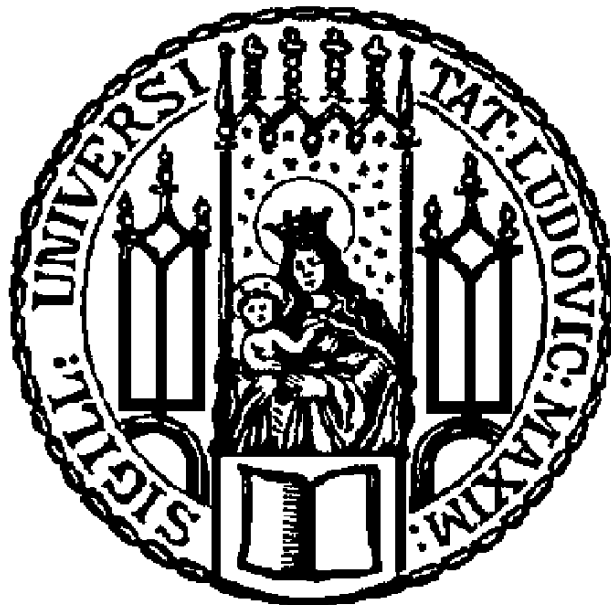


Sommersemester 2016

Multivariate dynamische Modellierung von Finanzmarktrisiken

Bachelorarbeit



Verfasser: Thomas Berger
Betreuer: Prof. Stefan Mittnik, Phd
Prüfer: Prof. Stefan Mittnik, Phd
Institut für Statistik
Ludwig-Maximilians-Universität München
Abgabedatum: 17. August 2016

Inhaltsverzeichnis

1	Einleitung	4
2	Grundlagen zur Analyse von Renditezeitreihen	5
2.1	Schwache Stationarität	5
2.2	Serielle Korrelation	7
3	Modellierung von Finanzzeitreihen	10
3.1	Grundlegender Aufbau	10
3.2	Modelle für univariate Zeitreihenanalyse	11
3.2.1	Autoregressives Modell	11
3.2.2	GARCH Modell	12
3.3	Multivariate Volatilitätsmodelle	14
3.3.1	Exponential weighted moving average	14
3.3.2	Dynamical Conditional Correlation	15
4	Risikomanagement	19
4.1	Value at Risk	19
4.1.1	Historische Simulation	20
4.1.2	Parametrische Methoden	21
4.2	Evaluierungsmethoden	22
4.2.1	Unconditional Coverage Test	22
4.2.2	Unabhängigkeitstest	23
4.2.3	Conditional Test	25
5	Datenanalyse	26
5.1	Vorstellung und Aufbereitung der Daten	26
5.2	Deskriptive Analyse	27
5.3	Value at Risk basierende Modellanalyse	30
6	Kritik und Fazit	39
7	Appendix	42
7.1	Datenbeschreibung	42
7.2	Deskriptive Analyse	44
7.3	Modellbasierte Analyse	46
7.3.1	Notation	46
7.3.2	VaR Grafiken	47

Zusammenfassung

Diese Bachelorarbeit beschäftigt sich mit der multivariaten Modellierung von Finanzmarktrisiken. Zuerst werden Grundlagen der Zeitreihenanalyse beschrieben und diese anhand Beispiele erläutert. Danach wird die Modellierung von Finanzzeitreihen erklärt. Dabei wird besonders auf die multivariaten Modelle eingegangen. Anschließend wird sich mit dem Value at Risk und seinen Evaluierungsmethoden auseinandergesetzt. Schließlich wird bei der Datenanalyse eine deskriptive Beschreibung der Daten erfolgen, welche für die darauffolgende Modellanalyse unerlässlich ist. Bei der Modellanalyse wird bei einem gleichgewichteten Portfolio von 14 ETFs eine Value at Risk Prädiktion durchgeführt. Die Güte der Modelle wird dabei mit Evaluierungsmethoden geprüft.

1 Einleitung

Wenn ein Investor eine Kapitalanlage erwirbt, ist er nicht nur an der zukünftig erwarteten Rendite, sondern vor allem auch am Risiko der Anlage interessiert. Die Volatilität ist ein Maß für das Risiko einer Anlage und ist für die zeitliche Schwankung, zum Beispiel einer Aktie, verantwortlich. In dieser Arbeit wird die Volatilität über die bedingten Standardabweichungen ihrer Log>Returns definiert. Die zeitliche Änderung der Volatilität ist eine typische Eigenschaft von Renditezeitreihen, welche unter dem Begriff Heteroskedastizität bekannt ist. Dies bedeutet, dass es wahrscheinlicher ist, dass größere Ausschläge auf große und dementsprechend niedrigere Ausschläge auf niedrige folgen. Diesen stilisierten Fakt nennt man Volatilitätscluster (vgl. R.S. Tsay (2013;S.170)).

Weil eine direkte Beobachtung von Volatilitäten nicht möglich ist, hat die Modellierung von Volatilitäten und deren Prognose eine beträchtliche Relevanz. Das *autoregressive conditional heteroskedasticity* (ARCH)-Modell von Robert Engle (1982) war das erste Modell, welches eine bedingte Varianz für die Renditen einer Kapitalanlage zuließ und hat die Finanzzeitreihenanalyse maßgeblich geprägt. Bollerslev(1986) erweiterte dieses Modell zum *generalized* ARCH-Modell (GARCH). Die GARCH Familie und ihre Erweiterungen gehören zur Kategorie der univariaten bedingten Volatilitätsmodelle. Für die gemeinsame Betrachtung von mehreren Kapitalanlagen sind für die Modellierung von zeitabhängigen Kovarianzmatrizen allerdings multivariate Volatilitätsmodelle notwendig. Das möglichst adäquate Prognostizieren der Kovarianzmatrix ist von enormer Wichtigkeit in Anwendungen der Anlageaufteilung und Portfolioauswahl, um bevorstehende Risiken adäquat zu kalkulieren (vgl. J. Danielsson (2011;S.57)).

Im Folgenden werden in Kapitel 2 die grundlegenden Konzepte der Zeitreihenanalyse wie die Stationarität erläutert. Danach folgt in Kapitel 3 die theoretische Beschreibung der in der Analyse hauptsächlich verwendeten Modelle. In Kapitel 4 wird der im Risikomanagement populäre Value at Risk und verschiedene Backtesting Methoden erläutert. Schließlich erfolgt in Kapitel 5 die Datenanalyse, welche mit der deskriptiven Analyse beginnt und anschließend in die Modellanalyse übergeht. Dabei erfolgt eine Berechnung des Value at Risk mit mehreren Modellen. Diese werden mit Backtesting Verfahren überprüft und ausgewertet. Dem folgt zum Schluss noch ein abschließendes Fazit.

2 Grundlagen zur Analyse von Renditezeitreihen

2.1 Schwache Stationarität

Eine Zeitreihe ist eine nach der Zeit chronologisch geordnete Folge von Beobachtungen, wie zum Beispiel die täglichen Schlusskurse des Deutschen Aktienindex (DAX). Bei der Zeitreihenanalyse wird generell versucht, die typischen Eigenschaften wie die serielle Korrelation oder Volatilitätscluster abzubilden. Um jedoch mit einer Analyse zu beginnen, muss die jeweilige Zeitreihe in einer für Modellierungsprozesse passenden Form vorliegen. Dabei ist es meistens notwendig, dass die Zeitreihe schwach stationär ist. Sei Y_1, Y_2, \dots eine schwach stationäre Zeitreihe, dann gilt:

- $E(Y_t) = \mu$ eine endliche Konstante für alle t
- $Var(Y_t) = \sigma$ eine positive Konstante für alle t
- $Cov(Y_t, Y_s) = \gamma(|t - s|)$ für alle t und s für die Funktion $\gamma(k)$

Also sind die Eigenschaften eines schwach stationären Prozesses, dass der Mittelwert und die Varianz nicht von der Zeit abhängen. Die Kovarianzfunktion hängt nur vom *lag*, also dem Abstand zwischen t und s ab, und nicht von der Zeit direkt. Das heißt, dass beispielsweise Y_2 und Y_5 dieselbe Kovarianz besitzen wie Y_7 und Y_{10} . Der Abstand beträgt nämlich jeweils drei Einheiten. Das Adjektiv „schwach“ bezieht sich darauf, dass nur Mittelwert, Varianz und Kovarianz als stationär angenommen werden und nicht andere verteilungsspezifische Charakteristiken, wie etwa die Schiefe, die Form oder die Quantile (vgl. D. Ruppert (2015; S.308)).

Die Funktion $\gamma(s)$ steht dabei für die Autokovarianzfunktion, welche Bestandteil der Autokorrelationsfunktion ist:

$$\rho_t = \frac{Cov(Y_t, Y_{t-s})}{\sqrt{Var(Y_t) \times Var(Y_{t-s})}} = \frac{Cov(Y_t, Y_{t-s})}{Var(Y_t)} = \frac{\gamma(s)}{\gamma(0)}$$

wobei wegen der schwachen Stationarität $Var(Y_t) = Var(Y_{t-s})$ gilt. Aus der Definition ergibt sich, dass $\rho_0 = 1$, $\rho_l = \rho_{-l}$ und $-1 \leq \rho_l \leq 1$. Die Folge $\{\rho_l\}$ wird dabei Autokorrelationsfunktion (ACF) genannt. Eine schwach stationäre Zeitreihe ist genau dann seriell unkorreliert, wenn $\rho = 0$ gilt für alle s . (vgl. R.S. Tsay (2013; S.45)).

Für eine zufällige Stichprobe $\{Y_t | t = 1, \dots, T\}$ ist die geschätzte Autokorrelationsfunktion (ACF) definiert als

$$\hat{\rho}_s = \frac{\sum_{t=s+1}^T (Y_t - \bar{y})(Y_{t-s} - \bar{y})}{\sum_{t=1}^T (Y_t - \bar{y})^2}, \quad 0 \leq s < T - 1$$

wobei der Stichprobenmittelwert $\bar{y} = \sum_{t=1}^T Y_t / T$ ist (vgl. R.S. Tsay (2013; S.45)).

Viele Finanzzeitreihen weisen oft keine schwache Stationarität auf. Die *Änderungen* einer Zeitreihe, welche nach Anwendung einer Log-Transformation vorliegen, hingegen schon. Ob eine Zeitreihe schon selbst oder aber erst durch geeignetes Transformieren schwach stationär wird, ist dabei statistisch gesehen irrelevant (vgl. D. Ruppert (2015; S.308-309)). Deswegen werden nicht die Preise $P_t, t = 0, 1, 2, \dots$ eines Wertpapiers zur Analyse verwendet, welche meistens nicht stationär sind, sondern die logarithmierten Differenzen der Preise. Diese sogenannten *Log>Returns* sind folgendermaßen definiert:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln P_t - \ln P_{t-1}$$

Diese beschreiben approximativ die relative Änderung

$$\frac{P_t - P_{t-1}}{P_{t-1}}$$

eines Preises zum Zeitpunkt t , was durch die Taylor-Approximation $\ln(1 + x) \approx x$ für betragsmäßig kleine x bewiesen wurde (vgl. J.P. Kreiss (2006; S.295)).

Der Vorteil eines schwach stationären Prozesses ist, dass er mit relativ wenigen Parametern modelliert werden kann. Wenn jedes Y_t nur aus ihrem eigenen Mittelwert μ_t ermittelt werden könnte, dann wäre eine adäquate Schätzung von μ_t nicht möglich. So hat man aber bei einem schwach stationären Prozess einen vom Zeitindex t unabhängigen Mittelwert μ , welcher durch \bar{y} erwartungstreu geschätzt werden kann (vgl. D. Ruppert (2015; S.309)).

2.2 Serielle Korrelation

Im folgenden werden anhand der wöchentlichen Log>Returns eines Indexfonds, welcher aus Unternehmensanleihen der vereinigten Staaten von Amerika besteht, beispielhaft Techniken zum Aufzeigen von serieller Korrelation verdeutlicht. Die Log>Returns belaufen sich dabei über einen Zeitraum von Mitte Januar 1999 bis Ende April 2016.

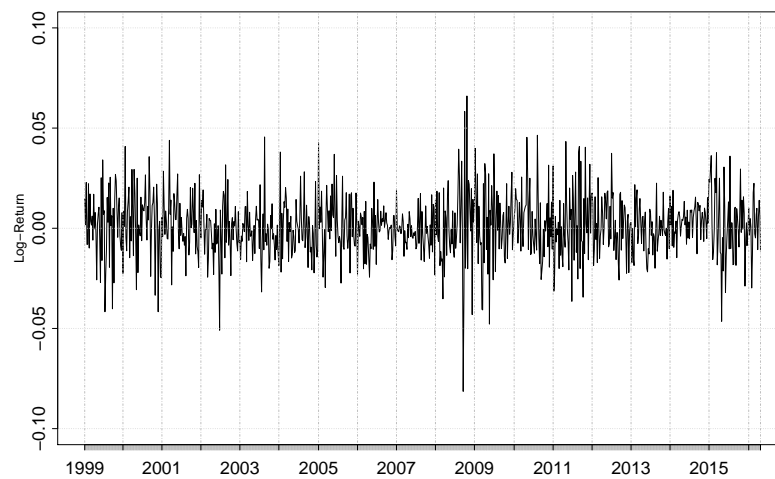


Abbildung 1: wöchentliche Log>Returns von Unternehmensanleihen der Vereinigte Staaten

An dieser Grafik kann man Zeiten mit niedriger Volatilität im Jahr 2006/2007, aber auch die Jahre mit hoher Volatilität wie zum Beispiel Mitte 2008 bis Ende 2009, erkennen. Dies charakterisiert die Volatilitätscluster, welche in der Einleitung besprochen wurden. Um festzustellen ob die Zeitreihe r_t serielle Korrelation aufweist, kann man sample ACF Plots betrachten.

