



Copula-GARCH-Modelle

– Bachelorarbeit –

im Studiengang Wirtschaftsmathematik
am Fachbereich Mathematik und Informatik

eingereicht von
Márton Eifert

betreut von
Prof. Dr. Hajo Holzmann

Marburg, 6. September 2010

Inhaltsverzeichnis

Abbildungsverzeichnis	iii
Tabellenverzeichnis	iv
Einleitung	1
1. Finanzökonomische Modellierung	3
1.1. Spezifika von Finanzzeitreihen	4
1.1.1. Begriffe, Kontext und Annahmen	4
1.1.2. Volatilität und bedingte Heteroskedastizität	9
1.1.3. Empirische Befunde im Überblick (Stylized Facts)	10
1.2. Univariate Volatilitätsmodelle: GARCH	12
1.2.1. Definition, Modellaufbau und Eigenschaften	12
1.2.2. Parameterschätzung	14
1.2.3. Schwachpunkte und Erweiterungen	17
1.3. Multivariate Volatilitätsmodelle: MGARCH	19
1.3.1. Ausgangssituation	19
1.3.2. Modellvarianten: Überblick	20
1.4. Überblick	23
2. Copula-Funktionen	25
2.1. Motivation	26
2.2. Theoretisches Konzept der Copula-Funktionen	27
2.2.1. Definitionen	27
2.2.2. Zentrale Aussagen	29
2.2.3. Weitere Eigenschaften	30
2.2.4. Bedingte Copula	32
2.3. Copula-Funktionen in der Statistik	33
2.3.1. Simulation von Zufallszahlen	33
2.3.2. Parametrische Schätzung	41
2.4. Ausblick	49
3. Copula-GARCH-Modelle	51
3.1. Copula-basierte Volatilitätsmodelle	52
3.1.1. Probabilistische Definition	53
3.2. Simulationsstudie	56
3.2.1. Modellaufbau	57
3.2.2. Algorithmen zu Simulation und Schätzung	59
3.2.3. Ergebnisse der Simulationsstudie	61
3.2.4. Anwendung auf reale Daten	61
3.3. Diskussion: Vorzüge und Schwachpunkte	62

Schlusswort	64
A. Ergebnisse der Simulationsstudie	67
B. Ergebnisse der Anwendung	71
Literaturverzeichnis	74

Abbildungsverzeichnis

1.1.	Zeitlicher Verlauf des Preises und der Log>Returns der BMW-Stammaktie (3562 Werktagesdaten zwischen 02.01.1995 und 26.01.2009)	5
1.2.	Dichte-Plots der Standardnormalverteilung sowie einer linksschiefen (negativ asymmetrischen) und einer leptokurtischen (steilgipfligen) Verteilung, die aus je zwei verschiedenen Normalverteilungen gemischt wurde.	8
1.3.	Zeitlicher Verlauf der Log>Returns der BMW-Stammaktie mit geschätzter Volatilität (3562 Werktagesdaten zwischen 02.01.1995 und 26.01.2009) . . .	10
1.4.	Zeitlicher Verlauf der Log>Returns der BMW-Stammaktie (3562 Werktagesdaten zwischen 02.01.1995 und 26.01.2009) im Vergleich mit einem simulierten GARCH(1,1)-Prozess	14
2.1.	Scatterplots der Gumbel- beziehungsweise Clayton-Copula	37
2.2.	Scatterplots der Gauß- beziehungsweise t-Copula	40
A.1.	Dichteplots der Schätzer, Option (a) <i>Gauß-Copula ohne TVC</i>	69
A.2.	Dichteplots der Schätzer, Option (b) <i>t-Copula ohne TVC</i>	69
A.3.	Dichteplots der Schätzer, Option (c) <i>Gauß-Copula mit TVC</i>	70
A.4.	Dichteplots der Schätzer, Option (d) <i>t-Copula mit TVC</i>	70
B.1.	Preisverläufe [oben] und Log>Returns [unten] der Commerzbank- und Deutsche-Bank-Aktien im Vergleich (3562 Werktagesdaten zwischen 02.01.1995 und 26.01.2009)	72
B.2.	Volatilitätsverlauf der Commerzbank- und Deutsche-Bank-Aktien im Vergleich [oben] und zeitabhängige Korrelation mit den Optionen (c) <i>Gauß-Copula mit TVC</i> [rot] und (d) <i>t-Copula mit TVC</i> [blau] und jeweils mittlere Korrelationen (horizontale Linien) [unten] (3562 Werktagesdaten zwischen 02.01.1995 und 26.01.2009)	73

Tabellenverzeichnis

A.1. Überblick über die ursprünglichen Voreinstellungen und gewählten Parameter der Simulationsstudie	67
A.2. Mittlere Schätzergebnisse und Standardfehler für die Parameter aus Option (a) bis (d) sowie p-Werte des Shapiro-Tests auf Normalität	68
B.1. Schätzergebnisse für die Parameter aus den Optionen (a) bis (d) bei Anwendung auf reale Daten (Commerzbank- und Deutsche-Bank-Aktien) . .	71

Einleitung

Quantitative statistische Modelle versuchen die Realität mittels mathematischer Zusammenhänge möglichst genau zu beschreiben. Gerade im finanzwirtschaftlichen Bereich – generell in der Ökonometrie, speziell aber auch im Risikomanagement – sind quantitative Erfassungsmöglichkeiten der Wirklichkeit von enormer Bedeutung: so ist der *Value-at-Risk* ein besonders wichtiges Standardrisikomaß, dessen Konzept mittlerweile vielfach für die Quantifizierung von Risiken angewandt wird. Auch andere Risikomaße als statistische Kennzahlen, wie etwa die klassische *Standardabweichung*, der *Conditional-Value-at-Risk* und viele weitere gehören zu den häufig verwendeten Erfassungsmöglichkeiten für die finanzwirtschaftliche Verlustwahrscheinlichkeit.

Sämtliche Risikomodelle beruhen auf Annahmen bezüglich des entsprechenden Marktes, auf dem diese angewandt werden. Mehrere grundlegende Techniken und Methoden setzten lange Zeit voraus, dass es sich bei der Verteilung von Finanztiteln um eine Normalverteilung handelt. Dies impliziert die Situation einer symmetrischen Verteilung mit Normalgipfel. Allerdings sind derartige Annahmen aufgrund von empirischen Untersuchungen mittlerweile mehrfach widerlegt und verworfen worden, da sie mit den tatsächlichen Beobachtungen kollidierten: zum Beispiel bildet, wie sich herausgestellt hat, eine Verteilung mit dickeren Flanken (eine sogenannte *Heavy-Tailed-Verteilung*) oder mit stärkerer Wölbung als die Normalverteilung (eine sogenannte *leptokurtische Verteilung*) heutige ökonometrische Sachverhalte um einiges besser ab. Da man mit den Normalitätsannahmen auch gleichzeitig davon ausging, dass die lineare Korrelation nach Pearson ein vermeintlich adäquates Abhängigkeitsmaß sei, musste auch hier nun auf allgemeinere Konzepte zurückgegriffen werden. Spätestens seit [5] gewannen hier Copula-Funktionen und auf deren Konzept basierende Abhängigkeitsmaße in Finanzmodellen zunehmend an Bedeutsamkeit.

Die vorliegende Arbeit möchte dieser Problemstellung, der sich zahlreiche Modelle in den letzten zwei bis drei Jahrzehnten gegenübersehen, nachgehen, indem zunächst die nötige Vorarbeit durch eine Einführung in die finanzökonometrische Modellierung und anschließend in die Welt der Copulas gemacht werden soll.

Die finanzökonometrische Modellierung analysiert zeitabhängige Beobachtungen von Finanztiteln (zum Beispiel Renditeverläufen von Aktien) über einen längeren Zeithorizont hinweg. Diese bilden die Grundlage einer umfassenden Risikoanalyse sowie von Prognosemöglichkeiten. Mit den Aspekten der Zeitreihenanalyse im Finanzkontext befasst sich deshalb das erste Kapitel; zuerst mit den Grundbegriffen – wie etwa der Volatilität – und Annahmen in der univariaten Situation und dann mit der nicht minder wichtigen multivariaten Betrachtungsweise. Gerade die Zusammenhänge innerhalb einer Zeitreihenkomponente (serielle Korrelation oder Abhängigkeit), genauso wie komponentenübergrei-

fende Querbeziehungen (Kreuzkorrelation oder -abhängigkeit), sind hier nämlich für den Statistiker von Interesse. Hier sollen daher gängige sogenannte *Volatilitäts-* oder *bedingt-heteroskedastische Modelle* – für univariate wie multivariate Zeitreihen – vorgestellt werden, welche die obigen Normalverteilungsannahmen durch möglichst realistische Merkmale zu ersetzen versuchen.

Mit einem theoretischen Überblick über zentrale Aussagen und Eigenschaften von eingangs genannten Copula-Funktionen und einem praktischen Ausblick in die möglichen Anwendungsbereiche soll anschließend das zweite Kapitel die Grundlagen für den nachfolgenden Kontext schaffen.

In den beiden Kapiteln 1 und 2 wird jeweils das Hauptaugenmerk auf die parametrische Schätzung gelegt, da diese inhaltlich auf das dritte Kapitel dieser Arbeit vorbereiten sollen. Dort werden die beiden vorangehenden Kapitel verknüpft: Viele multivariate Zeitreihenmodelle für Finanztitel haben das Konzept der Copula-Funktion versucht zu integrieren, so dass sich eine eigene Modellklasse der sogenannten *copula-basierten multivariaten Dynamik* (kurz CMD-Modellklasse) gebildet hat. Diese besitzt eine leicht zu implementierende Unterklasse, die die praktischen Vorzüge der Copula-Funktionen mit den bereits weit entwickelten und einfacher gestrickten univariaten Volatilitätsmodellen für die Marginalzeitreihen verbindet. Zusammen lassen sich damit multivariate Volatilitätsmodelle relativ einfach konstruieren und auch für die Parameterschätzung hat diese Modellklasse einige vereinfachende Implikationen. Deshalb befasst sich dieses Kapitel mit den titelgebenden *Copula-GARCH-Modellen*, deren Aufbau und Konzept umfassend dargestellt wird und für welche einige gängige Schätzmethoden aus schätztheoretischer Sicht beschrieben werden. Die dort aufgezeigte Theorie soll am Ende des Kapitels in einer Simulationsstudie veranschaulicht werden und die Einfachheit der Implementierung bestätigen.

Mit einem Fazit als Zusammenfassung, aber auch als kritische Betrachtung der im Laufe der drei Kapitel dargestellten Erkenntnisse und Ergebnisse und einem Ausblick auf aktuelle Forschungs- und Anwendungsgebiete der Copula-Theorie zusammen mit der GARCH-Modellierung runden wir die Arbeit am Ende ab.

Kapitel 1

Finanzökonometrische Modellierung

Als Einstieg in das Thema dieser Arbeit befassen wir uns zuerst mit der finanzökonometrischen Modellierung im Allgemeinen, das heißt der adäquaten Nachbildung der zeitlichen Dynamik von Finanztiteln. Diese zeitliche Dynamik, abgeleitet von einem stochastischen Prozess, spiegelt sich in Form einer *Zeitreihe* wider:

Definition 1.1 (Zeitreihe):

Eine *Zeitreihe* ist eine zufällige Beobachtung (auch Realisation oder Pfad genannt) eines stochastischen Prozesses $(X_t)_{t \in \mathbb{T}}$ mit *diskreter Zeitindexmenge* \mathbb{T} .

Zeitreihen tauchen in unterschiedlichsten Situationen auf, selbst dem Nichtmathematiker begegnen sie im Alltag nicht selten. So treffen wir fast täglich auf aktuelle Aktienkurse oder andere Börsendaten (Finanzzeitreihen) und ebenso finden sich Zeitreihen in Natur- und Ingenieurwissenschaften, in Psychologie und Medizin, den Sozialwissenschaften, in Politik und nicht zuletzt der Mathematik wieder, die diesem Themengebiet die nötige theoretische Würdigung bieten kann.

Bemerkungen:

- Obige Definition besagt, dass wir mit der Zeitreihe $(X_t(\omega))_{t \in \mathbb{T}}$ eine vom zufälligen Elementarereignis $\omega \in \Omega$ abhängige Realisationsfolge erhalten, wobei (Ω, \mathcal{F}, P) ein unbekannter Wahrscheinlichkeitsraum sei, der uns auch im restlichen Verlauf dieser Arbeit zugrunde liegen soll.
- Wir wollen aus Gründen der Einfachheit die Schreibweise $(X_t)_{t \in \mathbb{T}}$ für den abstrakten Zufallsprozess und $(X_t(\omega))_{t \in \mathbb{T}}$ für ein konkret eingetretenes Ereignis $\omega \in \Omega$ synonym verwenden.
- Im realen Experimentfall ist eine Stichprobe natürlich nur von endlichem Umfang. Wir schreiben dann $\mathbb{T} := \{1, \dots, T\}$ mit einer ganzen Zahl $T > 0$, die auch *Stichprobenumfang* oder *Sample-Size* genannt wird. Ohne Einschränkung nehmen wir also positive ganzzahlige Zeitpunkte, meist beginnend mit $t = 1$, an. Für allgemeine theoretische Betrachtungen wird \mathbb{T} nicht näher definiert (beispielsweise kann $\mathbb{T} = \mathbb{N}$

oder gar $\mathbb{T} = \mathbb{Z}$ sein). Definition 1.1 stellt eine strikte Trennung von den zeitkontinuierlichen Prozessen dar, die nicht in dieses Gebiet fallen – der Zustandsraum des Zufallsprozesses kann dagegen ungeachtet dessen sowohl diskret als auch kontinuierlich sein.

In den nächsten Teilen dieses Kapitels durchleuchten wir zuerst (in Abschnitt 1.1) die besonderen Charakteristika von Finanzzeitreihen, einschließlich wichtiger Begrifflichkeiten und zu berücksichtigender Komponenten. Anschließend beschäftigen wir uns in Teil 1.2 mit geeigneten ökonometrischen Modellen sowohl für die direkt beobachteten als auch nicht beobachtbaren zugrundeliegenden Komponenten einer Finanzzeitreihe. Dort werden wir speziell das erwähnte univariate GARCH-Modell einführen und diskutieren sowie multivariate Verallgemeinerungen vorstellen. Das Ende des Kapitels wird schließlich durch eine kurze Zusammenfassung dieses Kapitels in Punkt 1.4 eingeleitet.

1.1. Spezifika von Finanzzeitreihen

Anfangen bei Aktienkursen (oder allgemein Kursen von Vermögenswerten) eines bestimmten Unternehmens über Wechselkurse zwischen verschiedenen Währungen bis hin zu Zinsverläufen beispielsweise auf Kreditmärkten existiert wohl gerade in der Finanzwelt eine schier unüberschaubare Zahl zeitvariierender Daten. In diesem Kontext besonders wichtige Begriffe wie etwa *Rendite* oder *Volatilität* sollen in den kommenden Unterabschnitten kurz eingeführt und erläutert werden, denn wir werden diese Fachtermini von nun an häufiger antreffen. Danach verschaffen wir uns einen groben Überblick über geeignete Annahmen hinsichtlich der typischen Verteilungen von Finanzzeitreihen, aber auch Spezifika wie die Einbeziehung der *bedingten Heteroskedastizität* in neuartigen Finanzmodellen werden hier vorgestellt.

