underlying market return r_{Mt}. Another interesting effect is shown in figure 3, where we compare the autocorrelation for the GARCH filtered returns $|\eta_G(ti)|$ with $|r_{ti}|$ averaged over all stocks i. The former has much smaller correlations on a time scale of years. The large statistics allow even to resolve the peaks at multiple of three months due to the dividend pay days. Some autocorrelation in the filtered returns remains visible. This is mainly due to the restrictions on our covariance matrix.[4]

Figure 4 finally gives an overview about the beta values that can be obtained within the RMG. Note that these betas are normalized as stated in equation 12. The top left panel shows the development of beta values for some financial stocks, the bottom right shows the beta of IT related companies. Stocks from both groups show strong similarity within their group. The bottom left panel shows the beta for stocks from other sectors. They develop much more diverse. The bottom right panel illustrates how the distribution of beta values has developed over time. In time of crisis we observe peaks for the beta values but

---

[4]It is also possible to calculate de-garched returns using only $diag(H_t)$. The resulting acf is practically identical to that of η.
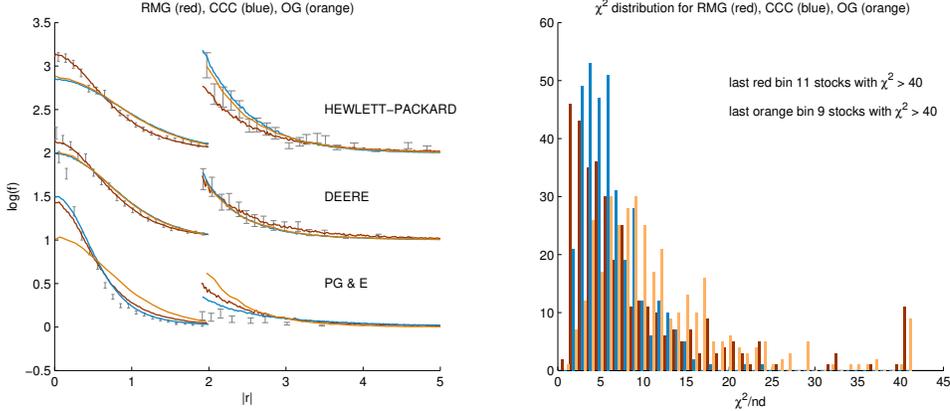
Figure 5: *The left panel shows the pdf of returns compared with the prediction of RMG (black), UVG (blue) and OG (red). Left part gives the histogram of χ^2 for all stocks from RMG (red) , OG and UVG(white)*

also a much wider distribution in general.

4   Comparison with other GARCH models

In the following we will compare the estimation results of our model with those of other multivariate GARCH models. We start by comparing the predicted pdf of stock returns with the six parameter model discussed in section 3 with the CCC and OGARCH model (hereafter abbreviated with OG). The CCC model uses a constant correlation matrix supplemented with N univariate GARCH for r_{ti}. In OG we include all eigenmodes $P_\nu \cdot r_t$, where the constant projectors P_ν are obtained from $\bar{H}$ by covariance targeting. In both we use i.i.d. Gaussian η for the noise. The left part of figure 5 shows that the pdf for the same two stocks as in figure 1 is in reasonable agreement with the three models. RMG and OG fail in cases where the leading eigenvector does not dominate H_t, as the third example for PG&E shows.

For a more comprehensive comparison we use the χ^2/n_d ratio equal to the sum of quadratic deviation in units of the squared error divided by the number n_d of bins. A histogram of these ratios from the 356 stocks of the S&P market is shown for RMG, OG and CCC in the right part of figure 5.

All three distributions exhibit a peak around a value of 4-5 which corresponds to a 5% confidence level. Values between 10-20 lead still to a qualitative description. In contrast to CCC both RMG and OG have a small fraction (5%) of outliers mainly from the energy sector as PG&E. On average RMG performs better than OG which may be due to the systematic error by covariance targeting. We stress that in CCC and OG 712 parameter have to be determined. The only six parameter used in RMG lead to a much more parsimonious description of the data.

Another way of assessing the results is of course the comparison of the estimated H_ts with those obtained from the other models. In the following we

11

use the DCC model as a benchmark and compare the resulting time series of covariances. We use 10 day averages. For the DCC and the OG we estimate the model repeatedly with 10 different stocks, since an estimation with the entire data set is not possible.