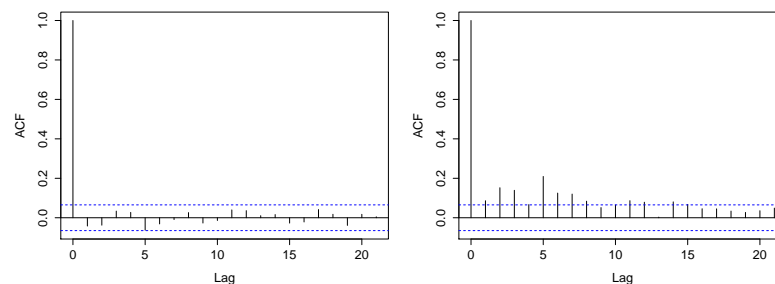


Abbildung 2: ACF-Plots

Die zwei gestrichelten blauen Linien sind Standardfehlergrenzen, die sich bei $\pm 2/\sqrt{T}$ befinden (vgl. R.S. Tsay (2013;S.45)). Diese Grenzen werden verwendet, um die Nullhypothese, dass der Autokorrelationskoeffizient Null ist, zu überprüfen. Die Nullhypothese wird abgelehnt, falls die Autokorrelationsfunktion die Grenzen überschreitet. Bei einem gewöhnlich verwendeten Signifikanzniveau von 5% erwartet man, dass bei einem von 20 Lags die Standardfehlergrenze überschritten wird (vgl. D. Ruppert (2015;S.312-313)). Das linke Diagramm in Abbildung 2 zeigt die ACF der Log>Returns. Sie weist keine signifikante serielle Korrelation auf, außer bei Lag 5. Die quadrierten Log>Returns, dargestellt im rechten Diagramm, sind andererseits offensichtlich autokorreliert. Folglich ist die Serie abhängig, aber nicht seriell korreliert. Eine weitere Methode um festzustellen ob eine Zeitreihe Autokorrelationen aufweist ist der Ljung-Box Test:

$$Q(m) = T(T+2) \sum_{l=1}^m \frac{\hat{\rho}_l^2}{T-l}$$

Die Nullhypothese ist $\rho_1 = \dots = \rho_m = 0$. Diese wird zugunsten der alternativen Hypothese $H_a : \rho_l \neq 0 \exists l \in \{1, \dots, m\}$ verworfen, falls $Q(m) > \chi_{\alpha}^2$. χ_{α}^2 ist das $100(1-\alpha)$ -Quantil einer Chi-Quadrat Verteilung mit m Freiheitsgraden (vgl. R.S. Tsay (2013;S.47-48)). Im weiteren Verlauf der Arbeit wird der p-Wert von $Q(m)$ als Richtwert bezüglich des Verwerfens oder des Nichtverwerfens der Nullhypothese herangezogen.

```
Box-Ljung test
data:  corpUs_IBOXIG
X-squared = 9.1763, df = 7, p-value = 0.2402
```

Diese Ljung-Box Statistik stimmt mit unserer Annahme überein, dass die Log>Returns keine serielle Korrelation aufweisen. Ganz im Gegensatz zur Ljung-Box Statistik der quadrierten Log>Returns:

```
Box-Ljung test
data:  corpUs_IBOXIG
X-squared = 116.62, df = 7, p-value < 2.2e-16
```

Diese bestätigt unsere Folgerungen bei den ACF-Plots. Die Ergebnisse implizieren generell, dass die Finanzzeitreihe seriell unkorreliert aber abhängig ist, weil die quadrierten Werte im Gegensatz zu den unveränderten Werten seriell korreliert sind (vgl. R.S. Tsay (2013;S.178)).

Wenn man anstatt einer univariaten Zeitreihe eine N -dimensionale multivariate Zeitreihe \mathbf{r}_t betrachtet, kann man auch hier die Präsenz von bedingter Heteroskedastizität durch einen multivariaten Ljung-Box Test feststellen. Wie in Kapitel Drei noch genauer erläutert wird, kann \mathbf{r}_t durch $\boldsymbol{\mu}_t + \mathbf{a}_t$ modelliert werden. Dabei wird im folgenden angenommen, dass $\boldsymbol{\mu}_t$ bekannt ist und der Innovationsprozess \mathbf{a}_t vorhanden ist. Weil die Volatilität eng mit dem Zweiten Moment von \mathbf{a}_t zusammenhängt, wird für den Test \mathbf{a}_t^2 benötigt. Wenn keine bedingte Heteroskedastizität vorhanden ist, wird die bedingte Kovarianzmatrix $\boldsymbol{\Sigma}_t$ als zeitunabhängig angenommen. Das heißt, dass \mathbf{a}_t^2 nicht abhängig ist von \mathbf{a}_{t-i}^2 für $i > 0$. Daraus ergibt sich die Nullhypothese zu $H_0 : \boldsymbol{\rho}_1 = \boldsymbol{\rho}_2 = \dots = \boldsymbol{\rho}_m = 0$ gegen die Alternativhypothese $H_a : \boldsymbol{\rho}_i \neq 0$ für mindestens ein $i(1 \leq i \leq m)$, wobei $\boldsymbol{\rho}_i$ die lag- i Kreuzkorrelationsmatrix von \mathbf{a}_t^2 ist (vgl.R.S.Tsay (2014;S.401)). Die multivariate Ljung-Box Statistik ist dabei definiert als:

$$Q_k^*(m) = T^2 \sum_{i=1}^m \frac{1}{T-i} \mathbf{b}_i^\top (\hat{\boldsymbol{\rho}}_0^{-1} \otimes \hat{\boldsymbol{\rho}}_0^{-1}) \mathbf{b}_i$$

Die Stichprobengröße wird durch T bestimmt, k ist die Dimension \mathbf{a}_t und $\mathbf{b}_i = \text{vec}(\hat{\boldsymbol{\rho}}_i^\top)$. Unter der Nullhypothese ist $Q_k^*(m)$ dabei asymptotisch $\chi_{k^2 m}^2$ verteilt (vgl.R.S.Tsay (2014;S.401)).

3 Modellierung von Finanzzeitreihen

3.1 Grundlegender Aufbau

Gegeben sei ein $d \times 1$ dimensionaler Vektor der Log>Returns $\mathbf{r}_t = (r_{1,t}, \dots, r_{d,t})$. Unter der Beachtung der Menge F_t , welche alle Informationen von $\mathbf{r}_t, \mathbf{r}_{t-1}, \dots$ zum Zeitpunkt t enthält, kann \mathbf{r}_t durch

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{a}_t$$

modelliert werden. $\boldsymbol{\mu}_t = E(\mathbf{r}_t | F_{t-1})$ ist der bedingte Mittelwertsvektor zum Zeitpunkt t und $\{\mathbf{a}_t\}$ eine Folge von WR-Prozessen mit Mittelwert Null und Kovarianzmatrix $\boldsymbol{\Sigma}_a = \text{Cov}(\mathbf{a}_t)$. Die bedingte Kovarianz oder Volatilitätsmatrix zum Zeitpunkt t kann bedingt auf die Informationsmenge beschrieben werden:

$$\boldsymbol{\Sigma}_t = \text{Cov}(\mathbf{a}_t | F_{t-1}) = \text{Cov}(\mathbf{r}_t | F_{t-1}).$$

Multivariates Modellieren von Zeitreihen hängt in erster Linie mit dem zeitabhängigen Modellieren von $\boldsymbol{\mu}_t$ und $\boldsymbol{\Sigma}_t$ zusammen (vgl. D. Ruppert (2015; S.441)). Wobei sich diese Arbeit hauptsächlich mit der Modellierung der bedingten Kovarianzmatrix beschäftigt.

Die Beziehung zwischen dem Innovationsprozess und dem Volatilitätsprozess ist definiert durch

$$\mathbf{a}_t = \boldsymbol{\Sigma}_t^{1/2} \mathbf{z}_t, \quad \mathbf{z}_t \stackrel{iid}{\sim} F(0, \mathbf{I}_d),$$

wobei $\boldsymbol{\Sigma}_t^{1/2}$ eine positiv definite Matrix aus der Quadratwurzel von $\boldsymbol{\Sigma}_t$ ist, sodass $\boldsymbol{\Sigma}_t^{1/2} \boldsymbol{\Sigma}_t^{1/2} = \boldsymbol{\Sigma}_t$. $\{\mathbf{z}_t\}$ ist dabei eine Folge von unabhängig und identisch verteilten Zufallsvektoren mit Erwartungswert $E(\mathbf{z}_t) = 0$ und $\text{Cov}(\mathbf{z}_t) = \mathbf{I}_d$ (vgl. R.S. Tsay (2014; S.400)).

Im folgenden werden die verschiedenen multivariaten Modelle erläutert, welche bei der Datenanalyse verwendet werden. Weil aber für $\boldsymbol{\mu}_t = (\mu_1, \dots, \mu_d)$ und $\mathbf{a}_t = (a_1, \dots, a_d)$ jeweils univariate Modelle geschätzt werden, wird davor auf die univariate Modellierung von Finanzzeitreihen eingegangen. Dabei werden erst kurz die grundlegenden Eigenschaften des Autoregressiven Prozesses gezeigt und danach wird das GARCH Modell und dessen Parameterschätzung erläutert. Daraufhin werden die multivariaten Volatilitätsmodelle beschrieben. Hierbei wird auf die verschiedenen Modellgleichungen und ihre Bestandteile eingegangen. Anschließend wird beim DCC Modell die Parameterschätzung erläutert und

auf die Verknüpfungen zum GARCH Modell eingegangen.

3.2 Modelle für univariate Zeitreihenanalyse

3.2.1 Autoregressives Modell

Da in der Datenanalyse für die Erklärung des Mittelwerts ein Autoregressives Modell (kurz AR) verwendet wird, werden in diesem Abschnitt kurz dessen Modelgleichung gezeigt und auf die Eigenschaften des AR(1) Modells eingegangen. Die Modellgleichung eines AR(p)-Modells sieht folgendermaßen aus:

$$r_t = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + a_t$$

Hier ist r_t der Log Return einer Kapitalanlage zum Zeitpunkt t und a_t ist ein WR Prozess mit Mittelwert Null und Varianz σ_a^2 . Eine Zeitreihe x_t wird *Weißes Rauschen*(WR) genannt, falls $\{x_t\}$ ein Folge von unabhängig identisch verteilten Zufallsvariablen ist (vgl.R.S.Tsay (2013;S.50-51)).

In der späteren Datenanalyse wird mit einem AR(1) Modell

$$r_t = \phi_0 + \phi_1 r_{t-1} + a_t$$

gearbeitet, bei dem r_t bedingt auf den r_{t-1} folgende Eigenschaften bezüglich des Erwartungswerts und der Varianz besitzt:

$$E(r_t|r_{t-1}) = \phi_0 + \phi_1 r_{t-1}, \quad Var(r_t|r_{t-1}) = Var(a_t) = \sigma_a^2$$

Bei einem Ar(1) Modell ist r_t also um $\phi_0 + \phi_1 r_{t-1}$ zentriert mit einer Varianz von σ_a^2 . Bei diesem Modell hängt also der Wert r_t von seinen vergangenen gewichteten Werten ab. ϕ_0 und ϕ_1 können dabei durch die Maximum Likelihood Methode geschätzt werden.

3.2.2 GARCH Modell

Weil bei den multivariaten Modellen in Kapitel 3 ein GARCH Modell bei der Modellierung der bedingten Volatilität verwendet wird, wird in diesem Abschnitt der GARCH Prozess, seine Parameterschätzung und die wichtigsten Eigenschaften kurz erläutert.

Das GARCH Modell wurde von Bollerslev(1986) als Erweiterung zum ARCH Modell von Engle(1982) eingeführt. Sei $a_t = r_t - \mu_t$ die Innovation zum Zeitpunkt t , dann ist das GARCH(p,q) Modell definiert als:

$$a_t = h_t^{1/2} z_t, \quad h_t = \omega + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}$$

$\{z_t\}$ ist hier eine Folge von unabhängigen Zufallsvariablen mit Mittelwert 0 und Varianz 1. Es gilt $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ und $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$. Die letzte Bedingung mit α_i und β_i impliziert, dass die unbedingte Varianz von a_t endlich ist, wohingegen σ_t über die Zeit variiert. Die Verteilung von z_t kann eine Normalverteilung oder auch eine Verteilung mit schweren Rändern, wie die T-Verteilung sein (vgl.R.S.Tsay (2013;S.199)).

Wie man anhand der Modellgleichung erkennt, variiert die bedingte Varianz h_t über die Zeit und ist abhängig von den vergangenen quadrierten Innovationen $\{a_{t-i}^2\}_{i=1}^p$. Falls diese groß sind tendieren sie dazu, dass die bedingte Varianz h_t auch groß ist, aus welchem wiederum ein großes a_t folgt und vice versa. Dies modelliert den stilisierten Fakt der Volatilitätscluster, welcher in der Einleitung besprochen wurde.

Eine Ein Schritt Prognose für ein GARCH(1,1) Modell

$$h_t = \omega + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}$$

ergibt sich, wenn l der Ausgangszeitpunkt der Vorhersage ist

$$h_{l+1} = \omega + \alpha_1 a_l^2 + \beta_1 h_l$$

Die Parameter eines GARCH(1,1) werden über die Maximum Likelihood Methode berechnet. Die Likelihood ergibt sich unter der Annahme, dass z_t normalverteilt ist:

$$L(\alpha, \beta | a_m, \dots, a_n) = f(a_{m+1}, \dots, a_n | \alpha, \beta, a_1, \dots, a_m) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi h_t}} \exp \left\{ \frac{-a_t^2}{2h_t} \right\}$$

Wobei $t = 1, \dots, T$ die Zeitpunkte der bedingten Likelihood sind. Die Likelihood zu maxi-

mieren ist äquivalent zur Maximierung der Log-Likelihood, welche oft einfacher zu berechnen ist:

$$\begin{aligned} \ln(L) &= l(a_{m+1}, \dots, a_n | \alpha, \beta, a_1, \dots, a_m) \\ &= -\frac{1}{2} \sum_{t=1}^T \left[\ln(2\pi) + \ln(h_t) + \frac{a_t^2}{h_t} \right] \\ &= -\frac{1}{2} \sum_{t=1}^T \left[\ln(h_t) + \frac{a_t^2}{h_t} \right] + \text{constant} \end{aligned}$$

Die Maximierung dieses Terms geschieht meistens über numerische Optimierungsmethoden.