1.1.1. Begriffe, Kontext und Annahmen

In den folgenden Absätzen sollen unter anderem kurz die Begriffe „Rendite“ sowie „Shocks“, „Volatilität“ beziehungsweise „bedingte Heteroskedastizität“ definiert werden, damit die Begrifflichkeiten und der Kontext griffbereit sind.

Rendite

Primär handelt es sich bei den direkt beobachteten Daten um den Preisverlauf eines Finanztitels, formal die Zeitreihe $(P_t)_{t \in \mathbb{T}}$, beispielsweise von Wertpapieren. Doch diese ist in ihrer untransformierten Version für statistische Analysen relativ schlecht geeignet, außerdem weist $(P_t)_{t \in \mathbb{T}}$ meist lokale Trends auf, weshalb wir die Zeitreihe noch „stationarisieren“

müssen. Uns interessieren daher als indirekt beobachtete zeitvariable Daten vielmehr die sogenannten *Renditen* (englisch *Asset Returns* oder einfach *Returns*) eines Finanztitels. Zwei häufig benutzte Renditetypen stellen wir nun vor.

Definition 1.2 (Gewöhnliche Rendite):

Sei P_t der Preis für eine Anlage zum Zeitpunkt t . Die *gewöhnliche Netto-Rendite* (englisch *Simple Net Return*) wird definiert durch

$$R_t := \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Alternativ wird manchmal die *gewöhnliche Brutto-Rendite* (englisch *Simple Gross Return*) $\tilde{R}_t := R_t + 1 = \frac{P_t}{P_{t-1}}$ betrachtet. Diese interpretieren wir als den *relativen Ertrag*, der sich aus dem Verhältnis „neuer Preis zu alter Preis“ ergibt.

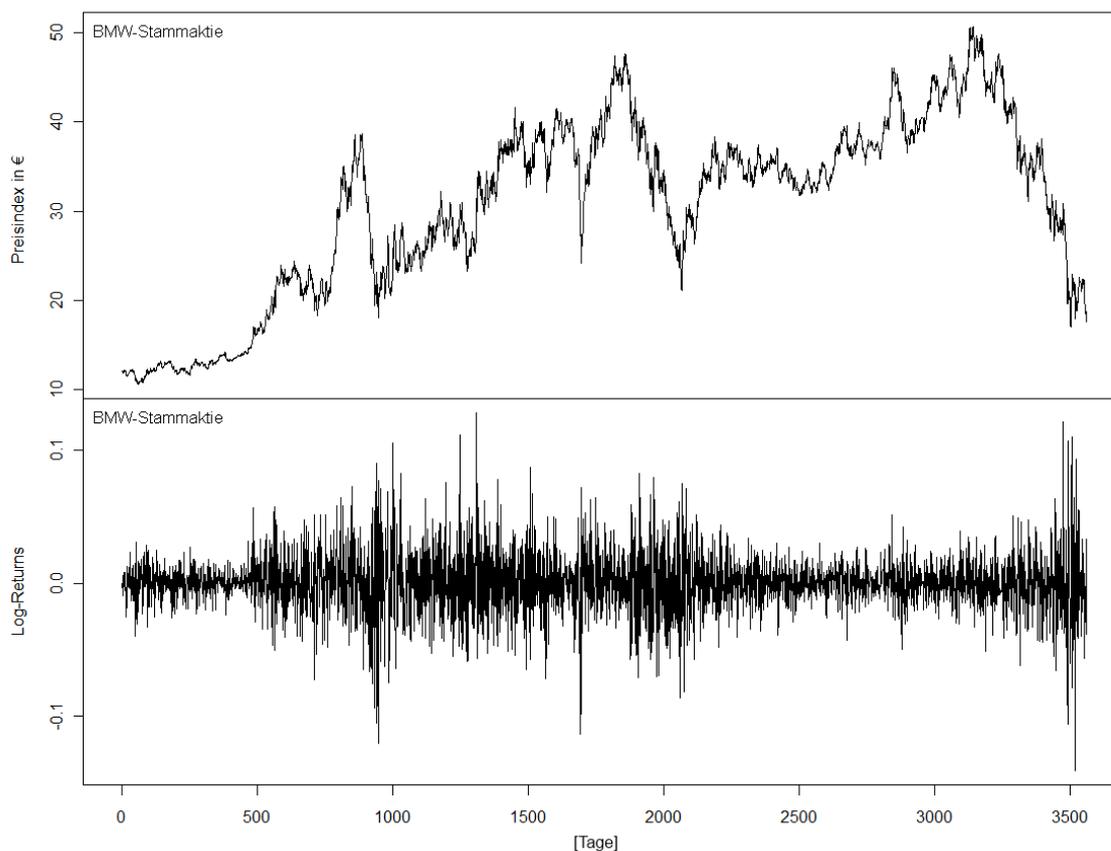


Abbildung 1.1.: Zeitlicher Verlauf des Preises und der Log>Returns der BMW-Stammaktie (3562 Werktagessdaten zwischen 02.01.1995 und 26.01.2009)

Definition 1.3 (Log-Rendite):

Sei P_t der Preis für eine Anlage zum Zeitpunkt t . Die *Log-Rendite* (englisch *Log-Return*) wird definiert durch

$$r_t := \ln(\tilde{R}_t) = \ln\left(\frac{P_t}{P_{t-1}}\right) = p_t - p_{t-1},$$

wobei man $p_t := \ln(P_t)$ als Abkürzung setzt.

Renditen sind eine brauchbare Transformation des direkt beobachteten Preises und können selbst wieder als Zeitreihe $(R_t)_{t \in \mathbb{T}}$ beziehungsweise $(r_t)_{t \in \mathbb{T}}$ aufgefasst werden. Die Betrachtung von Renditen einer Anlage hat mehrere Vorteile gegenüber der des Preises. Zum einen stellen sie ein normiertes Maß dar, beispielsweise den Ertrag aus einer zu Periodenbeginn getätigten Investition einer Geldeinheit, zum anderen sind diese Daten für statistische Analysen und Prognosen aufgrund der Stationarität besser modellierbar. Im weiteren Verlauf werden sowohl die englischen als auch die deutschen Begriffe gleichermaßen verwendet.

Wegen $\ln(1+x) \approx x$ als Approximation erster Ordnung gibt es für x nahe Null keine signifikanten Unterschiede zwischen der Verwendung der gewöhnlichen Renditen R_t und den Log>Returns r_t , was vor allem bei der Betrachtung von hochfrequentierten Daten wie etwa Tagesdaten gegeben ist (Renditen erreichen betraglich nur selten mehr als 10%). Neben diesen Varianten für Returns werden im Übrigen oft auch weitere Typen wie zum Beispiel *Portfolio-Return* oder *Excess-Return* verwendet.

Shocks

Als Shocks (oft auch *Innovationen*, *Störterme* oder *Residuen*) bezeichnet man im ökonomischen Modellierungskontext völlig zufällige Einflüsse, die neben (nichtzufälligen) erklärenden Variablen auf die Renditen einwirken. Formal handelt es sich bei der Shock-Zeitreihe $(z_t)_{t \in \mathbb{T}}$ meist um ein *weißes Rauschen* (auch *White-Noise-Prozess*, kurz *WN-Prozess*), also eine unabhängige und identisch verteilte Zeitreihe, die streng stationär mit Erwartungswert Null sowie konstanter Varianz σ^2 ist. Oft schreibt man kurz $z_t \sim \text{WN}(0, \sigma^2)$ und sehr häufig wird die Normierung $\sigma^2 = 1$ vorgenommen.

Zeitvariierende Momente und Verteilungsannahmen

Um die Illustration des Kontexts zu vervollständigen, benötigen wir noch weitere Grundbegriffe definiert aus wahrscheinlichkeitstheoretischer Sicht, etwa die zu einem bestimmten Zeitpunkt $t \in \mathbb{T}$ bekannte *Information* sowie die (auf die Information aus der Vergangenheit) *bedingten Momente* eines Prozesses.

Definition 1.4 (Zeitvariierende Information):

Bis zum Zeitpunkt t ist uns eine gewisse Grundgesamtheit an Information (beispielsweise über die bisherigen Kursverläufe) verfügbar, die wir formal als σ -Algebra

$$\mathcal{F}_t := \sigma(r_1, \dots, r_t) := \sigma(r_1^{-1}(B), \dots, r_t^{-1}(B) : B \in \mathcal{B}(\mathbb{R}))$$

definieren, welche eine Teilmenge der gesamten Prozessinformation $\mathcal{F} := \sigma(r_t : t \in \mathbb{T})$ bildet.

Auf diese Information aus der Vergangenheit \mathcal{F}_{t-1} können wir zum Beispiel die ersten beiden Momente, *Erwartungswert* und *Varianz*, bedingen.

Definition 1.5 (Zeitvariierende (bedingte) Momente):

Für den Renditeprozess $(r_t)_{t \in \mathbb{T}}$ definieren wir die ersten zwei zeitabhängigen und auf die Information \mathcal{F}_{t-1} bedingten Momente:

- (1.) Der *bedingte Erwartungswert* von r_t ist gegeben durch

$$\mu_t := \mathbb{E}[r_t | \mathcal{F}_{t-1}]. \quad (1.1)$$

- (2.) Die *bedingte Varianz* von r_t wird durch das zweite bedingte Moment des zentrierten Prozesses $\bar{r}_t := r_t - \mu_t$ definiert, also

$$\sigma_t^2 := \text{Var}(r_t | \mathcal{F}_{t-1}) = \mathbb{E}[\bar{r}_t^2 | \mathcal{F}_{t-1}]. \quad (1.2)$$

Im Finanzjargon – und somit in unserem Kontext – wird σ_t^2 auch *Volatilität* der Rendite r_t genannt.

Sind die r_t , $t \in \mathbb{T}$, unabhängig von der Information \mathcal{F}_{t-1} aus der Vergangenheit, so fällt die Bedingtheit weg und es ergeben sich die gewöhnlichen Kenngrößen *Erwartungswert* beziehungsweise *Varianz* von r_t . Alternativ erhält man die unbedingten Momente auch als Erwartungswert der bedingten Momente (Gesetz vom iterierten Erwartungswert).

Diese eben definierten Größen charakterisieren in ihren unbedingten Versionen die zugrundeliegende Verteilung einer Reihe von Daten im Allgemeinen zwar noch nicht eindeutig, als wichtigste Kenngrößen reichen sie jedoch häufig schon aus. Bei elliptischen Verteilungen genügen aber zum Beispiel bereits die ersten beiden Momente. Nun stellen wir uns die Frage: Was zeichnet gerade die Verteilungen speziell von Renditezeitreihen aus?

Bis vor etwa fünfzig Jahren gab es außer in L. Bacheliers Arbeiten aus Vorweltkriegszeiten (siehe [2]) keinen probabilistischen Zugang zu finanzökonometrischen Sachverhalten. So war man fast ausschließlich – ohne empirische Befunde – davon ausgegangen, dass die

Normalverteilung die wohl adäquateste Verteilung für Finanzzeitreihen wäre. Doch mittlerweile wurden diese Annahmen mehrfach widerlegt (siehe [22] oder [8]), da vor allem Renditezeitreihen oft zwei wichtige Merkmale aufweisen, die mit der Normalverteilungsannahme kollidieren: Als erstes sind die Verteilungen häufig *nicht symmetrisch*, sondern weisen eine *negative Schiefe* auf (man spricht auch von *Linksschiefe*), das bedeutet, dass negative Ereignisse wahrscheinlicher sind als entsprechende positive. Im Falle von Symmetrie wie bei der Normalverteilung wäre das dritte Moment, die *Schiefe*, gleich Null. Hierdurch wird auch eine der Ursachen des *Leverage-Effekts* erklärt. Weiterhin sind die Verteilungen sehr oft *leptokurtisch*, ein Phänomen, das auftritt, wenn an den äußeren Flanken der Verteilungsdichte die Wahrscheinlichkeiten für extreme Ereignisse generell um einiges höher sind als bei der Normalverteilung, so dass der Gipfel der Wahrscheinlichkeitsdichte vergleichsweise spitzer wird. Dieser Befund hängt mit dem vierten Moment, der *Wölbung* (oder auch *Kurtosis*), zusammen, die bei der Normalverteilung stets den konstanten Wert Drei hat und bei den empirischen Verteilungen von Finanzdaten deutlich darüber liegt. In einem solchen Zusammenhang verwendet man für die Randbereiche einer Verteilung auch öfters die englischen Bezeichnungen „fat tails“ oder „heavy tails“ oder es ist von *Heavy-Tailed-Verteilungen* die Rede, die vor allem in der Extremwerttheorie umfangreich verwendet werden.

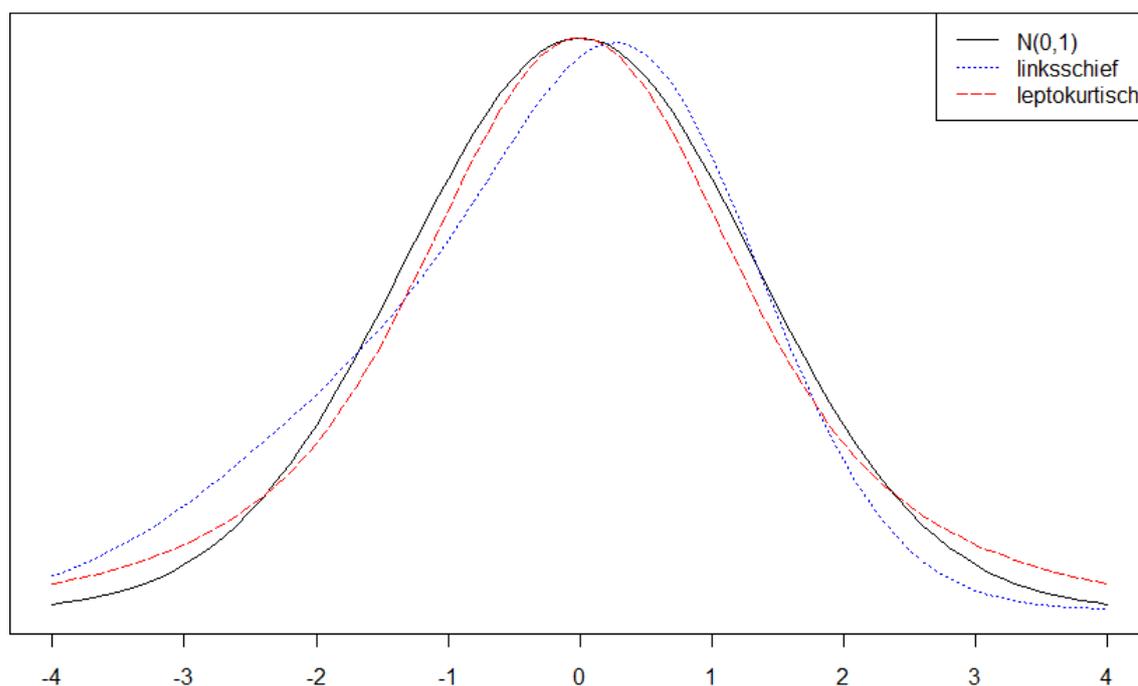


Abbildung 1.2.: Dichte-Plots der Standardnormalverteilung sowie einer links-schiefen (negativ asymmetrischen) und einer leptokurtischen (steilgipfligen) Verteilung, die aus je zwei verschiedenen Normalverteilungen gemischt wurde.