Figure 6 shows a histogram of RMSDs of pairwise comparisons of time series of estimated covariances. The figure shows similarities to the previous results, but now the DCC model serves as a benchmark. The deviations between DCC and our model are smallest, slightly better than the CCC model, the difference between DCC and the OG model is much larger. Interestingly the constant conditional correlations seems to work better than assuming constant principal components, especially in the long run.

It is also interesting to analyze which stocks are responsible for most of the deviations. A sectoral break-down of the results, shown in figure 7 reveals some patterns. The RMG describes most of the stocks well, but has slight problems with stocks from the energy sector, which are probably not always well described by our model, since they are not very representative for the market trend. A similar observation can be made for CCC. In this case however part of the deviation also stems from the IT sector.

In the next section we will show that the average correlation in the IT sector has slightly declined - against the general trend, which might explain this finding. In fact the OG model seems to have problems with all stocks that come from sectors where we have seen changes of relative importance or volatility over time.

5 Applications

5.1 Market transition

The large number of stocks in our sample allows to search for group specific regularities. An obvious question is if the correlations that can be extracted from H_t differ for stocks from specific sectors, and more interestingly, how they develop over time. Our sample period covers a time period during which we have seen pronounced changes in the market, namely the IT bubble, the financial crisis, and the growing importance of energy markets.

For this reason we use the estimated H_ts and then derive correlation matrices C_t such that

$$C_t = D^{-1} H_t D^{-1} \tag{23}$$

where D is the square root of the matrix with the main diagonal elements of H_t. We can use these correlations to calculate the median correlation for pairs of stocks from specific sectors (GICS classification).

Figure 8 shows some of these median correlations. We observe very high average within sector correlation for stocks in the IT and financial sector. But also some correlations between stocks of different sectors are rather high, for example when the consumer or materials sectors are involved.

In general we observe two important changes. First, the overall level of correlations has shifted upwards from 2002 until 2007. The second observation is
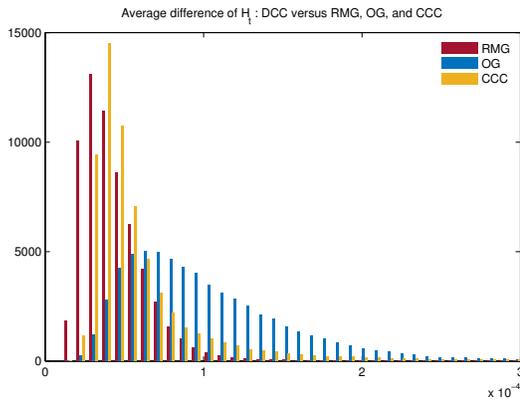
Figure 6: *Distributions of deviation of different H_t time series. We calculate the RMSD for all pairs of time series of estimated covariances and plot the histogram for the deviation of the time series derived from RMG, CCC, and OG versus the DCC model. We observe that RMG and CCC produce H_t that are relatively similar to DCC, while the difference to OG is large and more heterogeneous among the different covariances.*
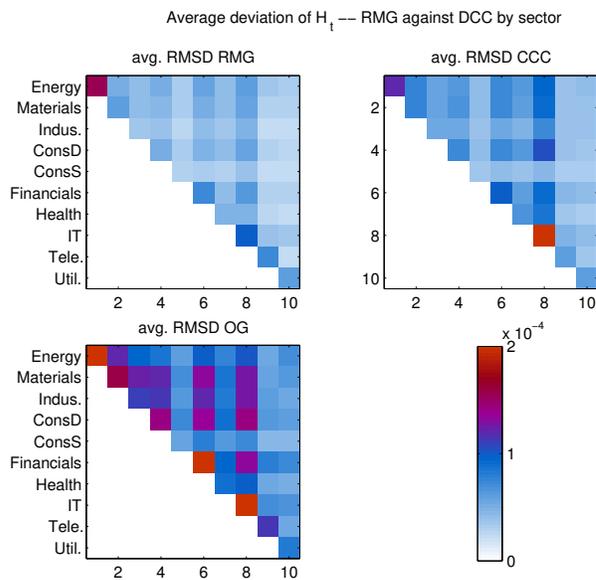