Ein Schwäche des GARCH Modells ist, dass es auf positive und negative Schocks gleich reagiert. Es nimmt also nicht Rücksicht auf den sogenannten "Leverage Effect". Bei diesem Effekt zieht ein negativer Marktschock eine höhere Volatilität nach sich, als ein positiver. Dies wird bei einigen Erweiterungen zum GARCH Modell, wie dem GJR-GARCH (Glosten et al. (1993)) berücksichtigt. Das GJR-GARCH(p,q) Modell ist definiert als:

$$h_t = \omega + \sum_{i=1}^p \left(\alpha_i a_{t-i}^2 + \gamma_i I_{t-i} a_{t-i} \right) \sum_{j=1}^q \beta_j h_{t-j}$$

wobei γ den Leverage Term repräsentiert. Die Indikatorfunktion I nimmt den Wert 1 an, falls $a \leq 0$ ansonsten Null.

3.3 Multivariate Volatilitätsmodelle

3.3.1 Exponential weighted moving average

Das Exponential weighted moving average Modell (kurz EWMA) ist eine einfache Anwendung um zeitvariierende Volatilitäten zu schätzen:

$$\begin{aligned}\boldsymbol{\Sigma}_t &= \lambda \boldsymbol{\Sigma}_{t-1} + (1 - \lambda) \mathbf{a}_{t-1} \mathbf{a}_{t-1}^\top \\ &= (1 - \lambda) \sum_{l=1}^{\infty} \lambda^{l-1} \mathbf{a}_{t-l} \mathbf{a}_{t-l}^\top\end{aligned}$$

wobei die Bedingung $0 < \lambda < 1$ für den Persistenzparameter λ gilt. Die bedingte Kovarianzmatrix kann dabei aus der Kovarianzmatrix und dem quadrierten Marktschock \mathbf{a}_{t-1} der Vorperiode gebildet werden. Ein niedriger Wert des Persistenzparameter λ führt dazu, dass ältere Änderungen der Volatilität wenig Einfluss auf die geschätzte Volatilität haben. Ein hohes λ steht dementsprechend für eine hohe Persistenz im Volatilitätsprozess. Wenn man die Rekursion mit einer positiv definiten Kovarianzmatrix $\boldsymbol{\Sigma}_s$ startet, ist $\boldsymbol{\Sigma}_t$ positiv definit für alle t (vgl. D. Ruppert (2015; S.432)).

Dieses Modell ist extrem sparsam und die resultierenden Volatilitätsmatrizen sind leicht zu aktualisieren, aber das Modell tendiert dazu jegliche Gütetests abzulehnen. Dies ist nicht überraschend, weil es schwer vorstellbar ist, dass ein einziger Persistenzparameter die Veränderung aller bedingten Varianzen und Kovarianzen adäquat bestimmt (vgl. R. S. Tsay (2014; S.414)).

In dieser Arbeit wird ein $\lambda = 0.96$ angenommen. Die Startwerte der ersten Kovarianzmatrix werden mit der Stichprobenkovarianzmatrix der In-sample Daten ermittelt.

3.3.2 Dynamical Conditional Correlation

- **Modellgleichung und Bedingungen**

Nichtlineare Kombinationen von univariaten Volatilitätsmodellen wurden aufgestellt um zeitvariierende Korrelationen zu erlauben. Engle(2002) generalisierte das Modell von Bollerslev(1990) um Dynamical Conditional Correlation (DCC) zu ermöglichen. In diesem multivariaten GARCH Modell wird Σ_t durch \mathbf{D}_t und \mathbf{R}_t

$$\Sigma_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t$$

beschrieben. Dieses Modell teilt das multivariate Modellieren in zwei Schritte auf. Im ersten Teil wird $\mathbf{D}_t = \text{diag}\{\sqrt{h_{1,t}}, \dots, \sqrt{h_{d,t}}\}$ modelliert, welche die Standardabweichungen enthält. Diese können zum Beispiel durch ein GARCH(1,1)-Modell

$$h_{i,t} = \omega_i + \alpha_i a_{i,t-1} + \beta_i h_{i,t-1}$$

geschätzt werden. Um die bedingten Standardabweichungen von \mathbf{D}_t zu modellieren, können beliebige GARCH-Modelle mit unterschiedlicher Ordnung verwendet werden. In dieser Arbeit wurde aber für die Diagonalelemente \mathbf{D}_t ausschließlich das GARCH(1,1)-Modell und das GJR-GARCH(1,1)-Modell verwendet (vgl.D.Ruppert (2015;S.440)).

Im zweiten Schritt wird dann die dynamische Abhängigkeit der Korrelationsmatrix modelliert. \mathbf{R}_t ist hierbei die bedingte Korrelationsmatrix der skalierten Störterme,

$$\boldsymbol{\varepsilon}_t = \mathbf{D}_t^{-1} \mathbf{a}_t \sim N(\mathbf{0}, \mathbf{R}_t)$$

wobei \mathbf{R}_t symmetrisch ist, weil es eine Korrelationsmatrix ist.

$$\mathbf{R}_t = \begin{pmatrix} 1 & \rho_{12,t} & \dots & \rho_{1n,t} \\ \rho_{12,t} & 1 & & \vdots \\ \vdots & & \ddots & \rho_{n-1,n,t} \\ \rho_{1n,t} & \dots & \rho_{n-1,n,t} & 1 \end{pmatrix}$$

Weil Σ_t als Kovarianzmatrix positiv definit sein muss, ist dies auch eine Bedingung die für \mathbf{R}_t bei der Modellierung beachtet werden muss (\mathbf{D}_t ist positiv definit, da die Diagonalelemente positiv sind). Außerdem ist es wichtig, dass \mathbf{R}_t als Korrelations-

matrix nur Elemente enthält, die zwischen -1 und 1 liegen. Das DCC-Model von Engle (2002) gewährleistet diese Bedingungen durch

$$\mathbf{Q}_t = (1 - a - b)\bar{\mathbf{Q}} + a\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}^\top + b\mathbf{Q}_{t-1}$$

$$\mathbf{R}_t = \text{diag}(\mathbf{Q}_t)^{-1/2}\mathbf{Q}_t\text{diag}(\mathbf{Q}_t)^{-1/2}$$

, wobei $\bar{\mathbf{Q}}$ die unbedingte Kovarianzmatrix von $\boldsymbol{\varepsilon}_t$ ist. a und b sind positive reelle Zahlen mit der Bedingung $0 < a + b < 1$, welche Stationarität und positive Definitheit von \mathbf{Q}_t versichert. Die Matrix \mathbf{Q}_t muss dabei reskaliert werden um eine adäquate Korrelationsmatrix \mathbf{R}_t zu erzeugen (vgl. R.S. Tsay (2014;S.428)).

Das DCC-Modell reagiert allerdings nicht spezifisch auf positive oder negative Schocks. Das Asymmetric Dynamical Conditional Correlation Modell (kurz: ADCC) geht dabei auf den Leverage-Effekt der Schocks ein. Das heißt, dass die Korrelationen ähnlich wie Volatilitäten asymmetrische Reaktionen auf positive und negative Schocks zeigen. Korrelationen sind zum Beispiel höher wenn beide Renditen negativ sind, als wenn sie bei gleicher Höhe positiv sind (vgl. L. Capiello et al. (2006;S.540-541)). Um diesen Effekt bei der Modellierung zu berücksichtigen wird beim ADCC Modell die Gleichung für \mathbf{Q}_t angepasst:

$$\mathbf{Q}_t = (\bar{\mathbf{Q}} - a\bar{\mathbf{Q}} - b\bar{\mathbf{Q}} - g\bar{\mathbf{N}}) + a\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}^\top + g\mathbf{n}_{t-1}\mathbf{n}_{t-1}^\top + b\mathbf{Q}_{t-1}$$

Wobei $\mathbf{n}_t = I[\boldsymbol{\varepsilon}_t < 0] \circ \boldsymbol{\varepsilon}_t$ ($I[\cdot]$ ist dabei eine $d \times 1$ Indikatorfunktion, welche den Wert 1 annimmt, falls die Bedingung wahr ist, sonst wird sie 0. " \circ " bezeichnet das Hadamard-Produkt) und $\bar{\mathbf{N}}_t$ wird empirisch durch $\bar{\mathbf{N}}_t = \frac{1}{T} \sum_{t=1}^T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^\top$ geschätzt (vgl. L. Capiello et al. (2006;S.543-545)).

Die folgende Bedingung ist notwendig und hinreichend für \mathbf{Q}_t um positiv definit zu sein:

$$a + b + \delta g < 1$$

wobei δ der maximale Eigenwert von $\bar{\mathbf{Q}}^{-1/2}\bar{\mathbf{N}}\bar{\mathbf{Q}}^{-1/2}$ ist. Wenn man $g = 0$ setzt, reduziert sich die Gleichung auf das DCC Modell. Bei der späteren Datenanalyse werden sowohl das DCC als auch das ADCC Modell verwendet.

- **Parameterschätzung**

In diesem Abschnitt wird beschrieben, wie die Parameter des DCC-Modells determiniert werden. Unter der Annahme, dass $\mathbf{a}_t | F_{t-1} \sim N(0, \mathbf{D}_t \boldsymbol{\Sigma}_t \mathbf{D}_t)$ gilt, wird zur Parameterschätzung für die multivariate Normalverteilung die Quasi Maximum Likelihood Schätzung verwendet. In der späteren Datenanalyse wird für \mathbf{a}_t auch eine multivariate t-Verteilung angenommen. In diesem Abschnitt wird die Parameterschätzung beispielsweise für die multivariate Normalverteilung gezeigt. Unter der Annahme der multivariaten Normalverteilung ergibt sich somit folgende Log-Likelihood (vgl. R.F. Engle (2002; S.342)):

$$\begin{aligned}
 L &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log |\boldsymbol{\Sigma}_t| + \mathbf{a}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{a}_t) \\
 L &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log |\mathbf{D}_t \mathbf{R}_t \mathbf{D}_t| + \mathbf{a}_t^\top \mathbf{D}_t^{-1} \mathbf{R}_t^{-1} \mathbf{D}_t^{-1} \mathbf{a}_t) \\
 L &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + 2 \log |\mathbf{D}_t| + \log |\mathbf{R}_t| + \boldsymbol{\varepsilon}_t^\top \mathbf{R}_t^{-1} \boldsymbol{\varepsilon}_t) \\
 L &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + 2 \log |\mathbf{D}_t| + \mathbf{a}_t^\top \mathbf{D}_t^{-1} \mathbf{D}_t^{-1} \mathbf{a}_t - \boldsymbol{\varepsilon}_t^\top \boldsymbol{\varepsilon}_t + \log |\mathbf{R}_t| + \boldsymbol{\varepsilon}_t^\top \mathbf{R}_t^{-1} \boldsymbol{\varepsilon}_t)
 \end{aligned}$$

Durch das Addieren und Subtrahieren von $\mathbf{a}_t^\top \mathbf{D}_t^{-1} \mathbf{D}_t^{-1} \mathbf{a}_t = \boldsymbol{\varepsilon}_t^\top \boldsymbol{\varepsilon}_t$ kann die Log-Likelihood in einen Volatilitätsterm und Korrelationsterm aufgeteilt werden. Wenn θ die Parameter der Volatilitätsgleichung und ϕ die zusätzlichen Parameter der Korrelationsgleichung repräsentiert, ergibt sich für die Summe der Log-Likelihood aus den beiden Termen:

$$L(\theta, \phi) = L_v(\theta) + L_c(\theta, \phi)$$

Der Volatilitätsterm ist

$$L_v(\theta) = -\frac{1}{2} \sum_t (n \log(2\pi) + \log |\mathbf{D}_t|^2 + \boldsymbol{\varepsilon}_t^\top \mathbf{D}_t^{-2} \boldsymbol{\varepsilon}_t)$$

und der Korrelationsterm

$$L_c(\theta, \phi) = -\frac{1}{2} \sum_t (\log |\mathbf{R}_t| + \boldsymbol{\varepsilon}_t^\top \mathbf{R}_t^{-1} \boldsymbol{\varepsilon}_t + \boldsymbol{\varepsilon}_t^\top \boldsymbol{\varepsilon}_t).$$

Der Volatilitätspart ist dabei die Summe der einzelnen GARCH-Likelihoods

$$L_v(\theta) = -\frac{1}{2} \sum_t \sum_{i=1}^N \left(\log(2\pi) + \log(h_{i,t}) \frac{a_{i,t}^2}{h_{i,t}} \right),$$

welche im multivariaten normalverteilten Fall getrennt maximiert werden können. Bei anderen multivariaten Verteilungen, wie der multivariaten t-Verteilung, muss aufgrund der Existenz eines Skalenparameters die Maximierung in einem Schritt durchgeführt werden, sodass der Skalenparameter für alle Modelle gleichzeitig geschätzt wird. Die Aufteilung der Schätzung in zwei Teilen ermöglicht eine schnellere Maximierung der Log-Likelihood.

- **DCC-Prognose**

Diese Arbeit begnügt sich mit der Ein-Schritt Prognose. Im Folgendem wird kurz das Problem der Mehrfachprognose erläutert. Aufgrund der Nichtlinearität des DCC Entwicklungsprozesses kann die Mehrfachprognose der Korrelation nicht direkt gelöst werden und wird dagegen auf Basis einer Approximation geschätzt (vgl.R.F.Engle, K.Sheppard (2001,S.19-21)).

\mathbf{Q}_{t+n} ergibt sich hierbei aus:

$$\mathbf{Q}_{t+n} = (1 - a - b)\bar{\mathbf{Q}}_t + aE_t(\mathbf{z}_{t+n-1}\mathbf{z}_{t+n-1}^\top) + b\mathbf{Q}_{t+n-1}$$

Hier ist $E_t(\mathbf{z}_{t+n-1}\mathbf{z}_{t+n-1}^\top) = \mathbf{R}_{t+n-1}$ und $\mathbf{R}_{t+n} = \text{diag}(\mathbf{Q})_{t+n}^{-1/2} \mathbf{Q}_{t+n} \text{diag}(\mathbf{Q})_{t+n}^{-1/2}$. Eine Approximation um \mathbf{R}_{t+n} ergibt sich, wenn man $\bar{\mathbf{Q}} \approx \mathbf{R}$ und $E_t[\mathbf{Q}_{t+1}] = E_t[\mathbf{Q}_{t+n}]$ festlegt.

$$E_t(\mathbf{R}_{t+n}) = \sum_{i=0}^{n-2} (1 - a - b)\bar{\mathbf{R}}_t (a + b)^i + (a + b)^{n-1} \mathbf{R}_{n+1}$$

4 Risikomanagement

4.1 Value at Risk

Da bei der Datenanalyse die Prädiktion des Value at Risk mit den in Kapitel 3 und 4 erläuterten Modellen im Vordergrund steht, wird in diesem Kapitel kurz der Value at Risk erklärt und danach auf nicht parametrische und parametrische Methoden des Value at Risk eingegangen. Der Value at Risk (VaR) ist wahrscheinlich das bekannteste Risikomaß. Entwickelt wurde es 1994 von Dennis Weatherstone, dem damaligen CEO von JPMorgan Chase, um mit einer bestimmten Wahrscheinlichkeit Aussage darüber treffen zu können, welcher Verlust in einem bestimmten Zeitraum nicht überschritten wird (vgl.C.Peitz (2016;S.61)).