Wir wollen hier auf eine formale Definition der eben genannten höheren Momente *Schiefe* und *Kurtosis* nicht näher eingehen, illustrieren aber wenigstens in Abbildung 1.2 ein

anschauliches Beispiel für linksschiefe (negativ asymmetrische) sowie leptokurtische (steilgipflige) Verteilungsdichten im Vergleich zur Dichte der Standardnormalverteilung.

Diese empirischen Befunde – lange Zeit unter anderem aus Mangel an geeigneten Modellen nicht beachtet – hatten mit ihrem Nachweis und anschließender allgemeiner Akzeptanz für eine Reform in der Finanzmarktmodellierung gesorgt, denn gerade für diesen Bereich sind möglichst realistische und genaue Modelle von enorm hoher Bedeutung. Aus diesem Grund geht man heutzutage oft von leptokurtischen und asymmetrischen Verteilungen wie beispielsweise der (schiefen) Student-t-Verteilung aus. Bei dieser sind Schiefe und Wölbung vom Freiheitsgrad-Parameter abhängig, den man deshalb als alternative Maßzahl für diese Momente verwenden kann.

1.1.2. Volatilität und bedingte Heteroskedastizität

Schwankungen der Rendite werden in der Finanzterminologie *Volatilitäten* genannt. Dass sich bei Finanzzeitreihen die Volatilität im Laufe der Betrachtungszeit ändert, ist (besonders bei hochfrequenten Daten) seit [6] mehrfach bestätigt worden und herrscht als generelle Auffassung mittlerweile seit einigen Jahrzehnten vor. Dieses Phänomen nennt man in der Statistik *bedingte Heteroskedastizität* und wird definiert als eine Ungleichmäßigkeit oder Zeitvariation der bedingten Varianz innerhalb einer Datenmessung, sprich: Innerhalb einer Reihe von zeitlich aufeinander folgenden Messungen ist die Volatilität der Daten nicht zeitkonstant (die Daten somit *nicht bedingt-homoskedastisch*). Die Volatilität eines Finanztitels ist jedoch nicht ad hoc beobachtbar, sondern muss geschätzt werden. Eine möglichst genaue und zuverlässige Schätzung der Volatilität ist auch deshalb wichtig für große Kreditinstitute, weil sie in direktem Verhältnis zu Risikomaßen steht. Zwar gibt es auch exogene makroökonomische Faktoren, von denen man ein gewisses Einflusspotential auf die Volatilität erwarten könnte, doch hat man empirisch festgestellt, dass diese am ehesten noch von eigenen Rendite- und Volatilitätswerten abhängig ist (endogene Einflüsse). In Abschnitt 1.2 befassen wir uns daher mit geeigneten Modellansätzen aus der Finanzzeitreihenanalyse. Volatilitäten haben einige interessante und – vor allem für das finanzwirtschaftliche Risiko-Management – wichtige Aspekte.

Die folgenden Absätze sollen einige dieser empirisch untermauerten Merkmale der Volatilität einer Finanzzeitreihe illustrieren, welche wir am Ende dieses Abschnitts noch einmal zusammenfassen werden.

Ein erstes empirisch ausführlich untersuchtes Erscheinungsbild ist, dass Volatilitäten in sogenannten *Clustern* auftreten (auch *Volatilitäts-Clustering* genannt), das heißt, kleine Schwankungen werden tendenziell von kleinen gefolgt, umgekehrt trifft man häufig auf große Schwankungen zusammen mit anderen großen innerhalb eines zeitlich zusammenhängenden Bereichs. Außerdem weist die Volatilität die Besonderheit auf, auf vergleichbar

starke positive sowie negative Ausschläge des Preises in unterschiedlichem Ausmaß zu reagieren – ein Merkmal, das *Leverage-Effekt* genannt wird und mit linksschiefen Verteilungen in Verbindung steht.¹

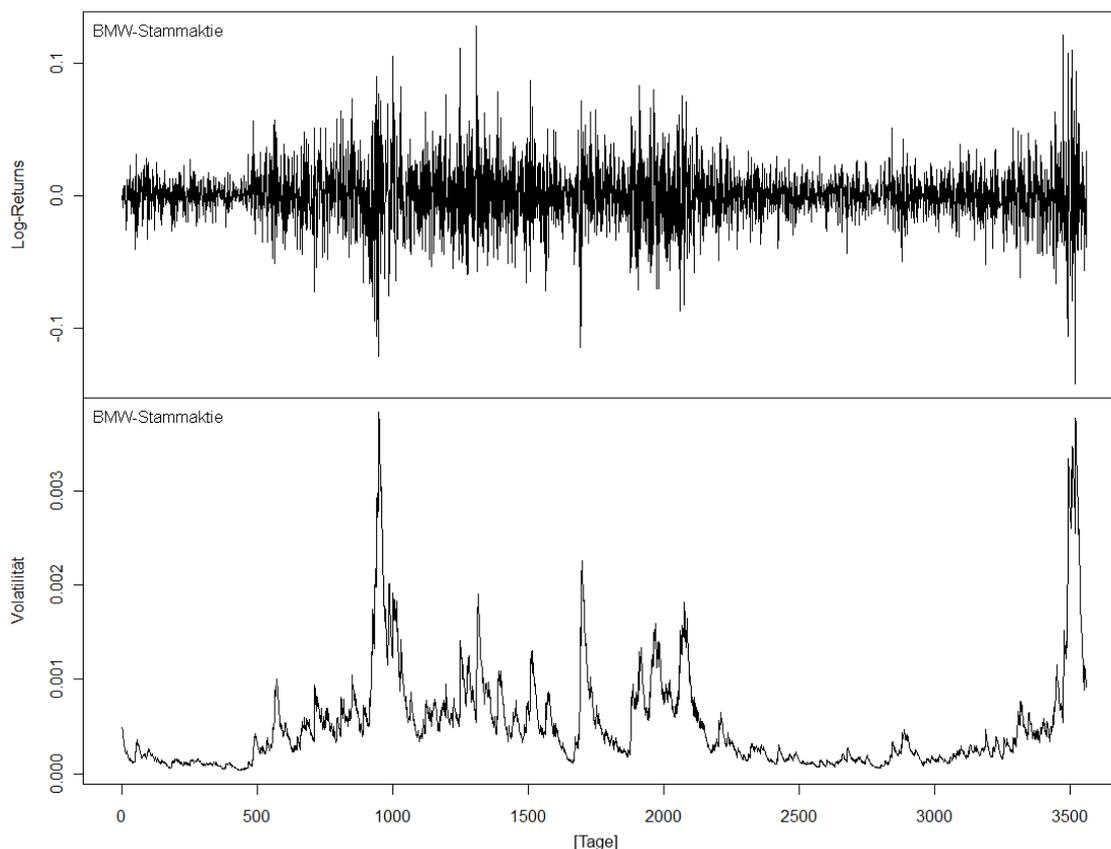


Abbildung 1.3.: Zeitlicher Verlauf der Log>Returns der BMW-Stammaktie mit geschätzter Volatilität (3562 Werktagesdaten zwischen 02.01.1995 und 26.01.2009)

Nun haben wir einige grundlegende Definitionen, Fachtermini und empirische Fakten über Finanzzeitreihen kennengelernt. Im kommenden Abschnitt sind diese noch einmal in einer Übersicht zusammengefasst. Anschließend wenden wir uns der bedingten Heteroskedastizität aus der Modellierungsperspektive zu.

1.1.3. Empirische Befunde im Überblick (Stylized Facts)

Zusammenfassend ergibt sich eine Vielzahl an Eigenschaften, die Renditezeitreihen als besondere finanzökonometrische Beobachtungsobjekte von anderen Zeitreihen hervorheben. In der Literatur oftmals als *stilisierte Fakten* (englisch *Stylized Facts*) bezeichnet, bilden unter anderem folgende Aspekte die wichtigsten Charakteristika einer Finanzzeitreihe,

¹Vergleiche [36], S. 99.

beruhend auf allgemein anerkannten Ergebnissen aus jahrelangen und umfassenden empirischen Untersuchungen:

- (1.) Renditezeitreihen sind nicht unabhängig und identisch verteilt, obwohl diese nur eine geringe serielle Korrelation aufweisen.
- (2.) Die Absolutbeträge und Quadrate von Renditezeitreihen sind dagegen stark seriell korreliert.
- (3.) Die bedingte Erwartung von Renditen ist nahezu gleich Null.
- (4.) Es liegt bedingte Heteroskedastizität vor, also eine zeitvariierende Volatilität, die im Zeitverlauf clusterweise auftritt.
- (5.) Manche Renditezeitreihen sind linksschief (*Leverage-Effekt*), viele sind leptokurtisch.

Auch im multivariaten Fall wurden ähnliche Merkmale festgestellt, auf die wir in Abschnitt 1.3 näher eingehen werden. So gilt:

- (6.) Die parallele Korrelation (Kreuzkorrelation mit Lag Null) variiert in der Zeit; sie ist zwischen den Komponenten signifikant höher als die zeitversetzte Kreuzkorrelation (Kreuzkorrelation mit Lag größer Null), welche selbst nur zwischen den Absolutbeträgen und Quadraten der Zeitreihenkomponenten hoch wird (vergleiche die Sachverhalte aus (1.) und (2.)).
- (7.) Extreme Ausschläge einer Zeitreihenkomponente gehen häufig mit extremen Ausschlägen anderer Komponenten einher.

Wie wir gesehen haben, sind diese Merkmale sehr nennenswerte Zeitreiheneigenschaften, die nur bei Renditen in solcher Form und in solchem Ausmaß diagnostiziert werden. Sie bedürfen deshalb auch eigener Modelle, um sie adäquat nachbilden zu können. Dieser Aufgabenstellung haben sich Ökonometriker seit Mitte des 20. Jahrhunderts auf der ganzen Welt gewidmet und noch heute werden ständig neue Modellvarianten für spezielle Zwecke entwickelt und vorgestellt, meist um die Schwächen anderer bereits bestehender auszubügeln. Daher wenden wir uns im kommenden Abschnitt der Einführung solcher *Volatilitäts-* oder auch *bedingt-heteroskedastischer Modelle* zu, die für einen ersten weltweiten Durchbruch in der Modellierung von Finanzzeitreihen gesorgt haben. Zuerst betrachten wir im Univariaten die sogenannten GARCH-Modelle, anschließend die multivariaten Erweiterungen mit Blick auf deren Fähigkeit, die Stylized Facts angemessen nachzubilden, aber auch mögliche Schwachpunkte.

1.2. Univariate Volatilitätsmodelle: GARCH

In den Anfängen der Ökonometrie gab es zunächst nur Modelle, die von einer zeitkonstanten (bedingten) Varianz für die Renditen r_t ausgingen. *Volatilitätsmodelle* oder *bedingtheheteroskedastische Modelle*, die also eine zeitliche Veränderung der Schwankung als Einflussfaktor und eigene grundlegende Modellkomponente integrieren, wurden erstmals in den 1980er Jahren eingeführt. Maßgeblich trugen hier die Statistiker R. Engle und T. Bollerslev zu einem Fortschritt in der finanzökonometrischen Modellierung bei, indem sie mit ihren revolutionären Entwicklungen, den ARCH- beziehungsweise GARCH-Modellen, welche die tatsächlichen Phänomene auf Finanzmärkten deutlich realistischer und vollständiger nachbilden und prognostizieren konnten als die bis dahin lange Zeit auch in der Praxis gebrauchten Modelle. Allein das GARCH-Modell von Bollerslev als Ergänzung von Engles ARCH-Klasse wird im nun kommenden Abschnitt kurz vorgestellt und durchleuchtet.

1.2.1. Definition, Modellaufbau und Eigenschaften

Die Betrachtungen beim ARCH- respektive GARCH-Konzept setzen zunächst bei der Aufspaltung der Rendite r_t in ein Produkt aus bedingter Varianz (oder Volatilität) $\sigma_t^2 = \text{Var}(r_t | \mathcal{F}_{t-1})$ und standardisierten zufälligen Shocks z_t an.

Der Entwickler der ARCH-Modellklasse machte für die Einbindung der bedingten Heteroskedastizität von Finanzzeitreihen erstmals den Ansatz, die Volatilität zeitdynamisch in Abhängigkeit von quadrierten historischen Renditedaten r_{t-i}^2 , $1 \leq i \leq p$, zu modellieren, sprich $\sigma_t^2 = f(r_{t-1}, \dots, r_{t-p})$. Diese Idee griff Bollerslev auf und entwickelte daraus die GARCH-Modellklasse, indem er die Volatilität zusätzlich noch mit vergangenen Volatilitätswerten σ_{t-j}^2 , $1 \leq j \leq q$, kombinierte, formal also $\sigma_t^2 = f(r_{t-1}, \dots, r_{t-p}, \sigma_{t-1}, \dots, \sigma_{t-q})$. Diese Bausteine zusammengesetzt führen uns zu folgender Definition.

Definition 1.6 (GARCH-Modell):

Ein *GARCH*(p, q)-Modell für die Renditezeitreihe $(r_t)_{t \in \mathbb{T}}$ mit Volatilität $(\sigma_t^2)_{t \in \mathbb{T}}$ besteht aus

(a) der *Renditegleichung*

$$r_t = \sigma_t \cdot z_t \quad (1.3)$$

und

(b) der *Volatilitätsgleichung*

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{T}. \quad (1.4)$$

Dabei seien $z_t \stackrel{iid}{\sim} \text{WN}(0, 1)$ zufällige Residuen und ohne Einschränkung gelte $\alpha_p \neq 0 \neq \beta_q$ für die Eindeutigkeit der Grade p und q . Im Spezialfall $q = 0$ reduziert sich die Volatilitätsgleichung (1.4) auf den ARCH-Ansatz von Engle und wir setzen die Bezeichnungen „GARCH($p,0$)-Modell“ und „ARCH(p)-Modell“ gleich.

Bemerkungen:

- Ein Prozess $(r_t)_{t \in \mathbb{T}}$ wie in obiger Definition 1.6 heißt *GARCH-Prozess mit Graden p und q* .
- Für die ersten beiden bedingten Momente eines GARCH-Prozesses gilt offensichtlich

$$\mathbb{E}[r_t | \mathcal{F}_{t-1}] = \sigma_t \cdot \mathbb{E}[z_t] = 0$$

sowie

$$\text{Var}(r_t | \mathcal{F}_{t-1}) = \mathbb{E}[r_t^2 | \mathcal{F}_{t-1}] = \sigma_t^2.$$

Außerdem gilt

$$\text{Cov}(r_t, r_{t+h}) = \mathbb{E}[\mathbb{E}[\sigma_t z_t \cdot \sigma_{t+h} z_{t+h} | \mathcal{F}_{t+h}]] = 0$$

für die Autokovarianz. Also ist ein GARCH-Prozess eine Zeitreihe aus unkorrelierten und zentrierten Zufallsvariablen, womit zwei weitere empirische Befunde für Finanzzeitreihen erfüllt wären (vergleiche Stylized Facts (1.) und (3.), Unterabschnitt 1.1.3).