Figure 7: *Deviation from DCC results by sector. We calculate the RMSD for all pairs of H_t time series and calculate the average RMSD for covariances of stocks from specific sectors. We plot these averages as color-coded values for the deviation of the H_t time series for RMG, CCC, and OG versus the DCC model. We observe that differences in the estimated covariances stem from specific sectors, especially energy, IT, and the financial sector.*
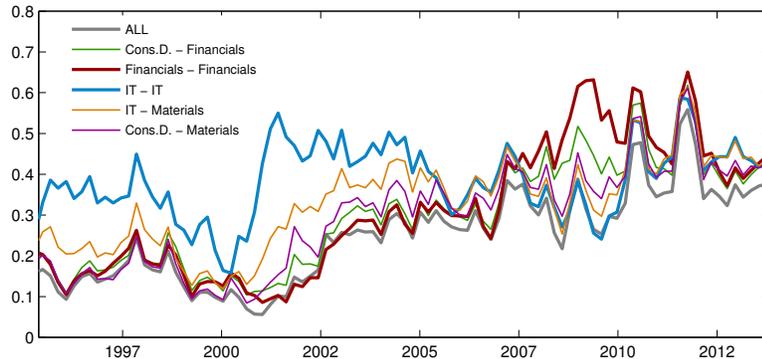
13

Figure 8: *Development of median correlation by sector, averaged over 50 days. We show the average of all stock correlations (bold grey line), as well as some of the sectors with the strongest correlations. The sector with the strongest inter-sector correlation has for a long time been the IT sector (blue), later the financial sector (red) has taken over this role.*

that the ranking between the sectors has changed, the financial sector surpasses the IT sector in terms of correlation around 2006.

These changes can be analyzed in more detail. In the analysis of Raddant and Wagner (2017) of the US, the UK, and the German stock market the same change has been found in the behavior based on the stock's beta values. In the years 1994-2006 trades of stocks with high beta and large volume were concentrated in the information technology sector, whereas in 2006-2012 those trades are dominated by stocks from the financial sector. The values of the β have been derived under the assumption that the covariance matrix of the returns have a large eigenvalue already at window sizes of 3 years. Since a $\beta > 1$ signals a risky investment, a market risk measure $R(t,s)$ has been defined for the sectors by multiplying $\beta_i > 1$ with the number $V(t,i)$ of traded shares in each window.

$$R(t,s) = A_S \sum_{i \epsilon s} \theta(\beta_i - 1.0)\beta_i(t) \, V(t,i) \tag{24}$$

The normalization constant A_S is chosen to have $\sum_s R(t,s) = 1$. When we apply this measure we see that before 2006 only the information technology sector and after 2006 only the financial sector exhibit large values of the risk measure.

However, the time and the duration of this transition had a systematic error of 1.5 years due to the window size. Repeating R with the β from RMG serves two purposes. Firstly it is a check whether in RMG the market property can be reproduced and secondly the transition time can be determined more accurately, since daily β are known from RMG. To reduce the noise on β we average $R(t,s)$ over one month. In figure 9 the risk parameters from equation (24) for the S&P market is shown as a function of time. The agreement with the previous determination (dashed lines) is good. With a time resolution of
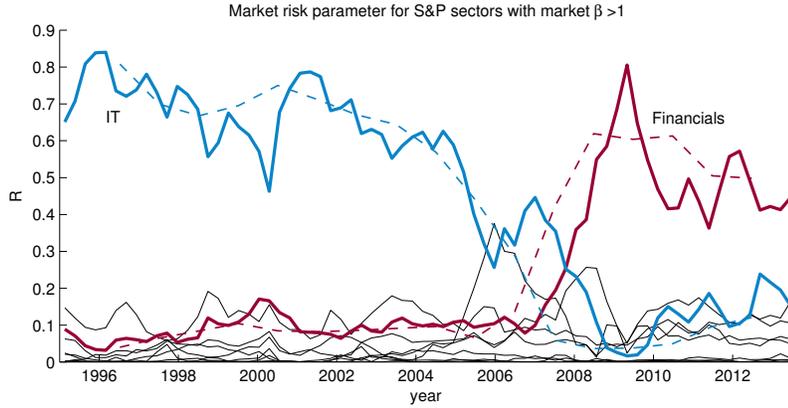
14

Figure 9: *Time dependence of the risk parameter described in equation (24) for the sectors of the S&P market. The dashed line shows the result of Raddant and Wagner (2017). The solid lines (blue for the information technology and red for the financial sector) use daily βs from the RMG.*

one month we can now safely say that the transition happens during the year 2006.