Der Value at Risk braucht also genau zwei Parameter, den Zeitraum T und das Konfidenzniveau $1 - \alpha$. Diese Koeffizienten gegeben ist der Value at Risk eine Grenze, welche besagt, dass der Verlust über den Zeitraum die Grenze zu einer Wahrscheinlichkeit gleich dem Konfidenzniveau, nicht überschreitet. Zum Beispiel wenn der Zeitraum eine Woche beträgt, der Konfidenzparameter 99% ($\alpha = 1\%$) und der VaR 5 Millionen Euro, so beträgt die Wahrscheinlichkeit, dass der Verlust 5 Millionen über den Zeitraum einer Woche überschreitet, 1%. Wenn \mathcal{L} der Verlust über die Haltedauer T ist, dann ist der $\text{VaR}(\alpha)$ das obere α -Quantil von \mathcal{L} (vgl.D.Ruppert (2015;S.553-554)).

Für eine stetige Verlustverteilung \mathcal{L} ergibt sich der $\text{VaR}(\alpha)$ aus

$$\text{VaR}(\alpha) = \inf(x : P(\mathcal{L} > x) \leq \alpha)$$

Dabei bezeichnet $\inf(S)$ die größte untere Schranke von S , zum Beispiel die größte Zahl die kleiner oder gleich S ist. Im folgenden wird die Annahme, dass die Log>Returns r_t stationär sind, getroffen. Das ist normalerweise eine angemessene Prämisse. Außerdem wird für die historische Simulation (HS) angenommen, dass die Log>Returns unabhängig sind. Dies ist aufgrund der Volatilitätscluster eine kritische Annahme (vgl.D.Ruppert (2015;S.555)).

Im folgenden liegt der Fokus auf die Historische Simulation für ein Portfolio von Wertpapieren. Danach wird auf die parametrischen Methoden wie das Multivariate GARCH für die Value at Risk Berechnung eingegangen. Die Ergebnisse werden später in der Datenanalyse präsentiert.

4.1.1 Historische Simulation

Eine Methode zur Bestimmung des VaR ist die historische Simulation. Dies ist eine nicht-parametrische Anwendung, das heißt, es wird für die Verlustverteilung des Value at Risk keine bestimmte Verteilungsannahme getroffen. Alle Annahmen über die zukünftigen Log>Returns basieren dabei auf der Verlustverteilung der vergangenen Werte und werden dabei nicht über ein statistisches Modell geschätzt. Bei der multivariaten Version der historischen Simulation muss zuerst die Portfolio Rendite der gewichteten Log>Returns berechnet werden:

$$\mathbf{R} = \mathbf{r}\mathbf{w}$$

wobei \mathbf{r} für die historischen Log>Returns der Kapitalanlagen steht und \mathbf{w} für den Gewichtsvektor der Renditen (vgl. J. Danielsson (2011; S.96)).

Danach kann der Value at Risk für den nächsten Tag (VaR_{t+1}) durch das $1 - \alpha$ -Quantil ($Q_{1-\alpha}$) der historischen Verteilung der Portfolio Renditen bestimmt werden:

$$VaR_{t+1} = -Q_{1-\alpha}(\mathbf{R}_t, \mathbf{R}_{t-1}, \dots, \mathbf{R}_1)$$

Unter der Annahme eines Fensters W von 1000 Beobachtungen ist die 5% VaR Schätzung einfach das Negative der fünfzigst kleinsten ($1000 \times 0.05 = 50$) Zahl, welche in der Stichprobe enthalten ist.

Die historische Simulation ist eine einfache Anwendung um Risiken vorherzusagen. Sie ist robust gegenüber Ausreißern und enthält keine Schätzfehler welche in den parametrischen Methoden auftreten können. Ein Nachteil ist, dass die resultierenden VaR Schätzungen erhebliche Sprünge enthalten können. Dies kann auftreten, wenn eine sehr große negative Zahl das Fenster verlässt oder dazukommt.

4.1.2 Parametrische Methoden

Wenn der Value at Risk für ein Portfolio von Kapitalanlagen geschätzt wird, ist eine parametrische Schätzung basierend auf multivariate normalverteilte Renditen ziemlich geeignet, weil die Portfolio Rendite dann univariat normalverteilt ist. Bei einem Portfolio von Kapitalanlagen können die Mittelwerte, Varianzen und Kovarianzen beispielsweise von multivariaten Modellen wie in Kapitel 4 beschrieben, geschätzt werden. Mit den Gewichten des Portfolios $w = (w_1, \dots, w_n)$ kann man anschließend den Erwartungswert des Portfolios μ_P berechnen (vgl.D.Ruppert (2015;S.565)):

$$\mu_P = \mu^\top w$$

Die Standardabweichung des Portfolios ergibt sich aus:

$$\sigma_P = \sqrt{w^\top \Sigma w}$$

Mit diesen Werten ergibt sich der VaR unter der Annahme normalverteilter Renditen des Portfolio:

$$\text{VaR}^{norm}(\alpha) = -\mu_P - \sigma_P \Phi^{-1}(\alpha)$$

Um die heavy tails oder die Steilgipfligkeit der Finanzrenditen adäquat darzustellen, wird manchmal anstatt einer Normalverteilung eine multivariate t-Verteilung für die Renditen angenommen. Basierend auf der Annahme einer multivariaten t-Verteilung ist der VaR definiert als:

$$\text{VaR}^t(\alpha) = -\mu_P - \sqrt{\frac{v}{v-2}} \sigma_P F_v^{-1}(\alpha)$$

$F_v^{-1}(\alpha)$ ist dabei das α -Quantil der t-Verteilung mit Freiheitsgraden v , welche geschätzt werden müssen (vgl.D.Ruppert (2015;S.564)).

Der Vorteil von parametrischen gegenüber nicht parametrischen Methoden wie der historischen Simulation ist die Berücksichtigung der stilisierten Fakten. So kann die Kovarianzmatrix beispielsweise durch eine DCC-Modell prädiziert werden, welches die Zeitabhängigkeit der Matrix modelliert.

4.2 Evaluierungsmethoden

In diesem Kapitel werden drei Tests (vgl. Christoffersen,1998) zur Prädiktionsgüte der VaR Modelle beschrieben. Wenn ein VaR Modell mit Konfidenzniveau $\alpha = 0.05$ gegeben ist, erwartet man, dass im Mittel 5% der realisierten Werte die VaR-Grenze verletzen. Ob eine Verletzung der VaR-Schwelle eingetreten ist wird dabei definiert als:

$$\eta_t = \begin{cases} 1 & r_t \leq -VaR_t \\ 0 & r_t > -VaR_t \end{cases}$$

η ist hier eine Folge von bernoulliverteilten Zufallsvariablen. Die Anzahl der Verstöße wird dabei mit $v_1 = \sum \eta_t$ notiert und die Menge an $\eta_t = 0$ mit $v_0 = W_T - v_1$. Das Zeitfenster in dem die VaR Schätzung stattfindet wird dabei mit W_T notiert (vgl.J.Danielsson (2011;S.145)).

Ob die Anzahl von v_1 dabei eine gewisse Grenze überschreitet, kann dabei durch den unconditional coverage Test von Christoffersen(1998) getestet werden.

4.2.1 Unconditional Coverage Test

Unter der Annahme das η eine Bernoulliverteilte Zufallsvariable ist die Nullhypothese für die Value at Risk Verstöße:

$$H_0 = \eta \sim B(p),$$

wobei B für die Bernoulli-Verteilung steht. Die Bernoulli-Dichte ist definiert als:

$$(1 - p)^{1-\eta_t} (p)^{\eta_t}, \quad \eta_t = 0, 1.$$

Die Wahrscheinlichkeit p kann durch $\hat{p} = \frac{v_1}{W_T}$ geschätzt werden. Die Likelihoodfunktion ist

$$L_U(\hat{p}) = \prod_{t=W_E+1}^T (1 - \hat{p})^{1-\eta_t} (\hat{p})^{\eta_t} = (1 - \hat{p})^{v_0} (\hat{p})^{v_1}.$$

Es ist zu beachten, dass der Datensatz für die Value at Risk Prädiktionen in einen Trainings- und Testdatensatz aufgeteilt wird, wobei der Testdatensatz in der Woche $W_E + 1$ beginnt.

Unter H_0 gilt $p = \hat{p}$, sodass die restringierte Likelihoodfunktion sich zu

$$L_R(p) = \prod_{t=W_E+1}^T (1-p)^{1-\eta_t} (p)^{\eta_t} = (1-p)^{v_0} (p)^{v_1}.$$

ergibt. Danach kann man einen Likelihood Ratio Test durchführen:

$$LR = 2(\log L_U(\hat{p}) - \log L_R(p)) \stackrel{asymptotic}{\sim} \chi_{(1)}^2.$$

Bei einem Signifikanzniveau von 5% wird die Nullhypothese abgelehnt wenn $LR > 3.84$ (vgl. J. Danielsson (2011; S.154)). Bei der Datenanalyse bezieht man sich bezüglich des Ablehnens der Nullhypothese auf den P-Wert.

4.2.2 Unabhängigkeitstest

Der Test zur Unconditional Coverage ignoriert die Zeitvariation in den Daten. Es ist aber auch wichtig zu testen, ob Häufungen von Verstößen auftreten. Es ist also von Interesse zu wissen, ob die Verstöße gleichmäßig verteilt oder aufeinander folgen. Man muss also die Wahrscheinlichkeit berechnen, wenn zwei Verstöße aufeinander folgen (z.B. p_{11}) und die Wahrscheinlichkeit eines Verstößes, falls kein Verstoß am vorherigen Tag aufgetreten ist (p_{01}). Im Allgemeinen ist die Wahrscheinlichkeit definiert als:

$$p_{ji} = P(\eta_t = i | \eta_{t-1} = j),$$

wobei i und j entweder Null oder Eins sind. Nach Christoffersen (1998) ist die Wahrscheinlichkeitsmatrix erster Ordnung definiert als:

$$\Pi_1 = \begin{pmatrix} 1 - p_{01} & p_{01} \\ 1 - p_{11} & p_{11} \end{pmatrix}.$$

Daraus ergibt sich die restringierte Likelihood-Funktion zu

$$L_R(\Pi_1) = (1 - p_{01})^{v_{00}} p_{01}^{v_{01}} (1 - p_{11})^{v_{10}} p_{11}^{v_{11}},$$

wobei v_{ij} die Anzahl der Beobachtungen ist wo j i folgt. Maximum Likelihood Schätzungen werden ermittelt, indem $L_R(\Pi_1)$ maximiert wird:

$$\hat{\Pi}_1 = \begin{pmatrix} \frac{v_{00}}{v_{00}+v_{01}} & \frac{v_{01}}{v_{00}+v_{01}} \\ \frac{v_{10}}{v_{10}+v_{11}} & \frac{v_{11}}{v_{10}+v_{11}} \end{pmatrix}.$$

Unter der Annahme der Nullhypothese, also das ein Verstoß heute keinen Einfluss auf einen Verstoß am nächsten Tag hat, ist $p_{01} = p_{11} = p$. Die Übergangsmatrix ist dann einfach:

$$\hat{\Pi}_0 = \begin{pmatrix} 1 - \hat{p} & \hat{p} \\ 1 - \hat{p} & \hat{p} \end{pmatrix},$$

wobei

$$\hat{p} = \frac{v_{01} + v_{11}}{v_{00} + v_{10} + v_{01} + v_{11}}$$

Die unrestringierte Likelihoodfunktion basierend auf der Nullhypothese verwendet die geschätzte Übergangsmatrix und ergibt sich zu:

$$L_U(\hat{\Pi}_0) = (1 - \hat{p})^{v_{00}+v_{10}} \hat{p}^{v_{01}+v_{11}}.$$

Der Likelihood ratio Test wird schließlich durch die beiden Likelihoods gebildet (vgl. Daniellson (2011;S.156)):

$$LR = 2(\log L_U(\hat{\Pi}_0) - \log L_R(\hat{\Pi}_1)) \stackrel{asymptotic}{\sim} \chi_{(1)}^2.$$

4.2.3 Conditional Test

Nach Christoffersen (1998) kann man aus den beiden eben diskutierten Tests einen gemeinsamen Test bilden, welcher eine Aussage darüber trifft, ob die tatsächlichen Verstöße signifikant von den erwarteten verschieden sind, und ob die Verstöße in Clustern auftreten. Dies kann über folgende Teststatistik evaluiert werden:

$$LR(joint) = LR(coverage) + LR(independence) \sim \chi_2^2.$$

An diesem Punkt könnte man meinen, dass es ausreicht, den Conditional Test durchzuführen. Dieser hat aber eine geringere Power, ein VaR Modell abzulehnen, welches nur eine der zwei Bedingungen erfüllt. Wenn zum Beispiel die tatsächlichen Verstöße nicht signifikant von den erwarteten verschieden sind, aber die Unabhängigkeitsannahme der Verstöße verletzt wird, dann hat der Unabhängigkeitstest eine größere Power das Modell abzulehnen als der zusammengesetzte Test (vgl. J. Danielsson (2011; S.159)).

Bei der Modellanalyse werden für die VaR Modelle alle drei Tests betrachtet.