- Insgesamt hat also ein GARCH(p,q)-Modell $1 + p + q$ zu schätzende Parameter. Es sei angemerkt, dass man mit dem reinen ARCH(p)-Modell zwar auch schon Volatilitätscluster nachbilden kann (vergleiche Stylized Fact (4.), Unterabschnitt 1.1.3). Allerdings benötigt man eine weitaus höhere Ordnung p und damit viel mehr Parameter – selbst im direkten Vergleich mit dem beispielsweise am häufigsten verwendeten Modell, GARCH(1,1), bei dem schon drei Parameter für den gewünschten Effekt ausreichen. Mit dem allgemeineren GARCH-Modell ist das Clustering also viel leichter zu modellieren.

Damit die Volatilität σ_t^2 nichtnegativ bleibt, muss an die Koeffizienten als Bedingung wenigstens $\omega > 0$ sowie $\alpha_i \geq 0$, $1 \leq i \leq p$, und $\beta_j \geq 0$, $1 \leq j \leq q$, gestellt werden.²

Man kann auch weitere Restriktionen an die Parameter einführen, zum Beispiel für die Endlichkeit von Schiefe und Wölbung, um die zugrundeliegende Verteilung näher charakterisieren und Abhängigkeitsstrukturen auch in den Tails modellieren zu können.

²Für strenge Stationarität eines GARCH-Prozesses sind noch weitere Parameterrestriktionen erforderlich, vergleiche zum Beispiel [21], die hinreichende und notwendige Bedingungen liefern.

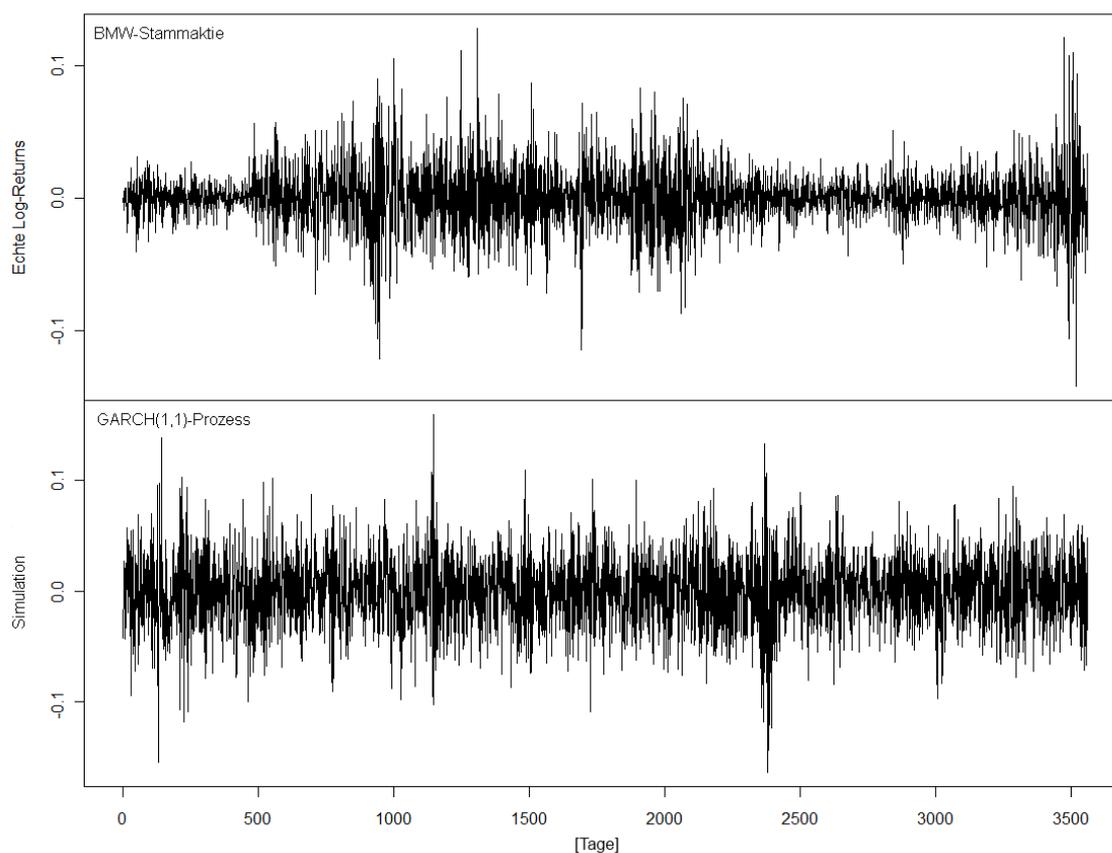


Abbildung 1.4.: Zeitlicher Verlauf der Log-Returns der BMW-Stammaktie (3562 Werktagessdaten zwischen 02.01.1995 und 26.01.2009) im Vergleich mit einem simulierten GARCH(1,1)-Prozess

1.2.2. Parameterschätzung

Die Parameterschätzung bei GARCH-Modellen erfolgt, wie in den meisten ökonomischen Anwendungen, klassisch mittels der *Maximum-Likelihood-Methode* (kurz ML). In [39] wird ausführlich auf ML-Schätzverfahren und -techniken unter finanzwirtschaftlichen Modellspekten eingegangen. Besonders ist jedoch zu erwähnen, dass Schätzverfahren, wie man sie für ARMA-Modelle verwendet, nur schwer auf GARCH-Parameter zu übertragen sind, obwohl der quadrierte Prozess $(r_t^2)_{t \in \mathbb{T}}$ einer ARMA-Gleichung genügt.³ Wir wollen hier lediglich auf die in [17] beschriebene ML-Schätzung der GARCH-Parameter eingehen.

Im Folgenden sei der Illustration halber angenommen, dass es sich beim Residuenprozess $(z_t)_{t \in \mathbb{T}}$ um Gaußsches weißes Rauschen handelt, dieser also unabhängig und identisch aus einer Normalverteilung hervorgeht. Daraus folgt, dass auch die bedingte Ver-

³Dies verdeutlichen beispielsweise [21].

teilung der Rendite r_t eine Normalverteilung ist.⁴ Sei ferner $\mathbb{T} = \{1, \dots, T\}$ die Zeitindexmenge mit endlichem Stichprobenumfang $T > m := \max(p, q)$ und schließlich $\boldsymbol{\theta} := (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)^\top$ der zum GARCH(p, q)-Modell aus Definition 1.6 gehörige Parametervektor, den wir schätzen wollen.

Die (auf alle vergangenen Realisationen) bedingte univariate Dichte von r_t sei mit $f_{|t-1}$ bezeichnet, wir erhalten also mit

$$f_{|t-1}(r_t; \boldsymbol{\theta}) := f(r_t; \boldsymbol{\theta} | r_{t-1}, \dots, r_1) = \frac{1}{\sigma_t(\boldsymbol{\theta})\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{r_t^2}{\sigma_t^2(\boldsymbol{\theta})}\right), \quad t \in \mathbb{T},$$

eine durch $\boldsymbol{\theta}$ parametrisierte Darstellung für die bedingte Normalverteilung.

Aus dem T -fachen Produkt der univariaten Dichten setzt sich die multivariate gemeinsame Dichte

$$h(r_1, \dots, r_T; \boldsymbol{\theta}) = \prod_{t=1}^T f_{|t-1}(r_t; \boldsymbol{\theta})$$

der *kompletten* Beobachtung r_1, \dots, r_T zusammen.

Dann gilt für die *bedingte Log-Likelihood-Funktion* $\mathcal{L}(\boldsymbol{\theta})$ die Schreibweise

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &:= \ln(h(r_1, \dots, r_T; \boldsymbol{\theta})) \\ &= \ln\left(\prod_{t=1}^T f_{|t-1}(r_t; \boldsymbol{\theta})\right) \\ &= \sum_{t=1}^T \ln(f_{|t-1}(r_t; \boldsymbol{\theta})) \end{aligned} \tag{1.5}$$

mit $t \in \mathbb{T}$ und $f_{|0} \equiv f$. Benennen wir weiterhin mit $\ell_t(\boldsymbol{\theta}) := \ln(f_{|t-1}(r_t; \boldsymbol{\theta}))$ die Summanden in Gleichung (1.5), so ergibt sich jeweils

$$\begin{aligned} \ell_t(\boldsymbol{\theta}) &= -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \sigma_t^2(\boldsymbol{\theta}) - \frac{r_t^2}{2\sigma_t^2(\boldsymbol{\theta})} = \\ &= -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \left(\ln \sigma_t^2(\boldsymbol{\theta}) + \frac{r_t^2}{\sigma_t^2(\boldsymbol{\theta})} \right), \quad t \in \mathbb{T}, \end{aligned} \tag{1.6}$$

und insgesamt wieder

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^T \ell_t(\boldsymbol{\theta}) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \left(\ln \sigma_t^2(\boldsymbol{\theta}) + \frac{r_t^2}{\sigma_t^2(\boldsymbol{\theta})} \right). \tag{1.7}$$

⁴In [36], S. 107f, findet sich alternativ auch eine beispielhafte Erläuterung der Parameterschätzung eines Student-t-verteiltern ARCH-Prozesses.

Der bedingte ML-Schätzer $\widehat{\boldsymbol{\theta}}$ für $\boldsymbol{\theta}$ ist definiert als das maximierende Argument von \mathcal{L} und damit das von allen Summanden ℓ_t gleichzeitig. Differenzieren wir einmal nach dem Parametervektor $\boldsymbol{\theta}$, so bekommen wir den ML-Schätzer gerade als Lösung des Gleichungssystems

$$\begin{aligned}\nabla \ell_t(\boldsymbol{\theta}) &:= \frac{\partial}{\partial \boldsymbol{\theta}} \ell_t(\boldsymbol{\theta}) = -\frac{1}{2} \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \sigma_t^2(\boldsymbol{\theta}) + \frac{\partial}{\partial \boldsymbol{\theta}} \frac{r_t^2}{\sigma_t^2(\boldsymbol{\theta})} \right) \\ &= -\frac{1}{2\sigma_t^2(\boldsymbol{\theta})} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{r_t^2}{2\sigma_t^4(\boldsymbol{\theta})} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ &= \frac{1}{2\sigma_t^2(\boldsymbol{\theta})} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left(\frac{r_t^2}{\sigma_t^2(\boldsymbol{\theta})} - 1 \right) \stackrel{!}{=} \mathbf{0}.\end{aligned}\tag{1.8}$$

Dabei gilt für die konkreten partiellen Ableitungen nach den einzelnen $\boldsymbol{\theta}$ -Komponenten:

$$\frac{\partial}{\partial \omega} \sigma_t^2 \equiv 1, \quad \frac{\partial}{\partial \alpha_i} \sigma_t^2(\boldsymbol{\theta}) = \sigma_{t-i}^2(\boldsymbol{\theta}) \quad (1 \leq i \leq p), \quad \frac{\partial}{\partial \beta_j} \sigma_t^2(\boldsymbol{\theta}) = r_{t-j}^2 \quad (1 \leq j \leq q),$$

so dass insgesamt der Gradient $\nabla \sigma_t^2(\boldsymbol{\theta})$ die Gestalt

$$\nabla \sigma_t^2(\boldsymbol{\theta}) := \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = (1, \sigma_{t-1}^2, \dots, \sigma_{t-p}^2, r_{t-1}^2, \dots, r_{t-q}^2)^\top$$

hat.

Das asymptotische Verhalten des Schätzers $\widehat{\boldsymbol{\theta}}$ wird im folgenden Satz beschrieben.

Satz 1.7 (Konsistenz und asymptotische Normalität des GARCH-Schätzers):

Vorausgesetzt, es existieren auch die ersten vier Momente des Residuenprozesses $(z_t)_{t \in \mathbb{T}}$, so stellt sich heraus, dass der resultierende bedingte ML-Schätzer $\widehat{\boldsymbol{\theta}}$ für die GARCH-Parameter konsistent und asymptotisch normalverteilt ist mit der Konvergenz

$$\sqrt{T} \cdot (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I(\boldsymbol{\theta})^{-1}), \quad T \rightarrow \infty,$$

in Verteilung, wobei $I(\boldsymbol{\theta}) := -\mathbb{E}[H(\boldsymbol{\theta})]$ die Fisher-Informationsmatrix ist.⁵

In Kapitel 3 werden wir im Zusammenhang mit Copula-GARCH-Modellen ein mehrstufiges ML-Schätzverfahren vorstellen, bei dem in einem eigenen Schritt die GARCH-Parameter zu schätzen sein werden und in einem zweiten Schritt die Parameter einer Copula (siehe Kapitel 2).

⁵ $H(\boldsymbol{\theta})$ ist dabei die Hesse-Matrix von $\ell(\boldsymbol{\theta})$, vergleiche [17], S. 89ff.

1.2.3. Schwachpunkte und Erweiterungen

Der Vorteil von GARCH-Modellen liegt – zumindest der ARCH-Klasse gegenüber – besonders in der Sparsamkeit bei der Parameterdimension von $\boldsymbol{\theta} = (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)^\top$, womit sich vor allem der Schätzaufwand im Vergleich zum ARCH-Modell enorm reduziert. Das am häufigsten verwendete GARCH(1,1)-Modell etwa, mit den insgesamt drei Parametern $\omega, \alpha_1, \beta_1$, genügt meist schon, um Charakteristika von Finanzzeitreihen nachzubilden, für die im ARCH-Modellkontext weitaus höhere Grade (und damit weitaus mehr Parameter) erforderlich wären. Außerdem können mit GARCH viel einfacher die Cluster-Effekte der Volatilität nachgebildet werden.

Doch GARCH-Modelle haben auch ein paar Schwächen. So können die klassischen Versionen unter ihnen zum Beispiel nicht die *Volatilitätsasymmetrie* (*Leverage-Effekt*, siehe Stylized Fact (5.)), Unterabschnitt 1.1.3) nachbilden, also die Eigenschaft, dass positive und negative Shocks unterschiedliche Auswirkungen auf die Schwankungen haben. Auch die Analyse von höheren Momenten bedarf noch weiterer Parameterrestriktionen, was stark einschränkend ist und bei höheren Graden leicht zu Komplikationen führen kann. Ferner mag man mit ihnen zwar Volatilitäten auf mechanische Art und Weise adäquat rekonstruieren können, allerdings geben uns GARCH-Modelle freilich noch keine Erklärung für die Ursachen des Zustandekommens solcher Phänomene.

Viele Versuche, die Schwächen der GARCH-Modelle auszubügeln, beginnen zumeist bei dem Ansatz, bereits vorhandene Modellvarianten noch weiter auszubauen. Stellvertretend seien dafür unter anderem folgende Kategorien genannt:

- **Asymmetrische GARCH-Modelle**

Bei *asymmetrischen GARCH-Modellen* versucht man den Effekt, den die Shocks auf die Volatilität ausüben (den ARCH-Teil), in zwei unterschiedlich gewichtete Anteile, nämlich die der positiven und die der negativen Shocks, zu spalten. Die unmittelbare Folge ist, dass man den *Leverage-Effekt* nun auf sehr simple Art und Weise nachbilden kann. Meist werden negative Shocks höher gewichtet als positive, so dass Ausschläge nach unten deutlich stärker und tendenziell häufiger vorkommen als Ausschläge nach oben. Dies ist empirisch schon in sehr vielen Finanzzeitreihen bekräftigt worden.

Beispiele für solche GARCH-Erweiterungen sind das EGARCH-Modell von Nelson (1991), das TGARCH-Modell von Zakoian (1994), das GJR-GARCH-Modell von Glosten, Jagannathan und Runkle (1993) sowie weitere. Ausführliches zu diesen Modellen sowie der Theorie über das Testen auf Asymmetrie-Effekte findet man auch in [17].