5.2 Leverage effect

The leverage effect consists in a negative correlation between volatility and future returns (Black, 1976). It is a relatively small effect (Schwert, 1989), but important for estimation of risk. Since GARCH models provide a measurement of the daily volatility they are well suited for an analysis of this effect.

In a first step we determine for each stock the time correlation $C_i(t)$ between the market volatility v_0 and the observed returns r_{ti}:

$$C_i(t) = \frac{1}{N_C} \sum_{t'} \left(v_0(t' - t) - \bar{v}_0 \right) r_{t'i} \tag{25}$$

with the normalization factor $N_C^2 = T \cdot var(v_0) \sum_t r_{ti}^2$. The stocks lead to a large variety of functions $C_i(t)$, although they exhibit the leverage effect in the sense of negative values of $C_i(t) \cdot sign(t)$. Therefore we use the asymmetry defined by

$$A_i = \frac{1}{t_m} \sum_{t=1}^{t_m} \left(C_i(t) - C_i(-t) \right) \tag{26}$$

with a maximum of the time lag t_m of two months. A_i corresponds to the difference in the area under $C(t)$ for positive and negative t.

It has been suggested by Black that the leverage effect is related to risk. To test this suggestion we show in figure 10 the asymmetry as a function of the mean value $\bar{\beta}_i$. Since $\beta(t)$ changes around 2006 we use two time series in the years from 1984 to 2005 and from 2007 to 2013. The asymmetries are clearly negative and increase slightly with $\bar{\beta}$ as indicated by the line connecting the
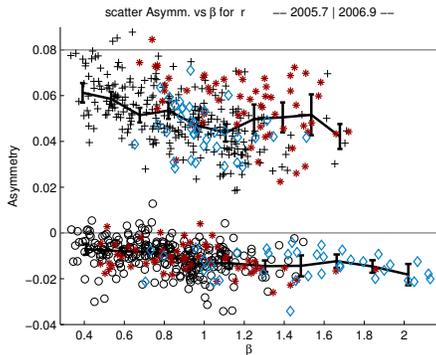
15

Figure 10: *Asymmetry of r vs $\bar{\beta}$ for stocks. Blue diamonds refer to technology, red stars to financials. The line corresponds to average $\bar{\beta}$ and stdev.*

Figure 11: *Asymmetry of η vs $\bar{\beta}$ for stocks. Blue diamonds refer to technology, red stars to financials. The line corresponds to average $\bar{\beta}$ and stdev.*

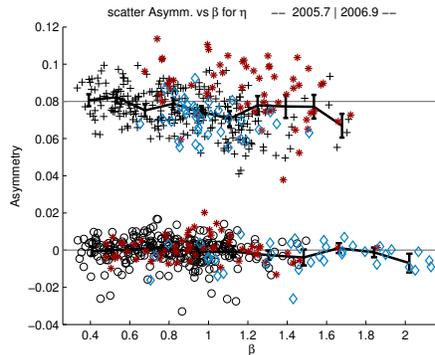mean of A. If our RMG describes the data successfully the leverage effect should disappear by replacing r_{ti} by the GARCH filtered returns $\eta_G(t)_i$ in equation (25). As shown in figure 11 this is in fact the case. The GARCH filtered returns have little autocorrelation and no significant leverage effect.

We conclude this section with a remark on predictions. These cannot be made using our RMG. For prediction of future returns the leverage effect has to be included in the recursion. Analogous to the GJR-Garch model (Glosten et al., 1993) an additional matrix proportional to $\delta_{ij} r_{ti} |r_{ti}|$ can be added on the RHS of equation (18). For large N the recursions for v_0 and β are unchanged. Only v_1 is affected. We repeated the fit including such a term. The likelihood improves, however the values of α and γ are inside the errors the same. Also the results contained in figures 10 and 11 remain the same.