5 Datenanalyse

5.1 Vorstellung und Aufbereitung der Daten

Der vorliegende Datensatz enthält Tagesdaten der Schlusskurse von 14 ETFs(Exchange Traded Funds) vom 04.01.1999 bis 27.04.2016. ETFs sind Investmentfonds, welche zum Beispiel die Wertentwicklung des DAX passiv nachbilden. Sie werden deswegen auch Indexfonds genannt und bilden dabei ein breites Spektrum von Aktien oder Anleihen ab (vgl.<https://de.scalable.capital/anlageuniversum/> (aufgerufen am 10.08.2016)).

Die 14 im Datensatz enthaltenen ETFs bestehen dabei aus sechs Aktienfonds, drei Staatsanleihen, zwei Unternehmensanleihen, einer besicherten Anleihe, ein Rohstoffindex und ein Immobilienfonds. Eine genauere Beschreibung der einzelnen Daten befindet sich im Appendix.

Für die Analyse wurden aus den täglichen Daten wöchentliche Log>Returns gebildet. Dafür wurde jeweils die logarithmierte Rendite von Freitag auf Freitag berechnet. Falls für einen Freitag keine Daten vorliegen, wurde der Preis des vorigen Tages zur Berechnung der Rendite herangezogen. Durch dieses Vorgehen wurden aus jeweils 4403 täglichen Schlusskursen der ETFs 903 wöchentliche Log>Returns, welche vom 15.01.1999 bis 27.04.2016 datieren, gebildet.

5.2 Deskriptive Analyse

In diesem Kapitel werden die Indexfonds zunächst deskriptiv untersucht. Desweiteren werden die logarithmierten Renditen der verschiedenen Kapitalanlagen grafisch veranschaulicht. Danach werden die Log>Returns mithilfe des Box-Ljung Tests auf Autokorrelationen und Volatilitätscluster untersucht. Die modellbasierte Analyse im nächsten Kapitel wird die Folgerungen der deskriptiven Analyse in der Modellwahl miteinbeziehen. Zuerst zeigen Boxplots die grundlegende Verteilung der Log>Returns. Es zeigt sich, dass die besicherten

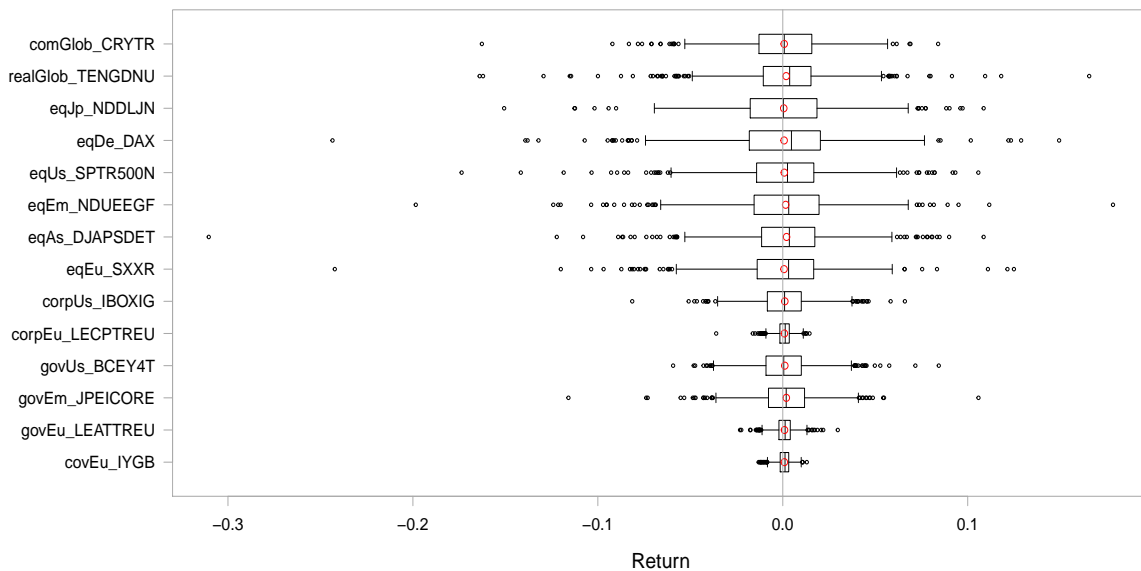


Abbildung 3: Boxplots der Indexfonds

Anleihen, die Staatsanleihen der EU und der Indexfonds der europäischen Unternehmensanleihen im Vergleich zu den restlichen Indexfonds eine geringe Varianz besitzen. Die Mittelwerte, welche durch den roten Kreis gekennzeichnet sind, befinden sich bei den meisten Indexfonds knapp über Null. Generell weisen die Aktienfonds ähnliche statistische Kennzahlen auf wie der Rohstoffindex und der Immobilienindex. Bei den Aktienfonds sieht man ausnahmslos eine große Spannweite der Daten. Der asiatische Aktienindex weist sogar eine Minimum von knapp über $-0,31$ auf, das heißt er hat innerhalb einer Woche um 30% an Wert verloren. Wie aus der grafischen Veranschaulichung der Daten im Appendix zu entnehmen ist, ergibt sich dieser Ausreißer aus der Finanzkrise 2008. Die Boxplots lassen durch die großen Ausreißer den Anschein erwecken, dass sich das Investieren insbesondere in Aktienfonds nicht sehr rentabel ist. Dies stimmt jedoch nicht, weil die Mittelwerte der

Aktien positiv sind, was auf eine langfristig positive Rendite hindeutet.

In folgender Tabelle werden die wichtigsten statistischen Kennzahlen beschrieben:

Tabelle 1: Tabelle von statistischen Kennzahlen der ETFs

ETFs	Mittelwert	Varianz	Schiefe	Kurtosis	$Q(7)$	$Q_{qu}(7)$
covEu_IYGB	0.00084	0.00001	-0.41766	0.72677	0.18985	0.00000
govEu_LEATTREU	0.00092	0.00003	-0.11742	2.20284	0.00717	0.00000
govEm_JPEICORE	0.00190	0.00030	-0.30674	4.51839	0.00033	0.00000
govUs_BCEY4T	0.00110	0.00025	0.33065	1.61366	0.60243	0.00000
corpEu_LECPTREU	0.00087	0.00002	-1.02100	5.88827	0.00135	0.00000
corpUs_IBOXIG	0.00110	0.00024	0.01902	1.58659	0.24024	0.00000
eqEu_SXXR	0.00067	0.00073	-1.04202	8.83311	0.00003	0.00000
eqAs_DJAPSDET	0.00205	0.00073	-1.89502	20.66563	0.00858	0.00000
eqEm_NDUEEGF	0.00155	0.00095	-0.47919	3.86428	0.01095	0.00000
eqUs_SPTR500N	0.00082	0.00077	-0.55105	3.34918	0.00002	0.00000
eqDe_DAX	0.00072	0.00111	-0.64819	4.61954	0.00610	0.00000
eqJp_NDDLJN	0.00053	0.00083	-0.18303	1.58373	0.67251	0.00491
realGlob_TENGDNU	0.00184	0.00070	-0.63186	6.84594	0.05337	0.00000
comGlob_CRYTR	0.00076	0.00055	-0.57830	3.05199	0.67276	0.00000

Die Schiefe wird dabei geschätzt durch:

$$\hat{S}(r) = \frac{1}{(T-1)\hat{\sigma}^3} \sum_{t=1}^T (r_t - \hat{\mu})^3.$$

r_t ist dabei die logarithmierte Rendite zum Zeitpunkt t , $\hat{\mu}$ ist der geschätzte Mittelwert der Datenreihe und $\hat{\sigma}$ ist die geschätzte Standardabweichung (vgl.R.S.Tsay (2013;S.23)). Die Schiefe weist generell auf die Asymmetrie der Daten hin. Die meisten ETFs sind in diesem Fall linksschief verteilt. Nur die US-Unternehmensanleihen und die US-Staatsanleihen haben positive Werte und haben deswegen eine rechtsschiefe Verteilung.

Die Kurtosis ist definiert als:

$$\frac{1}{(T-1)\hat{\sigma}^4} \sum_{t=1}^T (r_t - \hat{\mu})^4$$

Die Kurtosis ist ein Maß für die Spitzigkeit der Verteilung. Die Kurtosis der Normalverteilung beträgt drei (vgl.R.S.Tsay (2013;S.22-23)). Wie man anhand der Tabelle erkennen kann sind manche ETFs, wie zum Beispiel der japanische Aktienfonds, flachgipfliger als die Normalverteilung. Andere wie der asiatische Aktienfonds sind sehr steilgipflig. Die Kurtosis gibt auch Auskunft darüber ob die Verteilungsannahmen adäquat sind. Wenn eine Kapi-

talanlage eine leptokurtisch ist, also eine größere Kurtosis als Drei besitzt, ist das ein Indiz für schwere Ränder. Da die Kurtosis von manchen ETFs größer als drei ist, wird zusätzlich zur Normalverteilungsannahme die T-Verteilung bei der Modellanalyse verwendet, welche schwere Ränder adäquater als die Normalverteilung abbildet.

Die letzten beiden Spalten zeigen die P-Werte der Box-Ljung Tests. Bei der vorletzten Spalte wurde der Box-Ljung Test auf die Log>Returns des jeweiligen Indexfonds angewendet. Man sieht, dass einige ETFs, wie zum Beispiel der japanische Aktienfonds, deutliche serielle Korrelationen aufweisen. Für die Freiheitsgrade wurde die Zahl 7 verwendet, weil $7 \approx \ln(903)$ (vgl. R.S. Tsay (2014;S.49))

Die letzte Spalte enthält die P-Werte der quadrierten Log>Returns der jeweiligen ETFs. Anhand der P-Werte wird deutlich, dass Volatilitätscluster in allen ETFs existieren und somit ein GARCH Modell an die Daten angepasst werden kann.

5.3 Value at Risk basierende Modellanalyse

In diesem Abschnitt werden die in Kapitel 3 besprochenen Modelle auf die Daten angewendet. Dafür wird ein sogenanntes Backtesting-Verfahren durchgeführt. Backtesting ist nützlich, um die Schwächen einer Risikoprognose zu identifizieren und das Unterschätzen von VaR Prognosen aufzuzeigen (vgl. J. Danielsson (2011; S.143)).

Es wird also geprüft, ob die in Kapitel 3 genannten Modelle geeignet für die Prädiktion des Value at Risk sind. Dafür werden $\mathbf{r}_1, \dots, \mathbf{r}_{203}$ zur Parameterschätzung für beispielsweise die DCC-Modelle verwendet. $\mathbf{r}_1, \dots, \mathbf{r}_{203}$ beinhalten dabei die wöchentlichen Log>Returns der jeweiligen ETFs im Zeitraum vom 15.01.1999 bis zum 29.11.2002. Auf die restlichen Beobachtungen wird die Value at Risk Prädiktion durchgeführt. Dabei ist zu beachten, dass in dieser Arbeit ein rekursives Verfahren verwendet wird. Das heißt, dass sich die Anzahl der Beobachtungen auf jene die VaR Prädiktion basiert, im Verlauf des Verfahrens erhöht. Basiert $\text{VaR}_{204}(\alpha)$ noch auf $\mathbf{r}_1, \dots, \mathbf{r}_{203}$, so gründet sich die Prädiktion von VaR_{205} auf die wöchentlichen Log>Returns $\mathbf{r}_1, \dots, \mathbf{r}_{204}$ der ETFs. Die folgende Tabelle veranschaulicht dieses Vorgehen:

Tabelle 2: Rekursives Verfahren

Start	Ende	VaR
t=1	t=203	$\text{VaR}_{204}(\alpha)$
t=1	t=204	$\text{VaR}_{205}(\alpha)$
\vdots	\vdots	\vdots
t=1	t=901	$\text{VaR}_{902}(\alpha)$
t=1	t=902	$\text{VaR}_{903}(\alpha)$

Das Modell wird also nach jeder Prädiktion wieder an die Daten angepasst und die Parameter neu geschätzt. Mit dem gerade gefitteten Modell wird dann die VaR Prädiktion durchgeführt. Dies hat den Vorteil, dass jede Value at Risk Prädiktion auf die aktuellen Daten basiert und durch die Parameterschätzung mögliche Änderungen in den Daten aufgegriffen wird. So ergeben sich im Zeitraum vom 06.12.2002 bis 27.04.2016 für jedes verwendete Modell 700 wöchentliche VaR Prädiktionen. Dadurch, dass manche ETFs bei den Box-Ljung Testergebnissen (siehe Tabelle 1) serielle Korrelationen aufgewiesen haben, wurden für die parametrisch VaR Modelle für jedes ETF ein AR(1) Modell an die Daten angepasst. Mit diesen wurde dann auch der jeweilige Mittelwert für die VaR Prädiktion vorhergesagt. Da beim DCC Modell und seinen Abwandlungen für die Diagonale der Kovarianzmatrix GARCH Modelle benötigt werden, wurden hier für das DCC Modell ein GARCH(1,1) Modell gefittet. Beim ADCC Modell wurde hierfür ein GJR-GARCH Modell

verwendet. Für das DCC und das ADCC wurde jeweils eine multivariate Normalverteilung und eine multivariate t-Verteilung angenommen. Für die VaR Prädiktion wurde ein **gleichgewichtetes** Portfolio betrachtet.