- **GARCH-Modelle mit Volatilitätssprüngen**

Die Erfahrung hat gezeigt, dass die meisten Finanzzeitreihen schlagartig auf Nachrichten, formal auf eine Zugewinnung an Information, reagieren können. Solche plötzlichen Schwankungen können nur mit Modellen nachgebildet werden, die Sprünge – sei es nach oben oder nach unten – unterstützen. Hierfür werden gewöhnliche GARCH-Modelle mit sogenannten *Sprung-Modellen* (englisch *jump models*) kombiniert. Erwähnt seien als Beispiele unter anderem Beiträge von Jorion (1989), Vlaar und Palm (1993) und das noch recht junge GARJI-Modell von Maheu und McCurdy (2004), bei welchem es sich in Grundzügen um eine komplizierte Erweiterung von GARCH mit einer zeitvariierenden Sprungintensitätskomponente handelt.

- **Schwache GARCH-Modelle**

Eine empirische Studie von Nijman und Palm (1990) hat hervorgebracht, dass GARCH-Prozesse nicht temporal aggregierbar sind, das heißt, man kann weder die Schätzergebnisse noch das Modell überhaupt ohne Weiteres von hochfrequentierten (beispielsweise täglichen) Daten auf Daten mit geringerer Frequenz übertragen. Mit anderen (mathematischen) Worten: Die Klasse der gewöhnlichen GARCH-Modelle ist nicht abgeschlossen unter temporaler Aggregation. Zu diesem Zweck haben Nijman und Droste (1993) die Klasse der sogenannten schwachen GARCH-Modelle entwickelt, die diese Abgeschlossenheit als Eigenschaft aufweisen.

Neben den aufgezählten gibt es noch weitere Entwicklungen, von denen ein Großteil (auch obige Referenzen) in [17] oder [36] beschrieben wird. Auch folgender Ansatz kann dort in ausführlicherer Version nachgeschlagen werden.

- **Stochastische Volatilität**

Manche Modelle setzen noch eine Ebene weiter oben an: Neben Modellen, die die Volatilitätsgleichung in Form einer exakten deterministischen Funktion bilden, wie GARCH und obige Erweiterungen, gibt es auch Modelle mit zusätzlichen Zufallskomponenten. Zum letzten Typ gehören beispielsweise *Modelle der stochastischen Volatilität*, sogenannte *SV-Modelle*, die von S.J. Taylor fast zeitgleich zu den erstgenannten Typen (1982, 1986) eingeführt wurden. Darin wird neben den Residuen z_t aus der bekannten GARCH-Renditegleichung (1.3) eine weitere zufällige Fehlerkomponente in die Volatilität selbst eingebaut, so dass sie nur durch einen latenten Prozess beschrieben werden kann. Diese Betrachtung ist oft sehr sinnvoll, weshalb solche Modelle eine angemessene Alternative zu den klassischen Ansätzen darstellen, die aber auf Kosten des Schätzaufwands gehen kann, der sich natürlich mit jeder weiteren Zufallskomponente erhöht. Doch im Fortverlauf dieser Arbeit wollen wir nicht weiter auf die stochastische Volatilität eingehen – ausführlicher dargestellt wird sie in bereits erwähnten Werken.

1.3. Multivariate Volatilitätsmodelle: MGARCH

1.3.1. Ausgangssituation

Im multivariaten Fall liegen uns mehrere simultane Beobachtungen von Renditen $r_{t,1}, \dots, r_{t,n}$ vor, die komponentenweise nur minimal seriell und komponentenübergreifend nur minimal kreuzkorreliert sind. Beide Arten von Korrelation sind für die Absolutbeträge oder Quadrate der Renditeprozesse dagegen wesentlich höher (siehe Stylized Facts (1.) und (6.), Unterabschnitt 1.1.3).

Daher sind die Zusammenhänge zwischen quadrierten Renditezeitreihen sowie die Auswirkungen der Volatilität der einen Komponentenzeitreihe auf andere Komponentenzeitreihen (selten ist hierbei auch von *Ko-Volatilität* die Rede) von viel größerem Interesse. Hierfür brauchen wir nun multivariate bedingt-heteroskedastische Modelle, von denen einige nachfolgend vorgestellt werden sollen, aufbauend auf der klassischen GARCH-Idee im Univariaten aus Abschnitt 1.2; in Kapitel 3 werden wir schließlich eine eigene multivariate GARCH-Modellklasse einführen, die auf dem Wesen und Konzept der in Kapitel 2 vorgestellten Copula-Funktionen basiert.

Definition 1.8 (MGARCH-Modelle):

Multivariate GARCH-Modelle (kurz MGARCH) für n zeitgleich beobachtete Renditen, zusammengefasst in der Vektorzeitreihe $(\mathbf{r}_t)_{t \in \mathbb{T}} = (r_{t,1}, \dots, r_{t,n})_{t \in \mathbb{T}}^\top$, bestehen aus

(a) der Renditegleichung

$$\mathbf{r}_t = \Sigma_t^{1/2} \mathbf{z}_t, \quad t \in \mathbb{T}, \quad (1.9)$$

und

(b) einer Spezifikation der *bedingten Kovarianzmatrix* $\Sigma_t = \text{Cov}(\mathbf{r}_t | \mathcal{F}_{t-1}) = (\sigma_{t,ij})_{1 \leq i, j \leq n}$.

Dabei sei \mathbf{z}_t das multidimensionale Pendant zum unabhängigen und identisch verteilten weißen Rauschen z_t aus dem univariaten Fall.

Bemerkungen:

- Wir wollen für eine möglichst allgemeine Darstellung von nun an mit $\mathbb{I} = \{1, \dots, n\}$ die *Zeitreihenindex-* oder *Komponentenindexmenge* der n -dimensionalen Zeitreihe $(\mathbf{r}_t)_{t \in \mathbb{T}}$ bezeichnen.
- Die Matrix $\Sigma_t^{1/2}$ in obiger Gleichung ist die *Matrixwurzel* von Σ_t mit $\Sigma_t := \Sigma_t^{1/2} \cdot \Sigma_t^{1/2}$.
- In der Literatur herrscht bei der Bezeichnung „Volatilitätsmatrix“ ein wenig Uneinigkeit: manche Autoren sprechen bereits bei der gesamten bedingten Kovarianzmatrix Σ_t davon, während andere damit (konsequenterweise) lediglich die Diagonal-

elemente meinen, also nur die bedingten Varianzen, zusammengefasst in der Matrix $\text{diag}(\sigma_{t,11}, \dots, \sigma_{t,nn})$. Allgemeiner möchten wir daher auf den Begriff der *Kovolatilitätsmatrix* ausweichen.

Was nun fehlt, ist die Spezifikation einer Modellgleichung für Σ_t , die stets berücksichtigt, dass diese symmetrisch und positiv definit bleibt. In der Art und Weise eben dieser Spezifikation unterscheiden sich nun auch vor allem die meisten bis dato entwickelten MGARCH-Modellvarianten.

1.3.2. Modellvarianten: Überblick

In diesem Unterabschnitt stellen wir zwei Kategorien von Modellklassen vor:

- (1.) In der ersten wird die Kovolatilitätsmatrix Σ_t direkt modelliert. Darin enthalten sind allen voran die Modellklassen VEC und BEKK, welche als Pioniere in der multivariaten Finanzzeitreihenmodellierung gelten, enorm rechenaufwendig werden können und dennoch sehr häufig verwendet werden, da sie die realen Sachverhalte hervorragend nachbilden können.
- (2.) In der zweiten Kategorie befinden sich sogenannte *Conditional-Correlation*-Modelle (kurz CC-Modelle), die anstatt der direkten Modellierung von Σ_t bevorzugt auf einer Zerlegung $\Sigma_t = D_t R_t D_t^\top$ aus zeitvariierender bedingter Korrelationsmatrix R_t sowie der Diagonalmatrix

$$D_t := \text{diag}(\sqrt{\sigma_{t,11}}, \dots, \sqrt{\sigma_{t,nn}}) \quad (1.10)$$

aufbauen.

Modellgleichungen für Σ_t oder R_t direkt in einer Matrixdarstellung anzugeben, ist zwar möglich, aber oftmals relativ unübersichtlich und die Implementierung ist noch dazu enorm rechenintensiv. Daher wird häufig zunächst eine *Reparametrisierung* der symmetrischen Matrix Σ_t in Kategorie (1.) beziehungsweise R_t in Kategorie (2.) durchgeführt, bei der alle relevanten Komponenten in einem Vektor $\boldsymbol{\varsigma}_t$ gestapelt werden. Wegen der Symmetrie genügen so ohne Einschränkung diejenigen unterhalb der Diagonalen, zusätzlich zu denen entlang der Diagonalen. Der Vektor $\boldsymbol{\varsigma}_t$ wird anschließend einer einfachen vektoriellen Erweiterung des GARCH-Modells unterworfen.

In Kategorie (1.) erhalten wir so beispielsweise den gestapelten Vektor

$$\boldsymbol{\varsigma}_t = \text{st}_{\text{unten}}^{\text{diag}}(\Sigma_t) := (\sigma_{t,11}, \\ \sigma_{t,21}, \sigma_{t,22}, \\ \sigma_{t,31}, \sigma_{t,32}, \sigma_{t,33}, \\ \vdots \\ \sigma_{t,n1}, \dots, \sigma_{t,nn})^\top$$

aller Komponenten von Σ_t auf und unterhalb der Diagonalen.

Dagegen setzt sich in Kategorie (2.) der Vektor

$$\boldsymbol{\varsigma}_t = (\sqrt{\sigma_{t,11}}, \dots, \sqrt{\sigma_{t,nn}}, \boldsymbol{\rho}_t^\top)^\top$$

aus den bedingten Standardabweichungen und dem gestapelten Vektor

$$\boldsymbol{\rho}_t := \text{st}_{\text{unten}}(R_t) := (\rho_{t,21}, \\ \rho_{t,31}, \rho_{t,32}, \\ \vdots \\ \rho_{t,n1}, \dots, \rho_{t,nn-1})^\top$$

aller Komponenten von R_t unterhalb der Diagonalen zusammen. Die Diagonalelemente (das heißt die Korrelationen der Komponenten zu sich selbst) sind ohnehin immer konstant Eins und müssen deshalb nicht geschätzt werden.

Zwar hat $\boldsymbol{\varsigma}_t$ mit insgesamt $\frac{n(n+1)}{2}$ Elementen eine überschaubarere Form als mit manchen anderen Zerlegungen von Σ_t , doch auch diese bläht sich bei hohen Dimensionen recht schnell auf. Als Alternative wird auch oft die Cholesky-Zerlegung verwendet, die sogar ohne weitere (hier noch notwendige) Beschränkungen an die Komponenten die positive Definitheit der jeweiligen Matrix garantieren kann.⁶

In den Absätzen im Anschluss wird der Vollständigkeit halber jeweils kurz der zu den Modellkategorien gehörige Aufbau angegeben und auf Vorteile oder Schwachpunkt hingewiesen.⁷

⁶Nähere Details zur Reparametrisierung illustriert zum Beispiel [36].

⁷Diese und weitere MGARCH-Varianten werden in einer übersichtlichen Darstellung von [34] ausführlicher erläutert.

Modelle für die Kovolatilität Σ_t : VEC- und BEKK-GARCH

Das VEC-GARCH-Modell (von Bollerslev, Engle, Wooldridge, 1988) versucht das univariate GARCH-Modell aus Abschnitt 1.2 so allgemein wie möglich auf das Höherdimensionale zu erweitern. Die Kovolatilitätsgleichung der reparametrisierten Matrix Σ_t hat hierbei die Form

$$\varsigma_t = \boldsymbol{\omega} + \sum_{i=1}^p A_i \text{st}_{\text{unten}}^{\text{diag}}(\mathbf{r}_{t-i} \mathbf{r}_{t-i}^{\top}) + \sum_{j=1}^q B_j \varsigma_{t-j}. \quad (1.11)$$

Dabei wird diese generelle Darstellung für $n = 1$ wieder vollständig in die univariate Volatilitätsgleichung des GARCH(p, q)-Modells überführt, wenn wir $\boldsymbol{\omega}$, A_i und B_j als höherdimensionale Analoga zu den ω , α_i sowie β_j auffassen (vergleiche Gleichung (1.4) aus Definition 1.6).

Die Nachteile dieses Modells liegen unter anderem im rapiden Anstieg der Parameteranzahl. Diese beträgt allgemein $(p+q)\left(\frac{n(n+1)}{2}\right)^2 + \frac{n(n+1)}{2}$. Doch auch Parameterrestriktionen sind zu berücksichtigen, um die positive Definitheit von Σ_t am Ende zu gewährleisten. Somit ist auch die Parameterschätzung als Optimierungsverfahren sehr komplex und rechenlastig, obwohl (oder gerade weil) diese Modellklasse die wohl größte Flexibilität für die Modellierung von multivariaten Finanzzeitreihen bieten kann.

Das BEKK-Modell (von Baba, Engle, Kraft, Kroner, 1995) ist eine leichte Vereinfachung (und damit gleichzeitig eine Einschränkung) der obigen VEC-GARCH-Variante, bei der die positive Definitheit bereits in der Konstruktion der Kovolatilitätsgleichung (durch Multiplikation der Parameter-Matrizen *und* ihren Transponierten) vorgegeben wird.

Modelle für die bedingte Korrelation R_t : CC-GARCH

Bei *Constant-Conditional-Correlation*-GARCH-Modellen (kurz CCC-GARCH) variiert nur die Diagonalmatrix D_t aus Gleichung (1.10) mit der Zeit, da wir $R_t \equiv R$ als zeitkonstante positiv definite Matrix auffassen wollen. Daraus erhalten wir dann genau n univariate GARCH-Marginalmodelle für die Diagonalelemente nach dem Schema, das bereits in Abschnitt 1.2 vorgestellt wurde.

Diese Herangehensweise ist allerdings sehr restriktiv und bildet die Phänomene, welche finanzwirtschaftliche Zeitreihen auch wechselseitig verursachen, nicht sehr realistisch nach. Doch sie wird häufig als erster Schritt benutzt, um daran in anschließenden Schritten weitere Analysen und Modelle anknüpfen zu können.

Das *Dynamic-Conditional-Correlation*-GARCH-Modell (kurz DCC-GARCH) verallgemeinert obiges CCC-Konzept um die Zeitdynamik der Matrix R_t derart, dass Schätzungen in einzelnen Etappen aufgrund univariater GARCH-Marginalmodelle immer noch mög-

lich sind.⁸ Einen Vorschlag hierfür geben etwa [37] mit der Einführung des sogenannten *Varying-Correlation*-Modells (kurz *VC* oder *VC-GARCH*), bei dem die bedingte Korrelationsmatrix gemäß ARMA- oder GARCH-ähnlicher Modelle abgebildet wird.

In jeder sogenannten *Korrelationsgleichung* der DCC-GARCH-Modelle beträgt die Anzahl der zu schätzenden Parameter im Allgemeinen $\frac{n(n+1)}{2} + (n+1)(p+q)$. Als Schwäche wird beispielsweise die Restriktion angesehen, dass alle Korrelationen der selben Dynamik folgen, ein besonderer Fall, der in der Praxis natürlich eher selten eintritt.