6   Conclusions

In our GARCH model we use a two component form of the conditional covariance matrix H which avoids inversion of H and the problems of covariance targeting. The resulting H turns out to be similar as that of the DCC model. Apart from a small number of parameters and a computing time $T_{comp} \propto N$ our RMG has the advantage that daily β values relative to the market are determined. As an application we study a possible transition of the S&P market in 2006 observed earlier by (Raddant and Wagner, 2017). We repeated this analysis with a better time resolution using the $\beta(t)$ from RMG and found the same result. A second application refers to the time correlation between returns r_{ti} and the volatility (leverage effect). When one replaces r_{ti} by the de-garched returns $\eta_{Gi}(t)$ the leverage effect disappears.

Some improvement of the model could probably be achieved by extending the model to a three-factor version, in which the non-market volatility is described in a bit more detail. This might also improve the acf of de-garched returns.

References

Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics 31*, 282–299.

Alexander, C. (2001). A primer on the orthogonal GARCH model. *mimeo*.

Bauwens, L., S. Laurent, and J. V. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics 21*(1), 79–109.

Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy 96*(1), 116–131.

Diebold, F. X. and M. Nerlove (1989). The dynamics of exchange rate volatility: A multivariate latent factor ARCH. *Journal of Applied Econmetrics 4*, 1–22.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*(4), 987–1007.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Engle, R. F. (2009). *Anticipating correlations*. Princeton UP.

Engle, R. F. (2014). Dynamic conditional beta. *Working paper*.

Engle, R. F. and F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Economic Theory 11*, 122–150.

Engle, R. F. and J. Mezrich (1996). GARCH for groups. *Risk 9*, 36–40.

Engle, R. F., N. Shephard, and K. Sheppard (2008). Fitting vast dimensional time-varying covariance models. NYU Working Paper FIN-08-009, NYU.

Franses, P. H. and C. Hafner (2009). A generalized dynamic conditional correlation model for many asset returns. *Econometric Reviews 28*, 612–631.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Laloux, L., P. Cizeau, J. Bouchaud, and P. M (1999). Noise dressing of financial correlation matrices. *Phys.Rev.Lett. 83*, 1467.

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.

Marchenko, V. A. and L. A. Pastur (1967). Distribution of eigenvalues for some sets of random matrices. *Mat. Sb. (N.S.) 72(114)(4)*, 507–536.

Raddant, M. and F. Wagner (2017). Transitions in the stock markets of the US, UK and Germany. *Quantitative Finance 17(2)*.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance 44(5)*, 1115–1153.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance 19*, 425–442.

Wagner, F., M. Milakovic, and S. Alfarano (2010). What distinguishes individual stocks from the index? *Eur. Phys. J. B 73*, 23–28.

Appendix: Derivation of the Recursions

In the recursion equation (18) we separate the off-diagonal elements in α and γ

$$
\begin{aligned}
H_{t+1} &= H_t + \sum_{\nu=0,1} P_\nu(t) \cdot \left[\alpha_{\nu\nu}(r_t \cdot r'_t - H_t) + \gamma_{\nu\nu}(\bar{H} - H_t)\right] \cdot P_\nu(t) \qquad (27)\\
&+ P_0(t) \cdot \left[\alpha_{10}(r_t \cdot r'_t + \gamma_{10}\bar{H}\right] \cdot P_1(t) \ + \ P_1(t) \cdot \left[\alpha_{10}(r_t \cdot r'_t + \gamma_{10}\bar{H}\right] \cdot P_0(t)
\end{aligned}
$$

we notice that $\beta(t+1)$ is obtained from $\beta(t)$ by a rotation on the N dimensional sphere with an angle φ

$$
\beta(t+1) = \cos\varphi \ \beta(t) \ + \ \sin\varphi \ \Delta \qquad (28)
$$

with $\Delta^2 = N$ and $\Delta' \cdot \beta(t) = 0$. Inserting (28) into H_{t+1} we get

$$
\frac{1}{N}tr[H_{t+1} \cdot P_\nu(t)] = v_\nu(t+1) - (-)^\nu \sin^2\varphi \ A(t+1) \qquad (29)
$$

and

$$
\frac{1}{N}[H_{t+1} \cdot \beta(t)]_i = v_0(t+1)\beta_i(t) + \sin\varphi \ A(t+1)(\cos\varphi \ \Delta_i - \sin\varphi\beta_i(t)) \qquad (30)
$$