Bevor die Ergebnisse der verschiedenen Evaluierungsmethoden aufgezeigt werden, wird noch kurz beispielhaft auf die geschätzten Parameter eines DCC Modells, basierend auf r_1, \dots, r_{203} , mit multivariat normalverteilten Innovationen eingegangen. Außerdem wird noch das AIC für die verschiedenen Modelle, wenn jene über alle Log>Returns r_1, \dots, r_{903} gefittet wurden, kurz betrachtet.

```
*-----*
*          DCC GARCH Fit          *
*-----*

Distribution      :  mvnorm
Model            :  DCC(1,1)
No. Parameters   :  163
[VAR GARCH DCC UncQ] : [0+70+2+91]
No. Series       :  14
No. Obs.         :  203
Log-Likelihood   :  9105.962
Av.Log-Likelihood :  44.86
```

Optimal Parameters

```
-----
              Estimate  Std. Error  t value  Pr(>|t|)
[covEu_IYGB].mu      0.001010   0.000279   3.618946 0.000296
[covEu_IYGB].ar1     -0.001372   0.071424  -0.019207 0.984676
[covEu_IYGB].omega    0.000002   0.000001   1.293826 0.195725
[covEu_IYGB].alpha1  0.090687   0.016633   5.452347 0.000000
[covEu_IYGB].beta1   0.820783   0.032296  25.414137 0.000000
[govEu_LEATTREU].mu  0.000974   0.000312   3.123450 0.001787
[govEu_LEATTREU].ar1 -0.019276   0.069025  -0.279254 0.780050
[govEu_LEATTREU].omega 0.000002   0.000002   0.904287 0.365843
[govEu_LEATTREU].alpha1 0.093936   0.019112   4.915136 0.000001
[govEu_LEATTREU].beta1 0.828400   0.031302  26.465072 0.000000
[govEm_JPEICORE].mu  0.003115   0.001457   2.137310 0.032573
```

[govEm_JPEICORE].ar1	-0.083898	0.065573	-1.279459	0.200735
[govEm_JPEICORE].omega	0.000000	0.000001	0.062538	0.950135
[govEm_JPEICORE].alpha1	0.000000	0.000127	0.000005	0.999996
[govEm_JPEICORE].beta1	0.999000	0.000134	7428.301751	0.000000
[govUs_BCEY4T].mu	0.002012	0.001123	1.791086	0.073279
[govUs_BCEY4T].ar1	-0.033833	0.064605	-0.523694	0.600491
[govUs_BCEY4T].omega	0.000000	0.000001	0.119605	0.904796
[govUs_BCEY4T].alpha1	0.000000	0.000395	0.000004	0.999997
[govUs_BCEY4T].beta1	0.999000	0.000407	2456.755739	0.000000
[corpEu_LECPTREU].mu	0.000933	0.000279	3.339188	0.000840
[corpEu_LECPTREU].ar1	-0.006361	0.070569	-0.090134	0.928181
[corpEu_LECPTREU].omega	0.000002	0.000001	1.511860	0.130569
[corpEu_LECPTREU].alpha1	0.094155	0.014207	6.627333	0.000000
[corpEu_LECPTREU].beta1	0.815952	0.030423	26.820192	0.000000
[corpUs_IBOXIG].mu	0.001681	0.001135	1.481051	0.138593
[corpUs_IBOXIG].ar1	-0.039734	0.066960	-0.593397	0.552915
[corpUs_IBOXIG].omega	0.000000	0.000001	0.142280	0.886859
[corpUs_IBOXIG].alpha1	0.000000	0.000229	0.000040	0.999968
[corpUs_IBOXIG].beta1	0.999000	0.000218	4577.622519	0.000000
[eqEu_SXXR].mu	-0.000485	0.001713	-0.282910	0.777246
[eqEu_SXXR].ar1	-0.044014	0.076376	-0.576280	0.564426
[eqEu_SXXR].omega	0.000180	0.000092	1.964067	0.049522
[eqEu_SXXR].alpha1	0.198871	0.076539	2.598305	0.009369
[eqEu_SXXR].beta1	0.558751	0.162213	3.444547	0.000572
[eqAs_DJAPSDDET].mu	0.003335	0.001752	1.904148	0.056891
[eqAs_DJAPSDDET].ar1	0.152519	0.079913	1.908549	0.056320
[eqAs_DJAPSDDET].omega	0.000419	0.000140	2.989707	0.002792
[eqAs_DJAPSDDET].alpha1	0.264000	0.136602	1.932625	0.053282
[eqAs_DJAPSDDET].beta1	0.000000	0.252335	0.000000	1.000000
[eqEm_NDUEEGF].mu	0.000898	0.002589	0.346698	0.728818
[eqEm_NDUEEGF].ar1	0.051677	0.069739	0.741002	0.458692
[eqEm_NDUEEGF].omega	0.000001	0.000007	0.173165	0.862521
[eqEm_NDUEEGF].alpha1	0.000000	0.000739	0.000001	1.000000
[eqEm_NDUEEGF].beta1	0.999000	0.002586	386.322011	0.000000

[eqUs_SPTR500N].mu	-0.000508	0.002092	-0.242895	0.808087
[eqUs_SPTR500N].ar1	-0.114918	0.084388	-1.361791	0.173264
[eqUs_SPTR500N].omega	0.000002	0.000002	0.745063	0.456233
[eqUs_SPTR500N].alpha1	0.000000	0.002496	0.000007	0.999995
[eqUs_SPTR500N].beta1	0.999000	0.000914	1092.632368	0.000000
[eqDe_DAX].mu	-0.001232	0.002479	-0.497008	0.619184
[eqDe_DAX].ar1	0.030890	0.072116	0.428338	0.668405
[eqDe_DAX].omega	0.000324	0.000115	2.824178	0.004740
[eqDe_DAX].alpha1	0.187752	0.076099	2.467200	0.013617
[eqDe_DAX].beta1	0.568665	0.118905	4.782511	0.000002
[eqJp_NDDLJN].mu	-0.000692	0.004039	-0.171438	0.863880
[eqJp_NDDLJN].ar1	-0.056024	0.073039	-0.767038	0.443059
[eqJp_NDDLJN].omega	0.000001	0.000006	0.134358	0.893119
[eqJp_NDDLJN].alpha1	0.000000	0.002797	0.000007	0.999994
[eqJp_NDDLJN].beta1	0.998999	0.004134	241.627503	0.000000
[realGlob_TENGDNU].mu	0.001805	0.001478	1.221002	0.222085
[realGlob_TENGDNU].ar1	0.051962	0.106555	0.487657	0.625793
[realGlob_TENGDNU].omega	0.000001	0.000002	0.361149	0.717988
[realGlob_TENGDNU].alpha1	0.000000	0.000515	0.000005	0.999996
[realGlob_TENGDNU].beta1	0.999000	0.001286	776.876202	0.000000
[comGlob_CRYTR].mu	0.003122	0.004200	0.743400	0.457239
[comGlob_CRYTR].ar1	0.043186	0.070920	0.608946	0.542560
[comGlob_CRYTR].omega	0.000000	0.000002	0.143399	0.885975
[comGlob_CRYTR].alpha1	0.000000	0.009491	0.000000	1.000000
[comGlob_CRYTR].beta1	0.999000	0.010083	99.077874	0.000000
[Joint]dcca1	0.009158	0.003411	2.684962	0.007254
[Joint]dccb1	0.900573	0.024011	37.506928	0.000000

Information Criteria

Akaike -88.108

Wie man deutlich erkennt, hat dieses DCC Modell durch die hohe Anzahl an Kapitalanla-

gen sehr viele Parameter. Die meisten AR Parameter haben dabei hohe P-Werte genau wie bei einigen GARCH Parameterschätzungen. Die Mittelwertsgleichungen weisen also hauptsächlich nicht signifikante Parameter auf. Die DCC Parameter verdeutlichen mit $b = 0.90$ eine starke Persistenz der Korrelationsmatrix. Auf die Residuen der Autoregressiven Modelle kann man, wie in Ende Kapitel 2 beschrieben, einen multivariaten Box-Ljung Test anwenden.

```
multivariate Box-Ljung test
data: residuals(fit_DCCmnorm)^2
Q(7) = 7435.37, df = 1372, p-value = 0
```

Das verdeutlicht, dass bedingte Heteroskedastizität in den Daten vorliegt.

Im folgenden wird in einer Tabelle kurz das AIC der vier angepassten Modelle veranschaulicht

Tabelle 3: AIC für die Schätzung über den gesamten Datensatz

Modell	AIC
DCC_GARCH_mnorm	-90.864
DCC_GARCH_T	-92.068
ADCC_GJR-GARCH_mnorm	-90.807
ADCC_GJR-GARCH_T	-91.969

Das DCC_GARCH_T Modell ist nach dem AIC Kriterium das an die Daten am besten angepasste Modell. Das Akaike Informationskriterium ist dabei ein Maß für die Güte eines Modells. Es bildet sich aus der Log-Likelihood der Daten und den Anzahl der Parameter. Dabei wird das AIC so modelliert, dass eine hohe Loglikelihood zu einem niedrigen AIC und eine große Anzahl an Parametern zu einem hohen AIC führt (vgl. D. Ruppert (2015; S.109)). Anhand Tabelle 3 kann man auch erkennen, dass sich das AIC der jeweiligen Modelle bezüglich des Werts auf einen ähnlich niedrigen Niveau befinden.

Im folgenden werden die verschiedenen Tests der Modellevaluation in einer Tabelle veranschaulicht (Beschreibung der einzelnen Modelle befindet sich im Appendix). Das Signifikanzniveau wird dabei auf 5% festgelegt.

Tabelle 4: Modellevaluation für $\text{VaR}(\alpha = 0.05)$

Modell	Verstöße	%Verstöße	LR_{uc}	LR_{ind}	LR_{cc}
HS	39	5.57	0.50	0.00	0.00
EWMA	41	5.86	0.31	0.31	0.36
DCC_GARCH_mnorm	42	6.00	0.24	0.36	0.33
DCC_GARCH_T	46	6.57	0.07	0.26	0.10
ADCC_GJR-GARCH_mnorm	35	5.00	1.00	0.36	0.66
ADCC_GJR-GARCH_T	42	6.00	0.24	0.36	0.33

Erwartete Verstöße: $700 * 0.05 = 35$

Anhand Tabelle 4 erkennt man, dass die historische Simulation bezüglich der Anzahl der auftretenden Verstöße im Vergleich zu den anderen Modellen relativ gut abschneidet. Wie in Kapitel 4 beschrieben, ist es allerdings eine unbedingte Methode zur Bestimmung der Value at Risk Schätzung und berücksichtigt Eigenschaften wie das Volatilitätsclustering nicht. Bei einem Signifikanzniveau von 5% wird auch die Nullhypothese vom Unabhängigkeitstests und vom Conditional Coverage Test abgelehnt. Die Nullhypothese, dass die Anzahl der Verstöße signifikant von den erwarteten Verstößen verschieden ist, kann nicht abgelehnt werden.

Das EWMA Modell hat eine größere Anzahl an Verstößen als die Historische Simulation. Dafür kann die Nullhypothese bezüglich der Unabhängigkeit der Verstöße nicht abgelehnt werden. Das DCC Modell, dessen Innovationen multivariat normalverteilt sind, ist mit 42 Verstößen näher am erwarteten Wert von 35 dran als das DCC Modell mit den multivariat t-verteilten Innovationen. Beim diesem würde die Nullhypothese des Unconditional Coverage Test bei einem Signifikanzniveau von 10% abgelehnt werden.

Die Modelle, welche den Leverage Effekt der Finanzzeitreihen für die Volatilitäten und Korrelationen berücksichtigen, schneiden gegenüber den gewöhnlichen DCC Modellen besser ab. Das ADCC Modell, dessen Innovationen \mathbf{a}_t multivariat normalverteilt sind hat genau die erwartete Anzahl an Verstöße. Auch hier hat der Pendant mit der multivariaten t-Verteilung eine höhere Anzahl an Verstößen. Die Annahme einer multivariaten Normalverteilung der Innovationen \mathbf{a}_t führt im Allgemeinen zu besseren Evaluierungsergebnissen.

Abbildung 4 veranschaulicht die VaR Prädiktion über die Zeit (Das Vorzeichen des Value at Risk wurde zur besseren Veranschaulichung gegenüber der Definition in Kapitel 4 geändert).

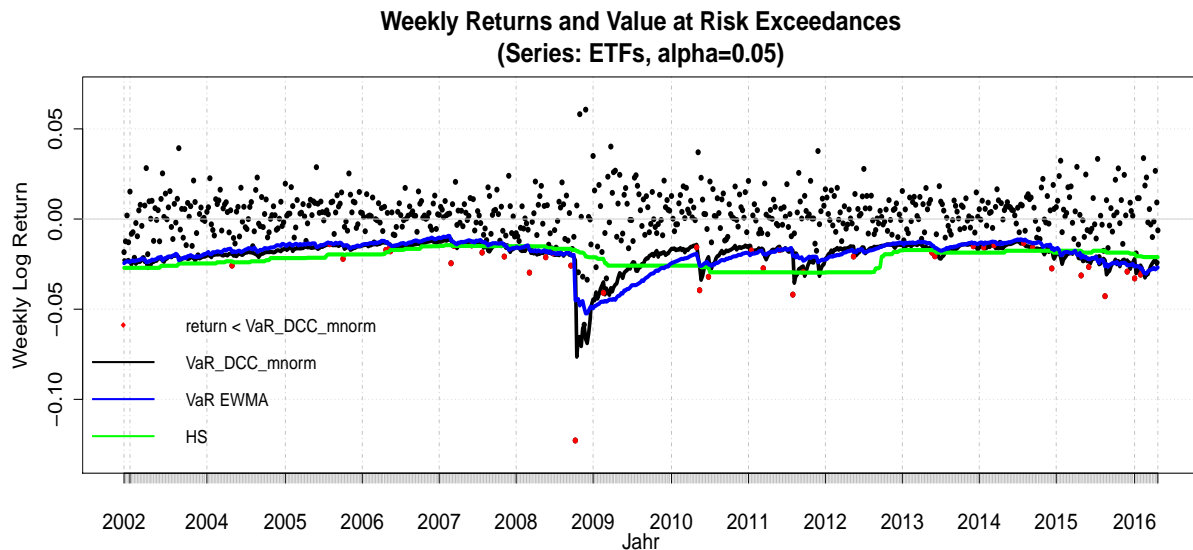


Abbildung 4: DCC Modell multivariat normalverteilt

In dieser Abbildung stehen die schwarzen Punkte für die realisierten Werte der Portfoliorendite und die roten Punkte stellen Überschreitungen der VaR Werte des DCC Modells dar. Diese Grafik veranschaulicht generell die unterschiedliche Performance der Modelle. Wie man anhand der Ergebnisse aus Tabelle 3 erschließen kann, ist die Abhängigkeit der Verstöße bei der historischen Simulation ersichtlich. Diese reagiert nicht so heftig auf die Finanzkrise 2008, wie beispielsweise das DCC Modell. Insgesamt scheint die historische Simulation robust gegenüber Ausreißer zu sein. Das EWMA Modell reagiert stärker auf die Volatilitätssprünge wie die historische Simulation, aber weniger stark als das DCC Modell. Wie in Kapitel 3 erwähnt, wird für $\lambda = 0.96$ angenommen., wohingegen sich die Persistenzparameter der DCC und ADCC Modelle für jede Woche neu berechnet werden. In Abbildung 5 wird der Verlauf der VaR Schätzung für das ADCC Modell gezeigt. Im Vergleich zum normalen DCC Modell reagiert es auf die Finanzkrise 2008 wesentlich heftiger und hat generell, wie in Tabelle 3 deutlich wurde, bessere Evaluierungsergebnisse bezüglich der Value at Risk Schätzung. Die Verläufe der zwei anderen VaR Modelle befinden sich im Appendix.