Eine ausführlichere Gegenüberstellung vieler gängiger, aber auch „exotischer“ MGARCH-Modelle und -Erweiterungen wird beispielsweise in [23] sowie [34] illustriert.

1.4. Überblick

Wir haben in diesem Kapitel unser Hauptaugenmerk auf Eigenschaften von Zeitreihen gelegt, die auf Finanzmärkten zu beobachten sind (Stylized Facts). Neben den Objekten des Interesses, den Renditen, haben wir die Shocks als zufällige unabhängige Komponenten sowie die Volatilität als wichtigste Bestandteile eines Finanzzeitreihenmodells kennengelernt.

Die Mehrheit der Finanzmodelle hat anfangs die bedingte Heteroskedastizität nicht berücksichtigt. Erst später in den 1980er Jahren gelangen erste Versuche, die zeitliche Variation und weitere besonderen Merkmale der Volatilität in das Modell zu integrieren, allen voran die GARCH-Modelle, mit denen wir uns in Grundzügen aus theoretischer und praktischer Sicht sowohl im Univariaten als auch im Multivariaten befasst haben. Seit deren Einführung wurden mittlerweile eine Vielzahl von Erweiterungen entwickelt, für unterschiedlichste Zwecke verwendet und mit Hilfe dieser eine Reihe weiterer Phänomene und Effekte beschrieben, die finanzwirtschaftliche Zeitreihen von anderen hervorheben, jedoch mit den vormaligen Modellen nicht genügend nachgestellt werden konnten.

Im Multivariaten haben wir uns mehrere MGARCH-Varianten angesehen. Die wohl offensichtlichste Anwendung multivariater GARCH-Modelle ist das Studieren von Zusammenhängen zwischen den Volatilitäten und Kovolatilitäten, die mehrere Finanztitel auf einem Markt zusammen im Zeitverlauf beeinflussen. Aber auch zwischen mehreren Märkten kann man nach Wechselwirkungen suchen und forschen: erhöht ein Shock großen Betrags auf einem Markt die Volatilität auf einem anderen Markt und, wenn ja, in welchem Ausmaß und mit welchem Vorzeichen?

⁸Ausführlicheres zu dieser Modellklasse gibt es in [7].

Mit diesen Fragen beschäftigen sich die Modelle zwar nicht direkt, aber mittelbar können sie Hinweise und Andeutungen auf die Ursachen aufdecken und ermöglichen so auch Vorhersagen. Dies ist vor allem im Hinblick auf die Entscheidungsfindung von großer Bedeutung wie etwa im Asset Pricing, Option Pricing oder in der Portfolio-Selektion. Besondere Bedeutung kommt den bedingt-heteroskedastischen Modellen auch im quantitativen Risikomanagement zu, da sie gerade bei der Abschätzung des Marktrisikos eines Finanztitels (beispielsweise durch den Value-at-Risk als Maßzahl) direkt die Volatilität als zu verarbeitende Komponente verarbeiten.

Ziel dieser Arbeit ist die Vorstellung einer eigenen neuartigen, multivariaten Modellklasse, die die ausgereiften Fähigkeiten von GARCH-Modellen für univariate finanzökonometrische Phänomene ausnutzt und das äußerst vielseitige Konzept der im nun folgenden Kapitel vorgestellten Copula-Funktionen als „Verbindung“ zu einem multivariaten Modell vereinigt.

Kapitel 2

Copula-Funktionen

Hat man für eine Reihe von Beobachtungen X_1, \dots, X_n (beispielsweise Renditen oder Risiken in einem Portfolio) alle univariaten Randverteilungen sowie deren lineare Korrelationsmatrix vorliegen, so kann man bei Normalverteilungsannahmen schnell die multivariate gemeinsame Verteilung aus diesen gegebenen Informationen konstruieren. Doch da viele Modelle in der finanzökonomischen Praxis mittlerweile nicht mehr von dieser vereinfachten Situation ausgehen, bedarf es eines allgemeineren Konzepts zur *Verknüpfung von univariaten Randverteilungen* und der formalen *Abhängigkeitsstruktur*. Hätte man diesen Baustein zur Hand, ließen sich mittels beliebiger Randverteilungen auch beliebige gemeinsame Verteilungsfunktionen generieren, beispielsweise auch ohne dass alle Marginalverteilungen vom selben Typ sein müssen. Umgekehrt könnte man die multivariate Verteilung aufspalten in seine univariaten Randbestandteile sowie die eben genannte Abhängigkeitsstruktur.

Die Lösung dieser Problemstellungen verspricht das Konzept der sogenannten *Copulas*. Mit Hilfe dieser lässt sich tatsächlich die multivariate Verteilung als funktionalen Ausdruck der univariaten Randverteilungen darstellen. Eine *Copula-Funktion* bildet genau diese funktionale Verknüpfung und den Baustein der formalen Abhängigkeitsstruktur, weshalb diese auch *dependence functions* oder *functional dependence* genannt wurden. Der Begriff „Copula“¹ selbst wurde in den 1950er Jahren vom Mathematiker A. Sklar ([35]) geprägt, obwohl diese unter obigen Bezeichnungen bereits ein Jahrzehnt früher in Arbeiten von W. Hoeffding beschrieben wurden. Im Laufe der nachfolgenden Jahre gewann diese Funktionenklasse ab etwa 1990 im Zusammenhang mit den revolutionären Entwicklungen von Finanzmarktmodellen (vergleiche Kapitel 1) zunehmend an Beliebtheit und Bedeutung. Spätestens nach [5] haben copula-basierte Zusammenhangsmaße die lineare Korrelation als Standardmaß für die Abhängigkeit in vielen Bereichen abgelöst und verliehen damit dem Copula-Konzept bei der Anwendung im Risiko-Management noch größeres Gewicht.

Dieses Kapitel beginnt in Abschnitt 2.1 mit einem kurzen Einstieg in wichtige Zusammenhänge zwischen Zufallsvariablen und deren Verteilungsfunktionen. Anschließend werden wir uns in Abschnitt 2.2 mit der Einführung in die Theorie der Copulas und mit zen-

¹lat. *copula*, „Verbindung“ oder „Verknüpfung“

tralen Aussagen beschäftigen, überwiegend angelehnt am Standardwerk [25], bevor wir Copulas im Abschnitt 2.3 aus einem praktischen Blickwinkel beleuchten. Dort werden die Simulationsmöglichkeiten erläutert und verschiedene Vorgehensweisen in der Maximum-Likelihood-Schätzung besonders ausführlich beschrieben, bei welcher wir nicht näher bestimmte Zufallsvektoren – zunächst zeitkonstant, danach als multivariate Zeitreihe – zugrundelegen werden. Vorrangiges Ziel dieses Kapitels ist nämlich eine Vorbereitung auf das Hauptthema in Kapitel 3 mit allen nötigen theoretischen und praktischen Hilfsmitteln. Als Abschluss versucht Abschnitt 2.4 einen Ausblick in tiefergehende und hier nicht näher behandelte Themengebiete zu liefern.

2.1. Motivation

Im gesamten Verlauf dieser Arbeit gehen wir stets stillschweigend von einem nicht näher beschriebenen Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) aus. Der Einfachheit halber werden wir fast immer Aussagen zu einem zweidimensionalen Zufallsvektor $(X, Y)^\top : \Omega \rightarrow \mathbb{R}^2$ mit Verteilungsfunktion H_{XY} sowie Randverteilungen F_X beziehungsweise F_Y treffen. Die Ergebnisse sind dann analog beziehungsweise induktiv auf Dimension $n > 2$ übertragbar.

Definieren wir zunächst den Begriff der Quasi-Inversen als Verallgemeinerung der Inversen von Verteilungsfunktionen.

Definition 2.1 (Quasi-Inverse):

Für eine Zufallsvariable X mit Verteilung F_X definieren wir mit

$$F_X^{(-1)}(y) := \inf\{x \in \mathbb{R} : F_X(x) \geq y\}$$

die sogenannte *Quasi-Inverse* der Funktion F_X .

Bemerkungen:

- Die Quasi-Inverse aus obiger Definition ist in der Statistik auch bekannt als Quantilsfunktion der entsprechenden Verteilung.
- Da eine Verteilungsfunktion nur rechtsseitig stetig und im Allgemeinen nur monoton, und nicht streng monoton wachsend ist, bildet der Begriff der Quasi-Inversen in obiger Definition eine Verallgemeinerung des Begriffs der Inversen. Im Falle strenger Monotonie gilt insbesondere die Invertierbarkeit von F_X , damit also $F^{(-1)} \equiv F^{-1}$.

Bevor wir uns mit dem Copula-Konzept auseinandersetzen, legen wir einen interessanten und für praktische Zwecke äußerst nützlichen Zusammenhang zwischen einer Zufallsvariable und ihrer Verteilung dar:

Satz 2.2 (Fisher, [11]):

Für eine stetige Zufallsvariable X und deren Verteilungsfunktion F_X ist die neue Zufallsvariable

$$F_X(X) := F_X \circ X : \Omega \rightarrow [0, 1] \quad (2.1)$$

uniform verteilt auf dem Intervall $[0, 1]$ (wir schreiben $F_X(X) \sim \mathcal{U}[0, 1]$), denn es gilt

$$P(F_X(X) \leq u) = P(X \leq F_X^{(-1)}(u)) = F_X(F_X^{(-1)}(u)) = u$$

mit $u \in [0, 1]$.

Diese Transformation (im Englischen ist hier von *probability integral transform* die Rede) hat eine entscheidende praktische Bedeutung: Die uniform verteilten Zufallsvariablen $U = F_X(X)$ und $V = F_Y(Y)$ enthalten keine Information mehr über die univariaten Verteilungen. Das bedeutet, die restliche Information, die noch in der gemeinsamen Verteilung H_{UV} von $(U, V)^\top$ „steckt“, enthält nur noch Daten über die Art der Abhängigkeit zwischen U und V . Dieses Sachverhalts bedient sich nun das theoretische Konzept der Copula-Funktionen.

2.2. Theoretisches Konzept der Copula-Funktionen

2.2.1. Definitionen

In seinen mathematischen Untersuchungen des funktionalen Zusammenhangs zwischen gemeinsamer Verteilung und den Marginalien eines Zufallsvektors hatte A. Sklar in den 1950er Jahren eine neue Klasse von Funktionen verwendet, die damals als *Copula-Funktionen* bekannt geworden sind. In diesem Unterabschnitt werden lediglich die wesentlichen Begrifflichkeiten sowie zentralen Ergebnisse der Theorie der Copula-Funktionen zusammengefasst, wobei der an tiefergehende Details interessierte Leser zum Beispiel auf das Einsteigsbuch [25] verwiesen sei.

Eine formale Definition der Copula-Funktion ist nun angebracht. Zuerst wird die allgemeinste (n -dimensionale) Version formuliert, doch aus Gründen der Einfachheit beschränken wir uns im anschließenden Kontext auf Dimension $n = 2$. Erst bei der Parameterschätzung betrachten wir wieder den allgemeinen Fall.

Definition 2.3 (*n*-Copula):

Wir nennen eine Funktion $C : [0, 1]^n \rightarrow [0, 1]$ eine *n*-Copula, wenn sie die folgenden Eigenschaften besitzt:

- (1.) Für alle $\mathbf{u} \in [0, 1]^n$ mit mindestens einer Komponente $u_i = 0$ gilt $C(\mathbf{u}) = 0$, $1 \leq i \leq n$.
- (2.) Für alle $\mathbf{u} \in [0, 1]^n$ mit $u_i = 1$ für alle $1 \leq i \leq n$ bis auf eine Komponente u_j gilt $C(\mathbf{u}) = u_j$, $1 \leq j \leq n$, $j \neq i$.
- (3.) Für alle $\mathbf{u}, \mathbf{v} \in [0, 1]^n$ mit $\mathbf{u} \leq \mathbf{v}$ (komponentenweise) gilt

$$V_C([\mathbf{u}, \mathbf{v}]) := \sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1+\dots+i_n} \cdot C(w_{1,i_1}, \dots, w_{n,i_n}) \geq 0, \quad (2.2)$$

wobei mit $w_{j,1} = u_j$ und $w_{j,2} = v_j$, $1 \leq j \leq n$, die insgesamt 2^n Eckpunkte des n -dimensionalen Hyper-Quaders $[\mathbf{u}, \mathbf{v}] \subset [0, 1]^n$ gemeint sind.

Copulas ermöglichen es, den eindimensionalen (Rand-)Verteilungen F_1, \dots, F_n von Zufallsvariablen X_1, \dots, X_n eine gemeinsame Verteilungsfunktion $H_{\mathbf{X}}$ des Zufallsvektors $\mathbf{X} := (X_1, \dots, X_n)^\top$ zuzuordnen. Dies ist zwar recht heuristisch formuliert, aber dass diese Funktionen tatsächlich einen solchen Zusammenhang herstellen können, ist einerseits nach Konstruktion recht einfach einzusehen, andererseits erklärt der zentrale Satz von Sklar im kommenden Unterabschnitt 2.2.2 genau diesen Sachverhalt und gibt uns sogar ein Eindeutigkeitskriterium.

Anstatt die zentralen Aussagen allgemein zu formulieren, werden wir nun die speziellen *2-Copulas* einführen (fortan einfach *Copulas*) und mit diesen die weiteren Besonderheiten von Copula-Funktionen erklären. Es ist zwar bei den meisten Aussagen möglich, analoge Schlussfolgerungen auf höhere Dimensionen zu übertragen (so gilt etwa für $n \geq 3$ und alle $2 \leq k \leq n$, dass jede k -dimensionale Randverteilung einer n -Copula wiederum eine k -Copula ist), doch auf Sonderfälle in der Dimension $n = 2$ wird nicht mehr explizit hingewiesen.²

Definition 2.4 (2-Copula):

Wir nennen eine Funktion $C : [0, 1]^2 \rightarrow [0, 1]$ eine *2-Copula*, wenn sie die Eigenschaften

- (1.) $C(u, 0) = 0 = C(0, v)$,
- (2.) $C(u, 1) = u$ und $C(1, v) = v$,
- (3.) $C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0$

für alle $u, u_1 \leq u_2$ und $v, v_1 \leq v_2$ aus $[0, 1]$ erfüllt.

²Für eine ausführlichere Behandlung siehe [25].

2.2.2. Zentrale Aussagen

Wie im letzten Abschnitt bereits angekündigt, liefert folgender Satz von Sklar die theoretische Grundlage für die praktische Anwendung von Copula-Funktionen.

Satz 2.5 (Sklars Theorem):

Seien X und Y zwei Zufallsvariablen mit Verteilungen F_X beziehungsweise F_Y sowie gemeinsamer Verteilung H_{XY} . Dann existiert eine Copula C so, dass

$$H_{XY}(x, y) = C(F_X(x), F_Y(y)) \quad (2.3)$$

für alle $x, y \in \mathbb{R}$ gilt. Die Copula ist genau dann eindeutig bestimmt, wenn F_X und F_Y beide stetige Verteilungsfunktionen sind.

Der Blickwinkel von Sklars Theorem lässt sich auch umkehren: Falls C eine Copula ist und F_X sowie F_Y die Verteilungsfunktionen von X beziehungsweise Y sind, dann ist jede Funktion H , die Gleichung (2.3) erfüllt, eine gemeinsame Verteilung von X und Y mit Marginalverteilungen F_X und F_Y und wir setzen $H := H_{XY}$.