with the abbreviation $A(t+1) = v_0(t+1) - v_1(t+1)/(N-1)$. Applying $\frac{1}{N}tr[P_\nu(t)$ onto the r.h.s of equation (27) we obtain the recursions for $v_\nu(t+1)$ with $\bar{\alpha}_\nu = 1 - \alpha_{\nu\nu} - \gamma_{\nu\nu}$

$$
v_\nu(t+1) - (-)^\nu \sin^2\varphi \ A(t+1) = \bar{\alpha}_\nu \ v_\nu(t) + \alpha_{\nu\nu}\rho^2_\nu(t) + \gamma_{\nu\nu}\frac{1}{N}tr[\bar{H} \cdot P_\nu(t)] \qquad (31)
$$

where the observed returns appear in the combinations given by equation (20). We denote application of $\beta(t)_i/N$ on the r.h.s. of equation (27) by D_i. D_i is given by

$$
D_i = \alpha_{10}r_{Mt}(r_{ti} - \beta_i(t)r_{Mt}) + \frac{\gamma_{10}}{N}\left([\bar{H} \cdot \beta(t)]_i - \beta_i(t)tr[\bar{H} \cdot P_0]\right)
$$

Setting D_i equal to equation (30) and using equation (31) we find an expression for Δ

$$
\sin 2\varphi \ A(t+1) \ \Delta_i = 2D_i \qquad (32)
$$

Due to $\Delta^2 = N$ equation (32) allows to express φ in terms of $A(t+1)$ and quantities known at t. Then from equation (31) $v_\nu(t+1)$ can be computed. It is interesting to note that $\beta(t)' \cdot D$ holds. Therefore the recursion (27) preserves the norm of β. For the recursion for $\beta(t+1)$ we finally get

$$
\beta_i(t+1) = \cos\varphi \ \beta_i(t) + \frac{1}{A(t+1)\cos\varphi}D_i \qquad (33)
$$

For the terms involving the unconditional expectation values $\bar{H}$ we use a decomposition analogues to H_t

$$
\bar{H} = \left(\bar{v}_0 - \frac{\bar{v}_1}{N-1}\right)\bar{\beta} \cdot \bar{\beta}' \ + \ \frac{N}{N-1}\bar{v}_1 \cdot 1 \qquad (34)
$$

Only the leading eigenvalue $N\bar{v}_0$, its eigenvector $\bar{\beta}$ and $tr[\bar{H}] = N(\bar{v}_0 + \bar{v}_1)$ are needed. As shown in Raddant and Wagner (2017) these terms can be extracted from the observed covariance matrix with statistical error of order $1/\sqrt{T}$. In the recursions the following expressions are required

$$\frac{1}{N}tr[\bar{H} \cdot P_\nu(t)] = \bar{v}_\nu - (-)^\nu (1 - m^2)\bar{A} \tag{35}$$

$$\frac{1}{N}[\bar{H} \cdot \beta(t)]_i = m\bar{A}\bar{\beta}_i + \frac{\bar{v}_1}{N-1}\beta_i(t) \tag{36}$$

with $\bar{A} = \bar{v}_0 - \bar{v}_1/(N-1)$ and $m = \bar{\beta}' \cdot \beta(t)/N$ denoting the overlap between $\beta(t)$ and $\bar{\beta}$. For the recursion initial values are needed. These are obtained in a similar way as $\bar{v}_\nu$ and $\bar{\beta}$ from the covariance matrix determined in the first four years.

The log likelihood with noise distribution $f(\eta)$ can be computed using

$$\eta = H_t^{-1/2} \cdot r_t \tag{37}$$

With the spectral decomposition for H_t we obtain

$$\eta = \frac{r_{Mt}}{\sqrt{Nv_0(t)}} \beta + \sqrt{\frac{N-1}{Nv_1(t)}} (r_t - r_{Mt}\beta) \tag{38}$$

The log likelihood L is given up to an uninteresting constant by

$$L = \frac{1}{2}\sum_t \left[\sum_i \ln f(\eta_i) - \ln v_0(t) - (N-1)\ln v_1(t)\right] \tag{39}$$

For Gaussian noise calculation of the $\ln(f)$ can be avoided. The complication of using a Student's t-distribution leads to an negligible increase of computing time compared to the calculation of $v_n(t))$ and $\beta(t)$. In any case the computing time increases only with $T \cdot N$.