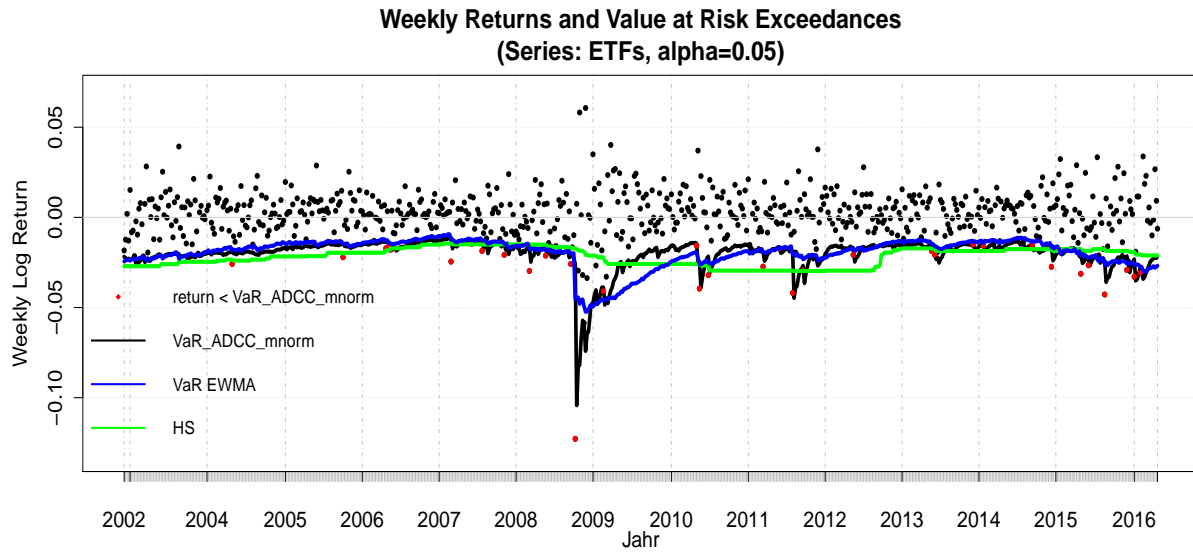


Abbildung 5: ADCC Modell multivariat normalverteilt

Um den Verlauf der Value at Risk Prädiktion nachzuvollziehen, kann es nützlich sein die geschätzten Parameter zu betrachten. Zur besseren Veranschaulichung wird die Gleichung von Q_t für das DCC und ADCC Modell nochmal dargestellt (siehe Kapitel 3):

$$Q_t = (1 - a - b)\bar{Q} + a\varepsilon_{t-1}\varepsilon_{t-1}^T + bQ_{t-1}$$

$$Q_t = (\bar{Q} - a\bar{Q} - b\bar{Q} - g\bar{N}) + a\varepsilon_{t-1}\varepsilon_{t-1}^T + gn_{t-1}n_{t-1}^T + bQ_{t-1}$$

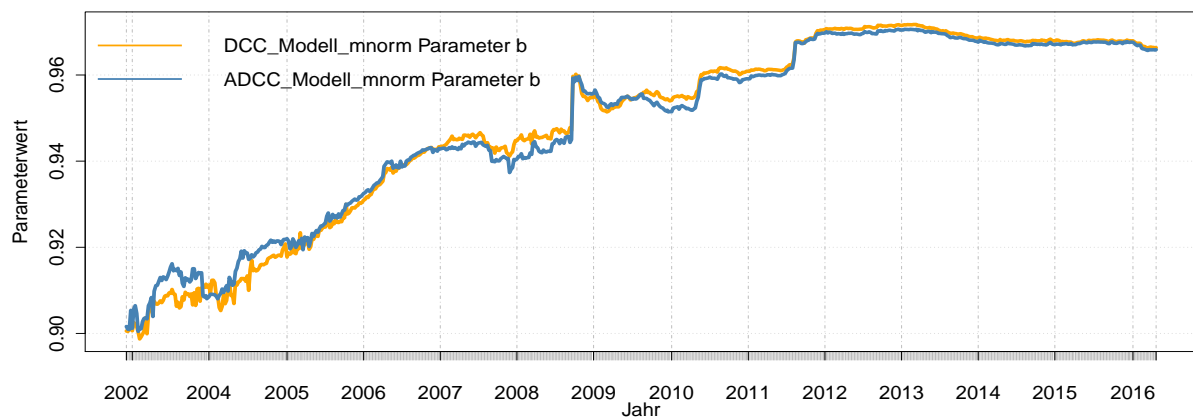


Abbildung 6: Parameter b

In Abbildung 6 erkennt man, dass sich der Persistenzparameter b im Laufe der Zeit erhöht und in der Finanzkrise 2008 es eine sprunghafte Erhöhung gibt. Das bedeutet, dass die vergangenen Q_t im Verlauf der Zeit höher gewichtet werden. Die Parameter der beiden Modelle haben insgesamt einen ähnlichen Verlauf.

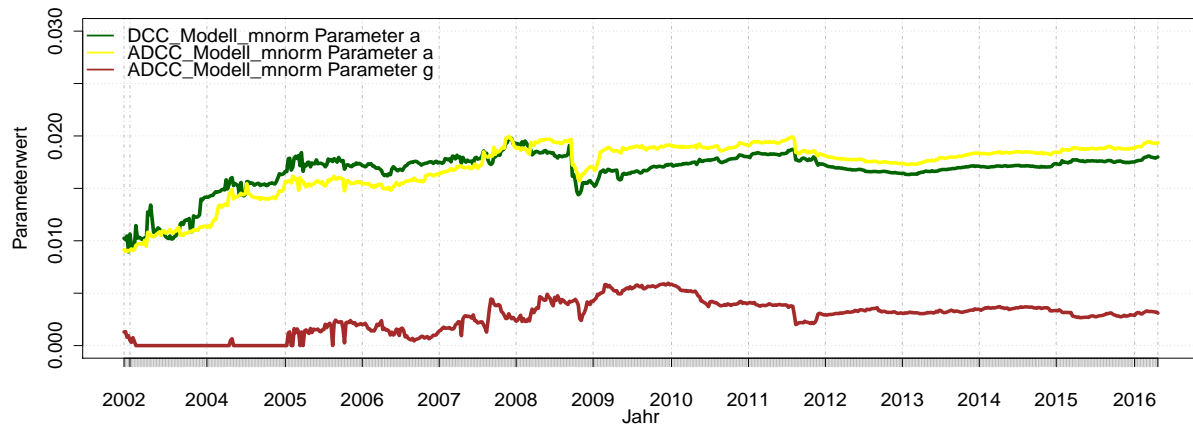


Abbildung 7: restliche Parameter

In Abbildung 7 wird die Gewichtung der vergangenen Störterme $a\epsilon_{t-1}\epsilon_{t-1}^T$ veranschaulicht. Die Parameterwerte für a steigen im Verlauf der Zeit und fallen dann kurzzeitig im Zeitraum der Finanzkrise 2008. Parameter g , welcher für die Asymmetrie der Daten zuständig ist, verharrt zu Beginn der Schätzung nahe an Null und steigt dann bis 2009 an und pendelt sich dann ungefähr auf den Wert 0.004 ein.

6 Kritik und Fazit

Wie sich in Kapitel 5 gezeigt hat, war die multivariate Normalverteilung bei der Modellevaluation jeweils die bessere Modellwahl für das Schätzen des Value at Risk, obwohl nach der deskriptiven Analyse die Mehrheit der ETFs leptokurtisch ist. Die historische Simulation hat von den Modellen die zweitniedrigsten Verstöße, ist aber kein unabhängiges Verfahren und reagierte kaum auf die Finanzkrise hinsichtlich der Value at Risk Prädiktion. Das EWMA Modell reagiert auf die Krise besser als die historische Simulation, aber schlechter als die übrigen DCC und ADCC Modelle. Wenn man den Leverage Effekt der Volatilitäten und der Varianzen in der Modellierung beachtet, also ein GJR-GARCH Modell für die Varianzen und ADCC Modell für die Korrelationen verwendet, erzielt man insgesamt die besten Ergebnisse bezüglich der VaR Prädiktion. Nach dem AIC Kriterium ist das DCC_GARCH_T Modell am besten an die Daten angepasst.

Generell ist die Performance von multivariaten Modellen besser, bei einer geringen Anzahl an Kapitalanlagen. Das liegt zu einem daran, dass man bei einer Anzahl von d Kapitalanlagen eine Kovarianzmatrix mit $k(k-1)/2$ verschiedenen Elementen schätzen muss und zum anderen, dass die Kovarianzmatrix Σ_t für alle t positiv definit sein muss (vgl. R.S. Tsay (2014;S.399)) Dies führt dazu, dass die Modelle bei der Schätzung teilweise nicht konvergieren und somit das Modellieren unmöglich ist.

Generell ist ein gleichgewichtetes Portfolio bezüglich dem Risikomanagement problematisch. Wie die Log>Returns der ETFs zeigen, wäre es optimal wenn man im Krisenfall eher einen größeren Anteil an beispielsweise besicherten Anleihen hält. So kann ein dynamisches Portfolio sicherlich zur Risikominimierung beitragen.

Literatur

- R.S. Tsay *An Introduction to Analysis of Financial Data with R* . Wiley Series in Probability and Statistics. Wiley, 2013
- D. Ruppert and D.S. Matteson. *Statistics and Data Analysis for Financial Engineering: with R examples*. Springer Texts in Statistics. Springer New York, 2015.
- J. Danielsson *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*. The Wiley Finance Series. Wiley, 2011
- R.S. Tsay *Multivariate Time Series Analysis : with R and financial applications*. Wiley Series in Probability and Statistics. Wiley, 2014
- J.P. Kreiss and G. Neuhaus *Einführung in die Zeitreihenanalyse Statistik und ihre Anwendungen*. Springer Berlin Heidelberg, 2010.
- L. Capiello, R.F. Engle, and K. Sheppard. Asymmetric Dynamics in the Correlation of Global Equity and Bond Returns. *Journal of Financial Econometrics* 4:4, 537-572, 2006
- R.F. Engle. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroscedasticity Models. *Journal of Business & Economic Statistics* ,20:3, 339-350, 2002
- R.F. Engle, and K. Sheppard. Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. Discussion Paper 2001-15, University of California, San Diego, 2001
- C. Peitz *Die parametrische und semiparametrische Analyse von Finanzzeitreihen: Neue Methoden, Modelle und Anwendungsmöglichkeiten* Springer Fachmedien Wiesbaden, 2016
- P. Christoffersen. Evaluating Intervals Forecasts, *International Economic Review*, 39, 841-862, 1998
- Scalable Capital Vermögensverwaltung GmbH, <https://de.scalable.capital/anlageuniversum/>
Zuletzt aufgerufen am 10.08.2016

- A. Ghalanos (2015) Package 'rmgarch'
URL: <https://cran.r-project.org/web/packages/rmgarch/rmgarch.pdf>
- A. Ghalanos (2015) Package 'rugarch'
URL: <https://cran.r-project.org/web/packages/rugarch/rugarch.pdf>
- B. Peterson ,and P. Carl (2014) Package 'PerformanceAnalytics'
URL: <https://cran.r-project.org/web/packages/PerformanceAnalytics/PerformanceAnalytics.pdf>
- J. Ryan, and J. Ulrich (2013) Package 'xts'
URL: <https://cran.r-project.org/web/packages/xts/xts.pdf>
- A. Dragulescu (2014) Package 'xlsx'
URL: <https://cran.r-project.org/web/packages/xlsx/xlsx.pdf>

7 Appendix

7.1 Datenbeschreibung

Aktienfonds:

- eqEu_SXXR: Aktien von 600 größten Unternehmen aus Europa
- eqAs_DJAPSDT: Aktien aus dem asiatischen pazifischen Raum
- eqEm_NDUEEGF: Aktien von Unternehmen aus Schwellenländern
- eqUs_SPTR500N: Aktien von 500 führenden US-Unternehmen
- eqDe_DAX: Aktien der dreißig größten gehandelten Unternehmen in Deutschland
- eqJp_NDDLJN: Aktien von japanischen Unternehmen

Immobilien:

- realGlob_TENGDNU: Börsennotierte Immobiliengesellschaften aus entwickelten Märkten

Rohstoffe:

- comGlob_CRYTR: Rohstoffindex

Besicherte Anleihen:

- covEu_IYGB: Besicherte Anleihen aus der Eurozone

Staatsanleihen:

- govEu_LEATTREU: Staatsanleihen der Eurozone
- govEm_JPEICORE: Staatsanleihen von US-Dollar dominierten Schwellenländer
- govUs_BCEY4T: US-Staatsanleihen

Unternehmensanleihen:

- corpEu_LECPTREU: Auf Euro laufende Unternehmensanleihen
- corpUs_IBOXIG: Auf US-Dollar laufende Unternehmensanleihen