Folgerung 2.6:

Als direkte Konsequenz und Alternativformulierung des Satzes 2.5 ergibt sich die Darstellung

$$C(u_1, u_2) = H_{XY}\left(F_X^{(-1)}(u_1), F_Y^{(-1)}(u_2)\right) \quad (2.4)$$

mit $u_1, u_2 \in [0, 1]$ und Quasi-Inversen $F_X^{(-1)}$ von F_X beziehungsweise $F_Y^{(-1)}$ von F_Y .

Welche praktischen Konsequenzen nun Sklars Theorem für uns hat, erkennt man sofort: Einerseits lassen sich gemäß Gleichung (2.3) durch die Wahl beliebiger univariater Verteilungen und einer beliebigen Copula-Funktion alle erdenklichen multivariaten Verteilungen völlig willkürlich zusammenbauen. Andererseits können wir nach Gleichung (2.4) die in einer gegebenen multivariaten Verteilung H_{XY} enthaltene Information über die Abhängigkeitsstruktur in Form der Copula-Funktion C extrahieren, nämlich durch Einsetzen der Quasi-Inversen der univariaten Marginalien F_X und F_Y .

Bemerkung (Copula als Verteilungsfunktion):

Eine Copula-Funktion C kann selbst als eine multivariate Verteilungsfunktion aufgefasst werden, nämlich als gemeinsame Verteilung zweier auf $[0, 1]$ gleichverteilter Zufallsvariablen U und V . Denn dann gilt $H_{UV}(u, v) = P(U \leq u \wedge V \leq v) = C(u, v)$, das heißt $C \equiv H_{UV}$.

Im nun folgenden Unterabschnitt sind einige der übrigen wichtigen Eigenschaften von Copulas, auch hinsichtlich späterer Kapitel dieser Arbeit, übersichtlich und in Kürze dargestellt.

2.2.3. Weitere Eigenschaften

Weitere für unseren Fortverlauf wichtige Merkmale von Copula-Funktionen sind ihre Differenzierbarkeit und die Zerlegbarkeit in absolutstetige und singuläre Komponente. Im Zuge dessen werden wir den Begriff „Copula-Dichte“ einführen. Abschließend übertragen wir die bisherigen theoretischen Aussagen auf die bedingte Version des Copula-Konzepts.

Treffen wir zuerst kurze Vorbereitungen, um danach sogenannte *Copula-Dichte* einzuführen.

Satz 2.7 (Differenzierbarkeit):

Eine Copula C ist *fast sicher partiell differenzierbar*, das heißt die partiellen Ableitungen $C_u(u, v) := \frac{\partial}{\partial u}C(u, v)$ und $C_v(u, v) := \frac{\partial}{\partial v}C(u, v)$ existieren fast sicher und es gilt stets

$$\mathbf{0} \leq \nabla C(u, v) \leq \mathbf{1} \quad (2.5)$$

komponentenweise für den Gradienten $\nabla C := (C_u, C_v)^\top$.

Da eine Copula C nach obiger Bemerkung als gemeinsame Verteilung zweier Zufallsvariablen $U, V \sim \mathcal{U}[0, 1]$ aufgefasst werden kann, ergibt sich für die partiellen Ableitungen

$$C_u(u, v) = P(V \leq v | U = u),$$

beziehungsweise analog

$$C_v(u, v) = P(U \leq u | V = v).$$

Es handelt sich also um die bedingte Verteilung von V , gegeben $U = u$, respektive die von U , bedingt auf $V = v$.

Satz 2.8 (Absolutstetige und singuläre Komponente):

Jede Copula-Funktion C besitzt die Zerlegung in die Summe

$$C \equiv A_C + S_C$$

aus *absolutstetiger Komponente*

$$A_C(u, v) := \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s, t) dt ds$$

und *singulärer Komponente*

$$S_C(u, v) := C(u, v) - A_C(u, v).$$

Bemerkung:

Die beiden Komponenten A_C und S_C erfüllen keineswegs die Copula-Eigenschaften aus Definition 2.4!

Definition 2.9 (Copula-Dichte):

Falls $C \equiv A_C$, dann heißt die Copula C *absolutstetig* und der Integrand aus Satz 2.8,

$$c(u, v) := \frac{\partial^2}{\partial u \partial v} C(u, v) = \frac{\partial^2}{\partial v \partial u} C(u, v),$$

wird *Copula-Dichte von C* genannt.

Die Copula-Dichte wird meist dazu verwendet, um beispielsweise gegebene (seien es reale oder simulierte) Daten an eine Copula anzupassen. Ist die Copula C die Verteilungsfunktion von uniform verteilten Zufallsvariablen U und V , so können wir entsprechend die Copula-Dichte c aus obiger Definition als gemeinsame Verteilungsdichte von U und V auffassen.

Mit Sklars Theorem (Satz 2.5) ergibt sich entsprechend für allgemeine absolutstetige Verteilungen H_{XY} sowie F_X und F_Y :

$$\begin{aligned} h_{XY}(x, y) &= \frac{\partial^2}{\partial x \partial y} H(x, y) \\ &= \frac{\partial^2}{\partial x \partial y} C(F_X(x), F_Y(y)) \\ &= \frac{\partial^2}{\partial F_X(x) \partial F_Y(y)} C(F_X(x), F_Y(y)) \cdot \frac{\partial}{\partial x} F_X(x) \cdot \frac{\partial}{\partial y} F_Y(y) \\ &= c(F_X(x), F_Y(y)) \cdot f_X(x) \cdot f_Y(y) \end{aligned}$$

mit den Dichten h_{XY} von $(X, Y)^\top$ sowie f_X und f_Y von X beziehungsweise Y . Daraus folgt das Verhältnis

$$c(F_X(x), F_Y(y)) = \frac{h_{XY}(x, y)}{f_X(x) \cdot f_Y(y)}, \quad (2.6)$$

welches genau dann 1 wird, wenn sich die gemeinsame Dichte $h_{XY}(x, y) = f_X(x) \cdot f_Y(y)$ als Produkt der Marginaldichten darstellen lässt. Dies ist nur bei Unabhängigkeit der Zufallsvariablen X und Y der Fall.

Mittels der Maximum-Likelihood-Methode (ML) schätzt man dann die Copula-Parameter einer zuvor nach gewissen Kriterien ausgewählten Copula-Familie, wie in Abschnitt 2.3 ausführlich beschrieben wird. Wie wir aber in Kapitel 1 erfahren haben, ist im finanzökonomischen Kontext meist die *bedingte Log-Likelihood-Funktion* in der ML-Methode das Objekt der Untersuchung. Um dies auch mit Copula-Funktionen zu ermöglichen, stellen wir im nachfolgenden Unterabschnitt das bisherige Konzept, übertragen auf die bedingte Version, dar.

2.2.4. Bedingte Copula

Wie wir in Kapitel 1 erfahren haben, ist es in der ökonometrischen Modellierung von Finanzzeitreihen ein ganz wesentlicher Aspekt, dass die Zufallsvariablen des Interesses bedingt auf Realisationen aus der Vergangenheit sind. Sei im Fortverlauf dieses Unterabschnitts W die (reellwertige) Zufallsvariable, auf die (beziehungsweise deren Werte aus $W(\Omega) \subset \mathbb{R}$) bedingt werden soll. Die von W erzeugte σ -Algebra symbolisieren wir mit $\mathcal{W} := \sigma(W)$.

Definition 2.10 (Bedingte Copula):

Eine 2-dimensionale, auf W *bedingte Copula* ist eine Funktion $C : [0, 1]^2 \times W(\Omega) \rightarrow [0, 1]$, die die Eigenschaften

- (1.) $C(u, 0 | w) = 0 = C(0, v | w)$,
- (2.) $C(u, 1 | w) = u$ und $C(1, v | w) = v$,
- (3.) $C(u_1, v_1 | w) - C(u_1, v_2 | w) - C(u_2, v_1 | w) + C(u_2, v_2 | w) \geq 0$

für alle $u, u_1 \leq u_2$ und $v, v_1 \leq v_2$ aus $[0, 1]$ und alle w aus $W(\Omega)$ erfüllt. Fassen wir die gesamte Information, auf die wir bedingen wollen, in \mathcal{W} zusammen, so schreiben wir schlicht $C(u, v | \mathcal{W})$. Man beachte jedoch, dass dieser Ausdruck – strikt betrachtet – wieder eine Zufallsvariable ist!

Diese bedingte Copula können wir ebenfalls, analog zum unbedingten Fall, auffassen als bedingte Verteilung zweier auf $[0, 1]$ uniform verteilter Zufallsvariablen U und V , gegeben die Information \mathcal{W} .

Mit der in [31] vorgestellten Tatsache, dass mit der Zufallsvariable X und deren bedingten Verteilung $F_X(\cdot | \mathcal{W})$ auch die Transformation $U := F_X(X | \mathcal{W})$ eine uniform verteilte Zufallsvariable ist, können wir auch eine bedingte Version von Sklars Theorem aufstellen.

Satz 2.11 (Bedingte Version von Sklars Theorem):

Seien X und Y zwei Zufallsvariablen mit bedingten Verteilungen $F_X(\cdot | \mathcal{W})$ beziehungsweise $F_Y(\cdot | \mathcal{W})$ sowie gemeinsamer bedingter Verteilungsfunktion $H_{XY}(\cdot, \cdot | \mathcal{W})$. Dann existiert eine bedingte Copula $C(\cdot, \cdot | \mathcal{W})$ so, dass

$$H_{XY}(x, y | \mathcal{W}) = C(F_X(x | \mathcal{W}), F_Y(y | \mathcal{W}) | \mathcal{W}) \quad (2.7)$$

für alle $x, y \in \mathbb{R}$ gilt. Die bedingte Copula ist genau dann eindeutig bestimmt, wenn $F_X(\cdot | \mathcal{W})$ und $F_Y(\cdot | \mathcal{W})$ beide stetige bedingte Verteilungsfunktionen sind.

Gilt umgekehrt, dass $C(\cdot, \cdot | \mathcal{W})$ und $F_X(\cdot | \mathcal{W})$ beziehungsweise $F_Y(\cdot | \mathcal{W})$ gegeben sind, so ist jede Funktion H , die Gleichung (2.7) erfüllt, eine bedingte multivariate Verteilung für X und Y und wir setzen $H := H_{XY}(\cdot, \cdot | \mathcal{W})$.

Folgerung 2.12:

Als direkte Konsequenz und Alternativformulierung des Satzes 2.11 ergibt sich die Darstellung

$$C(u_1, u_2 | \mathcal{W}) = H_{XY} \left(F_X^{(-1)}(u_1 | \mathcal{W}), F_Y^{(-1)}(u_2 | \mathcal{W}) | \mathcal{W} \right) \quad (2.8)$$

mit $u_1, u_2 \in [0, 1]$ und Quasi-Inversen $F_X^{(-1)}(\cdot | \mathcal{W})$ und $F_Y^{(-1)}(\cdot | \mathcal{W})$ von $F_X(\cdot | \mathcal{W})$ beziehungsweise $F_Y(\cdot | \mathcal{W})$.

Die Definition und Herleitung der bedingten Copula-Dichte $c(\cdot, \cdot | \mathcal{W})$ erfolgt recht ähnlich, so dass wir dies auslassen wollen. Analog zu Gleichung (2.6) können wir im absolutstetigen Fall ebenfalls schlussfolgern:

$$c(F_X(x | \mathcal{W}), F_Y(y | \mathcal{W}) | \mathcal{W}) = \frac{h_{XY}(x, y | \mathcal{W})}{f_X(x | \mathcal{W}) \cdot f_Y(y | \mathcal{W})}.$$

Die bedingte Dichte $h_{XY}(\cdot, \cdot | \mathcal{W})$ lässt sich also als Produkt aus bedingter Copula-Dichte $c(\cdot, \cdot | \mathcal{W})$ und den bedingten Marginaldichten $f_X(\cdot | \mathcal{W})$ respektive $f_Y(\cdot | \mathcal{W})$ schreiben.

Wir haben nun viele theoretische Eigenschaften von Copula-Funktionen kennengelernt, deren Verwendungen in der Praxis jedoch nicht immer unmittelbar ersichtlich sind. Deshalb wenden wir uns nun im nächsten Abschnitt der praktischen Seite zu, indem wir die Bedeutung der Copulas in der Anwendung hervorheben.

2.3. Copula-Funktionen in der Statistik

Copula-Funktionen werden auf vielerlei Weisen und zu unterschiedlichsten Zwecken eingesetzt. Wir befassen uns in diesem Abschnitt mit der künstlichen Erzeugung von Zufallsrealisationen aus einer gegebenen Copula heraus. Dabei werden wir speziell zwei Copula-Klassen präsentieren, die sehr häufig verwendet werden, um Simulationen oder Schätzungen durchzuführen. Beide besitzen nämlich jeweils eigene nützliche Besonderheiten, die wir in Grundzügen nennen wollen. Nachdem die Simulation von Zufallszahlen behandelt wurde, widmen wir uns ausführlich der Parameterschätzung bei Copula-Funktionen.

2.3.1. Simulation von Zufallszahlen

Eine typische Vorgehensweise zu Beginn einer Simulation ist beispielsweise die zufällige Erzeugung von Realisationen zweier reellwertiger Zufallsvariablen X und Y , die gemäß

der Funktionen F_X sowie F_Y verteilt sein sollen. Geschehen die Simulationen voneinander unabhängig, so erhält man stets stochastisch unabhängige Zufallszahlen. Nun möchten wir aber gezielt die Abhängigkeitsstruktur (und damit die Copula C) festlegen. Durch deren Wahl erzeugen wir entsprechende stochastisch abhängige und damit korrelierte Realisationen des Vektors $(X, Y)^\top$, verteilt nach $C(F_X(X), F_Y(Y))$. Hierzu sei folgende, recht intuitive Vorgehensweise genannt:

Algorithmus 2.13 (Inversionsmethode):

Sei C eine willkürlich ausgewählte bivariate Copula-Funktion.

- (1.) Generiere zwei unabhängige auf $[0, 1]$ uniform verteilte Zufallszahlen u und t .
- (2.) Setze $v := C_u^{(-1)}(t)$, wobei $C_u^{(-1)}$ die Quasi-Inverse der bedingten Verteilung

$$C_u(v) = P(V \leq v \mid U = u)$$

von V , gegeben $U = u$, sei (vergleiche auch Satz 2.7).

- (3.) Aus dem Paar (u, v) transformiere nun komponentenweise den gewünschten Vektor als Quantile

$$(x, y) = \left(F_X^{(-1)}(u), F_Y^{(-1)}(v) \right)$$

der selbst gewählten Marginalverteilungen F_X respektive F_Y .

Welchen Copula-Typ man in diesen Schritten wählt, hängt von vielerlei Faktoren ab. Sicherlich beeinflusst das übergeordnete Ziel der Simulation unsere Entscheidung. Beispielsweise eignen sich manche Copula-Typen nicht für die Modellierung bestimmter Abhängigkeitsstrukturen wie *Tail-Dependence* (selten auch *Abhängigkeit in den Flanken* oder *Randabhängigkeit* genannt), während andere Typen diese durchaus adäquat und sinnvoll beschreiben können. Daneben spielt aber auch die Einfachheit, mit der man solche Copulas erzeugen oder implementieren kann, eine große Rolle. Viele Vertreter bestimmter Typen werden nach gemeinsamen nützlichen Eigenschaften kategorisiert, die die Arbeit mit diesen enorm erleichtern. Zwei Klassen, die *archimedischen* und die *elliptischen Copulas*, sollen im Folgenden präsentiert werden.