7.2 Deskriptive Analyse

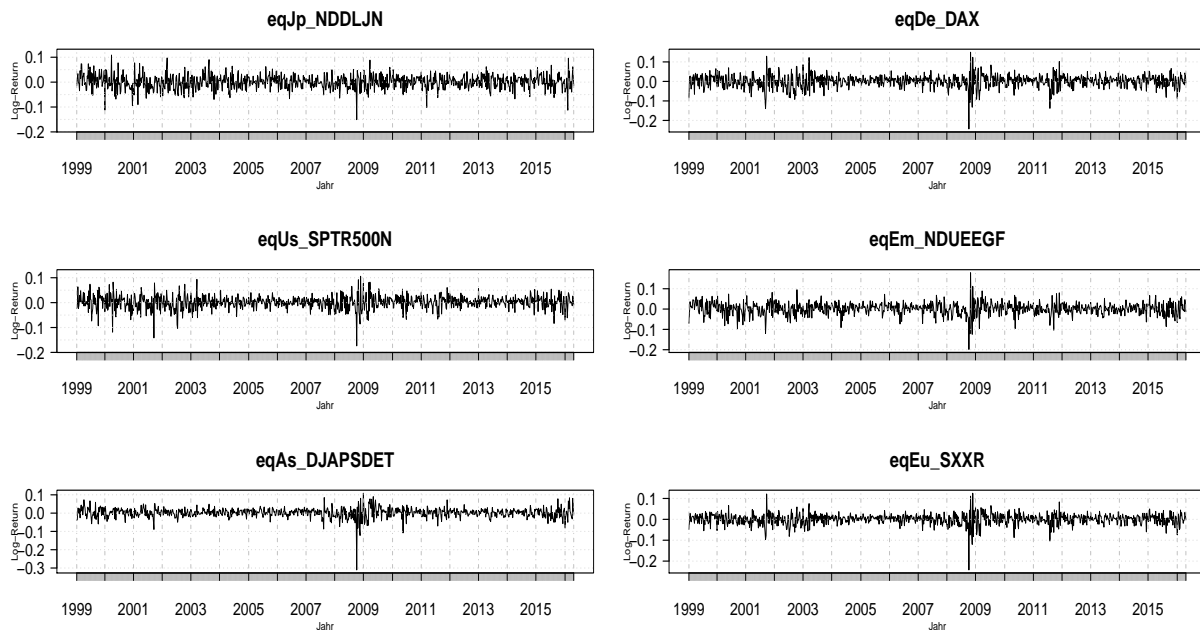


Abbildung 8: Wöchentliche Log>Returns der Aktienfonds von 15.01.1999 bis 27.04.2016

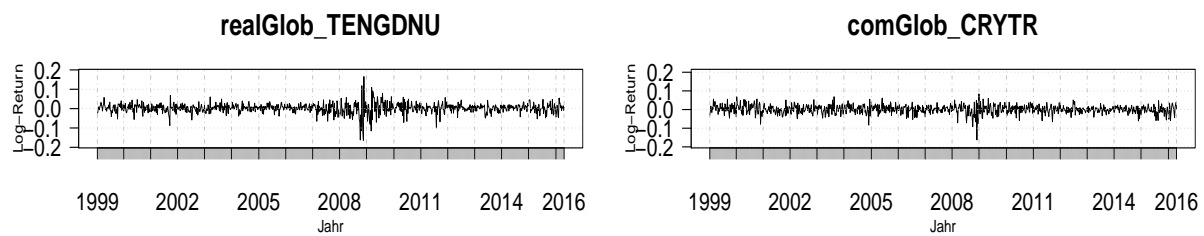


Abbildung 9: Wöchentliche Log>Returns des Immobilienfonds und Rohstofffonds von 15.01.1999 bis 27.04.2016

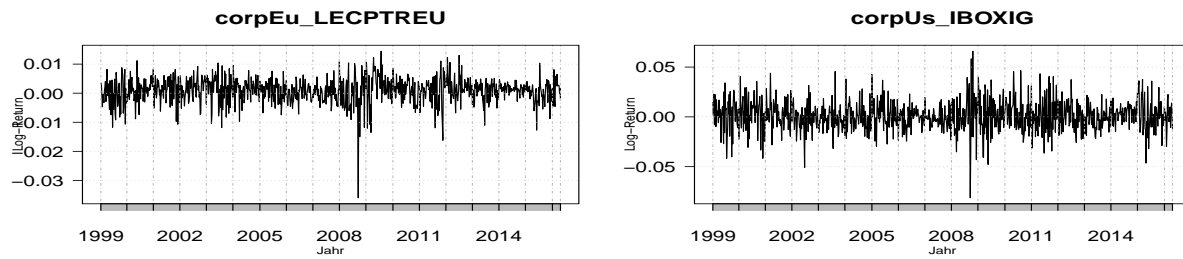


Abbildung 10: Log>Returns der Unternehmensanleihen von 15.01.1999 bis 27.04.2016

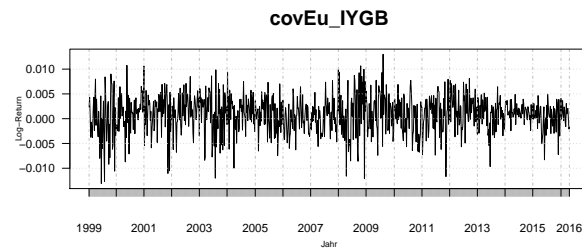


Abbildung 11: Log>Returns der Besicherten Anleihe von 15.01.1999 bis 27.04.2016

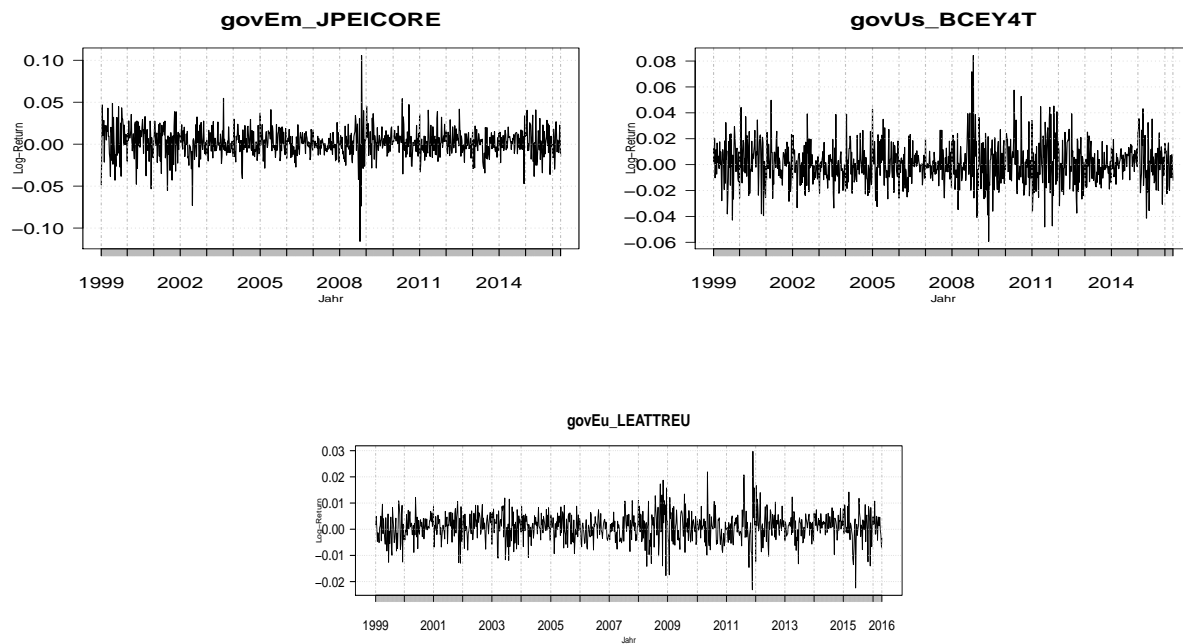


Abbildung 12: Log>Returns der Staatsanleihen von 15.01.1999 bis 27.04.2016

7.3 Modellbasierte Analyse

7.3.1 Notation

Für alle Parametrischen Verfahren wurde als Mittelwertsgleichung ein AR(1) Modell gewählt

- HS: Historische Simulation mit einem Fenster von 200 Beobachtungen
- EWMA : EWMA Modell mit Normalverteilungsannahme
- DCC_mnorm : GARCH(1,1) normalverteilt und DCC(1,1) multivariat normalverteilt
- DCC_T : GARCH(1,1) t-verteilt und DCC(1,1) multivariat t-verteilt
- ADCC_mnorm : GJR-GARCH(1,1) normalverteilt und ADCC(1,1) multivariat normalverteilt
- ADCC_T : GJR-GARCH(1,1) t-verteilt und ADCC(1,1) multivariat t-verteilt

7.3.2 VaR Grafiken

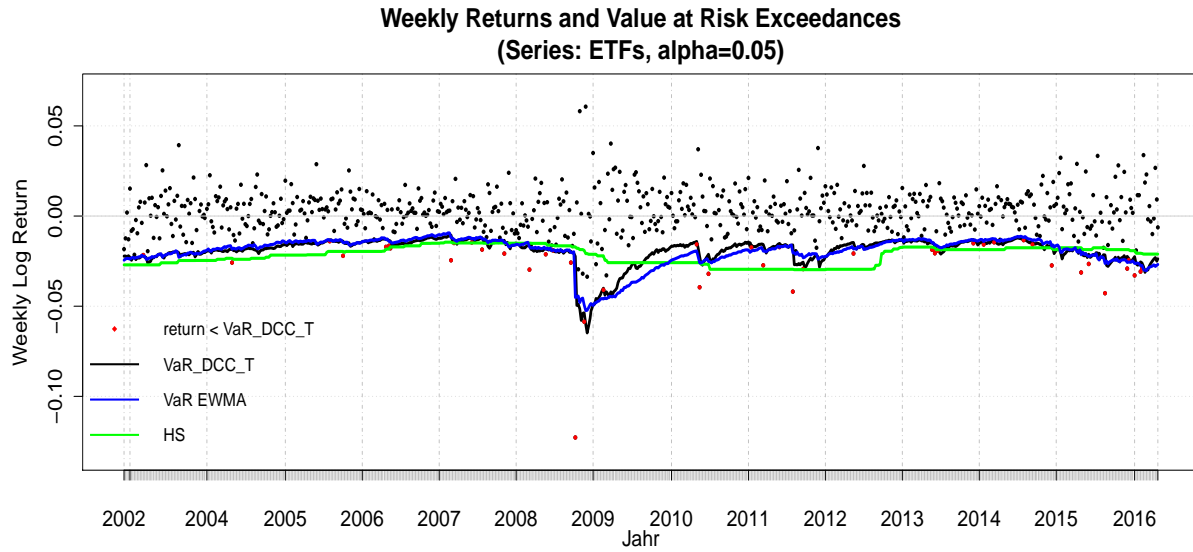


Abbildung 13: DCC Modell multivariat t-verteilt

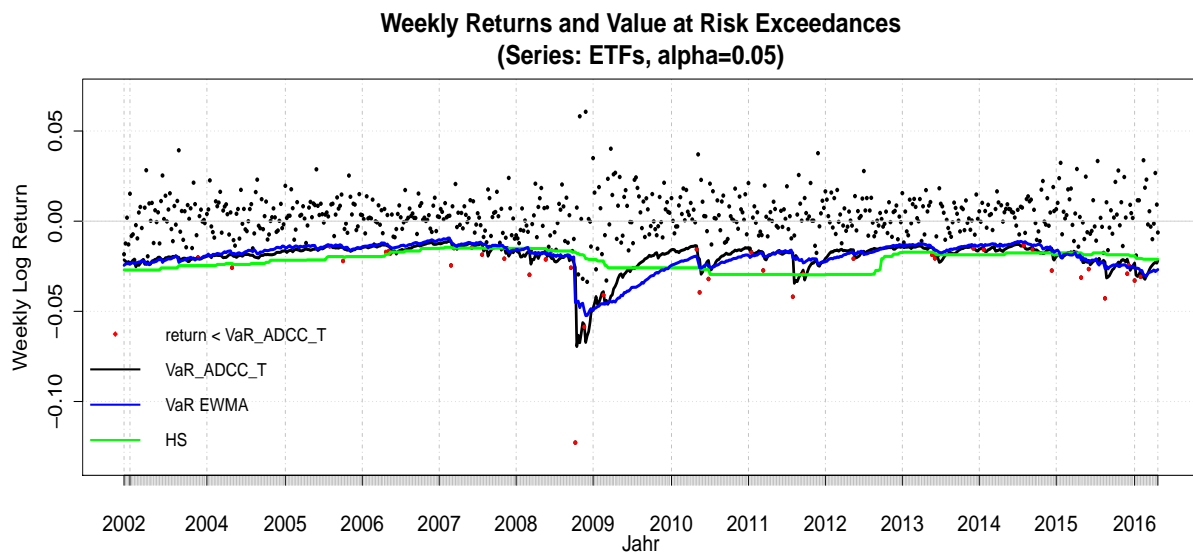


Abbildung 14: ADCC Modell multivariat t-verteilt

Abbildungsverzeichnis

1	wöchentliche Log>Returns von Unternehmensanleihen der Vereinigte Staaten	7
2	ACF-Plots	7
3	Boxplots der Indexfonds	27
4	DCC Modell multivariat normalverteilt	36
5	ADCC Modell multivariat normalverteilt	37
6	Parameter b	37
7	restliche Parameter	38
8	Wöchentliche Log>Returns der Aktienfonds von 15.01.1999 bis 27.04.2016 .	44
9	Wöchentliche Log>Returns des Immobilienfonds und Rohstofffonds von 15.01.1999 bis 27.04.2016	44
10	Log>Returns der Unternehmensanleihen von 15.01.1999 bis 27.04.2016 . . .	45
11	Log>Returns der Besicherten Anleihe von 15.01.1999 bis 27.04.2016	45
12	Log>Returns der Staatsanleihen von 15.01.1999 bis 27.04.2016	45
13	DCC Modell multivariat t-verteilt	47
14	ADCC Modell multivariat t-verteilt	47

Tabellenverzeichnis

1	Tabelle von statistischen Kennzahlen der ETFs	28
2	Rekursives Verfahren	30
3	AIC für die Schätzung über den gesamten Datensatz	34
4	Modellevaluation für $\text{VaR}(\alpha = 0.05)$	35

Inhalt der CD

Die beigelegte CD-ROM enthält folgende Dateien:

- Die Bachelorarbeit als pdf
- Den Ordner Grafiken, in dem alle in der Bachelorarbeit vorkommenden Grafiken im PDF-Format enthalten sind.
- Den Datensatz im ursprünglichen Zustand
- Den Ordner Rdata mit folgenden R-Dateien
 - weekly.RData
 - roll_multinormalverteilt.RData
 - roll_t.RData
 - roll_adcc_grj.RData
 - roll_adcc_t.RData
- Die folgenden R-Skripte
 - Datenaufbereitung
 - Datenanalyse
 - Funktionen

Erklärung der Urheberschaft

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Ort, Datum

Unterschrift