Archimedische Copulas

Zuerst sei die Klasse der archimedischen Copulas genannt, die alle oben genannten Vorteile (einfache Erzeugung, nützliche Eigenschaften) mit sich bringt und noch dazu eine Vielzahl von Repräsentanten besitzt. Wir beginnen gleich mit einer formalen Definition für diese Copula-Klasse und führen dabei den Begriff der „Generatorfunktion“ ein.

Satz 2.14:

Sei $\phi : [0, 1] \rightarrow [0, \infty]$ mit $\phi(1) = 0$ eine stetige, monoton fallende Funktion. Wir bezeichnen wie schon vorher mit

$$\phi^{(-1)}(t) := \begin{cases} \phi^{-1}(t), & \text{falls } 0 \leq t \leq \phi(0), \\ 0, & \text{falls } \phi(0) \leq t \leq \infty, \end{cases}$$

die Quasi-Inverse von ϕ . Sei außerdem eine Funktion $C : [0, 1]^2 \rightarrow [0, 1]$ gegeben durch

$$C(u, v) := \phi^{(-1)}(\phi(u) + \phi(v)). \quad (2.9)$$

Dann gilt: C ist genau dann eine Copula, wenn ϕ konvex ist.

Definition 2.15 (archimedische Copulas):

Copulas C mit der Gestalt aus Gleichung (2.9) heißen *archimedische*³ Copulas und die konvexe Funktion ϕ wird dabei *Generator-Funktion* (oder einfach *Generator*) von C genannt. Im Falle von $\phi(0) = \infty$ gilt die Invertierbarkeit und damit $\phi^{(-1)} \equiv \phi^{-1}$ und ϕ wird *strikt Generator* (sonst *nicht-strikter Generator*) genannt und auch die Copula C bezeichnen wir dann in expliziten Fällen mit *strikte archimedische Copula*.

Archimedische Copulas haben also den entscheidenden Vorteil gegenüber anderen Copula-Klassen, dass sie bereits nur mit Hilfe einer univariaten Generatorfunktion zu konstruieren sind.

Nun interessieren uns aber gerade die nützlichen Eigenschaften, die diese Klasse von Copula-Funktionen kennzeichnen. Diese sollen nun illustriert werden.

Satz 2.16 (Eigenschaften archimedischer Copulas):

Sei C eine archimedische Copula mit Generator ϕ . Dann gilt für alle $u, v, w \in [0, 1]$:

- (1.) C ist symmetrisch, das heißt $C(u, v) = C(v, u)$.
- (2.) C , aufgefasst als binärer Operator auf $[0, 1]$, ist assoziativ, das heißt also $C(C(u, v), w) = C(u, C(v, w))$.
- (3.) Für jede Konstante $k \in [0, \infty)$ ist auch $k \cdot \phi$ ein Generator von C .

Diese Eigenschaften sind für eine Vielzahl von Zwecken zwar sehr von Vorteil, jedoch kann man beispielsweise aus (1.) und (2.) im Höherdimensionalen folgern, dass die k -Copulas als Marginalien alle identisch sein müssen für $2 \leq k < n$. Außerdem ist die

³Die Bedeutung des Attributs „archimedisch“ rührt von der Interpretation einer Copula als binärer Operator her – Näheres dazu findet sich in erwähnten Standardwerken.

Tatsache, dass alle Familien in dieser Klasse bis auf ein oder höchstens zwei Parameter eindeutig festgelegt sind, ein weiterer Schwachpunkt. Diese Restriktionen an das Wesen der Abhängigkeitsstruktur gereichen den archimedischen Copulas zum Nachteil.

Als Letztes betrachten wir zwei wichtige Beispiele der Familie der archimedischen Copulas.

Beispiele (Archimedische Copulas):

- *Gumbel-Copulas* haben den (strikten) Ein-Parameter-Generator $\phi_\theta(t) := (-\ln t)^\theta$ mit Parameter $\theta \in [1, \infty)$. Daraus erhalten wir die (strikte) Ein-Parameter-Copula

$$C_\theta(u, v) := \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right).$$

- *Clayton-Copulas* nennen wir solche, die mit dem (für $\theta > 0$ strikten) Ein-Parameter-Generator $\phi_\theta(t) := \frac{1}{\theta}(t^{-\theta} - 1)$, mit $\theta \in [-1, \infty) \setminus \{0\}$, erzeugt wurden und die Gestalt

$$C_\theta(u, v) := \left[\max(u^{-\theta} + v^{-\theta} - 1, 0)\right]^{-\frac{1}{\theta}}$$

haben.

Abbildung 2.1 auf der nachfolgenden Seite stellt je zwei Beispiele einer Gumbel- und einer Clayton-Copula dar, zu je unterschiedlichen Parametern θ . Es handelt sich konkret um vier Scatterplots von je 500 simulierten Datenpaaren, die aus einer festen Copula zu festem Parameter erzeugt wurden. Die Copula-Parameter sind $\theta = 2,5$ in der ersten Zeile, in der zweiten Zeile $\theta = 10$. Links sind zu jedem der beiden Parameteroptionen eine Gumbel- (blau), rechts je ein Beispiel einer Clayton-Copula (grün) zu sehen.

Auf eine Zusammenstellung von weiteren Vertretern und von Zwei-Parameter-Familien wird an dieser Stelle verzichtet.⁴ Mit obigen Beispielen beenden wir die Thematisierung archimedischer Copulas und wenden uns nun den elliptischen Copula-Funktionen und ihren Besonderheiten zu.

⁴[25] bietet eine ausführliche tabellarische Übersicht über wichtige Vertreter dieser Copula-Familie.

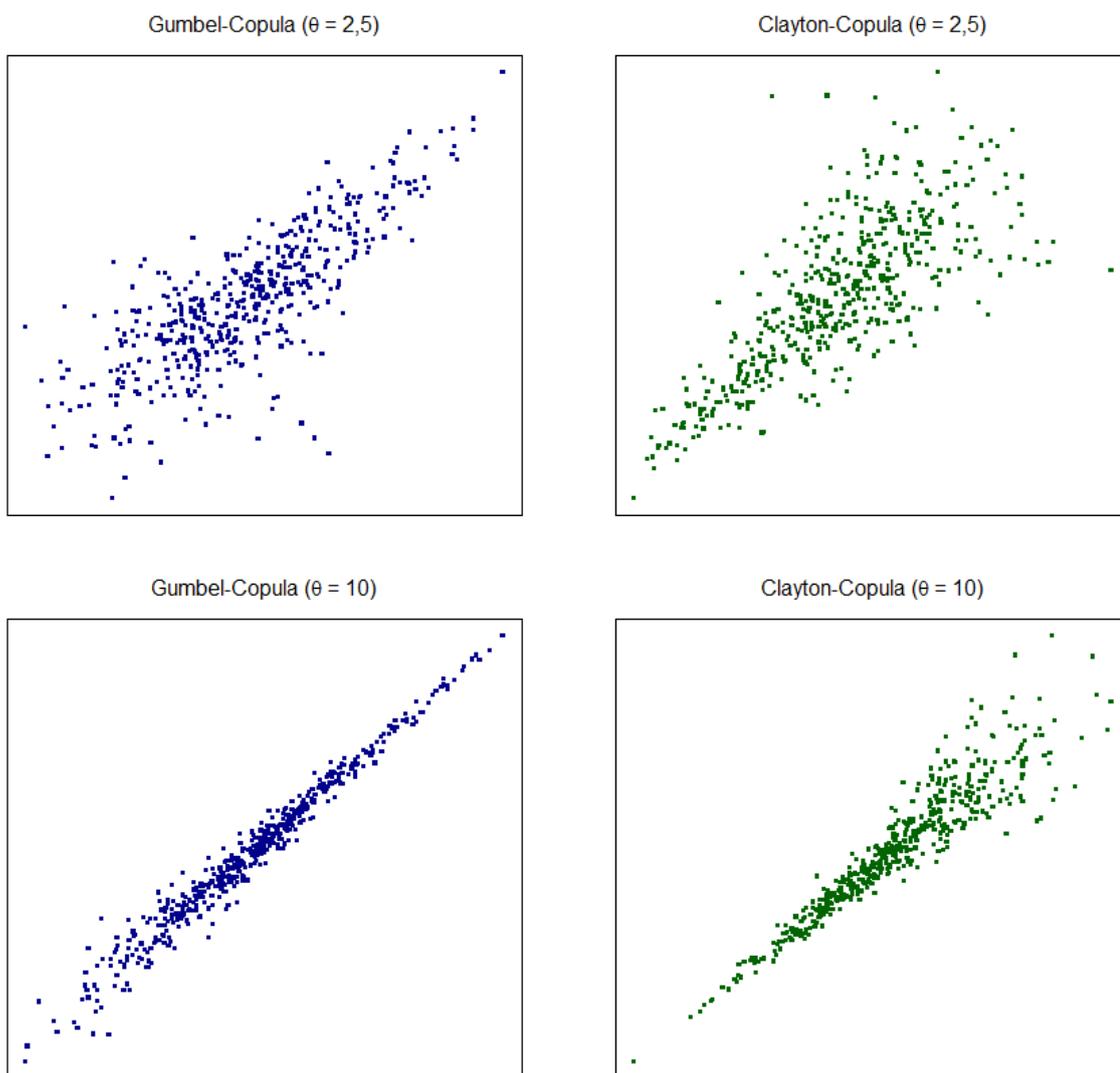


Abbildung 2.1.: Scatterplots der Gumbel- beziehungsweise Clayton-Copula

Elliptische Copulas

Der Begriff „elliptisch“ bei Copulas geht auf die Klasse der sogenannten *elliptischen Verteilungen* zurück, zu der beispielsweise die Normal- oder die Student-t-Verteilung gehören. Folgende Definition gibt zunächst Aufschluss über die charakteristischen Eigenschaften der Verteilungsklasse:

Definition 2.17 (Elliptische Verteilungen):

Für einen n -dimensionalen (reellen) Zufallsvektor \mathbf{X} , einen Vektor $\boldsymbol{\mu} \in \mathbb{R}^n$ und eine symmetrische, positiv definite $n \times n$ -Matrix Σ habe die charakteristische Funktion $\varphi_{\mathbf{X}-\boldsymbol{\mu}}$ von $\mathbf{X} - \boldsymbol{\mu}$ die Form

$$\varphi_{\mathbf{X}-\boldsymbol{\mu}}(\mathbf{t}) = \psi(\mathbf{t}^\top \Sigma \mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^n.$$

Dann bezeichnen wir die Verteilung von \mathbf{X} als *elliptisch* und schreiben $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \Sigma, \psi)$. Die Funktion ψ heißt auch *charakteristischer Generator* von \mathbf{X} , welcher zusammen mit $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}]$ und $\Sigma = \text{Cov}(\mathbf{X})$ die Verteilung eindeutig parametrisiert. Insbesondere bedeutet dies, dass elliptische Verteilungen durch die ersten beiden Momente bereits exakt charakterisiert werden.

Klären wir als Nächstes einige besondere Eigenheiten elliptischer Verteilungen, bevor wir eine elliptische Copula einführen.

Satz 2.18 (Eigenschaften elliptischer Verteilungen):

- (1.) Elliptische Verteilungen sind *radial* oder *kugelsymmetrisch*, das heißt für einen Zufallsvektor $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \Sigma, \psi)$ und für alle $\mathbf{a} \in \mathbb{R}^n$ gilt:

$$\mathbf{X} - \mathbf{a} \stackrel{\mathcal{D}}{=} \mathbf{a} - \mathbf{X}.$$

- (2.) Lineare Kombinationen $B\mathbf{X} + \mathbf{b}$ eines Zufallsvektors $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \Sigma, \psi)$ für eine $n \times n$ -Matrix B und einen Vektor $\mathbf{b} \in \mathbb{R}^n$ sind elliptisch verteilt mit Parameter $B\boldsymbol{\mu} + \mathbf{b}$, $B\Sigma B^\top$ und ψ :

$$B\mathbf{X} + \mathbf{b} \sim E_n(B\boldsymbol{\mu} + \mathbf{b}, B\Sigma B^\top, \psi).$$

- (3.) Alle Randverteilungen einer elliptischen Verteilung eines bestimmten Typs sind ebenfalls elliptische Verteilungen des selben Typs.

- (4.) Mit einem weiteren von $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \Sigma, \psi)$ unabhängigen Zufallsvektor $\mathbf{Y} \sim E_n(\boldsymbol{\lambda}, \Sigma, \chi)$ ist auch deren Faltung eine elliptische Verteilung mit Parameter $\boldsymbol{\mu} + \boldsymbol{\lambda}$, Σ und $\psi * \chi$:

$$\mathbf{X} + \mathbf{Y} \sim E_n(\boldsymbol{\mu} + \boldsymbol{\lambda}, \Sigma, \psi * \chi).$$

Es sei angemerkt, dass die Umkehrung zu Eigenschaft (3.) im Allgemeinen nicht gilt, denn durch die willkürliche Wahl einer Copula kann man auch eine gemeinsame Verteilung nach Sklars Theorem (Satz 2.5) generieren, die zwar ausschließlich elliptische Randverteilungen besitzt, selbst jedoch nicht elliptisch sein muss.

Definition 2.19 (Elliptische Copula):

Sei $\mathbf{X} \sim H := E_n(\boldsymbol{\mu}, \Sigma, \psi)$ eine Zufallsvariable mit eindimensionalen Randverteilungen $F_i := E_1(\mu_i, \sigma_i^2, \psi_i)$, $i = 1, \dots, n$. Dann nennen wir Copula-Funktionen C *elliptisch*, wenn sie die Form

$$C(\mathbf{u}) = H(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \quad (2.10)$$

für $\mathbf{u} \in [0, 1]^n$ besitzen.

Elliptische Copulas sind also schlicht die Copulas von elliptisch verteilten Zufallsvariablen. Der Vorteil dieser Copula-Klasse liegt in der Parametrisierung durch Parameter von elliptischen Verteilungen. Doch gerade die implizite Darstellung, die (im Gegensatz zu den archimedischen Copulas mit einer expliziten Generatorfunktion) direkt aus Folgerung 2.6 von Sklars Theorem abgeleitet ist, ist bei den elliptischen Copulas kein unerheblicher Nachteil. So kann man Zufallsvariablen nach bereits vorgestellter Inversionsmethode (Algorithmus auf Seite 34) nur numerisch erzeugen. Auch wegen der recht restriktiven Eigenschaften elliptischer Verteilungen (Kugelsymmetrie, Abgeschlossenheit der Typen-Klasse und dergleichen) ist dieser Typ nicht immer die beste Wahl. Für die Modellierung von Abhängigkeitsstrukturen eignen sich dagegen Copulas vom elliptischen Typ ganz besonders, vor allem da es sich beim Copula-Parameter um den (linearen) Korrelationskoeffizienten ρ beziehungsweise die Korrelationsmatrix R handelt.

Bevor wir uns mit dem nächsten Anwendungsbereich von Copulas befassen, sollen abschließend ein paar der wichtigsten Vertreter eben vorgestellter Copula-Klasse erwähnt werden.⁵

Beispiele (Elliptische Copulas):

- Die Familie der *Gauß-Copulas* hat die Gestalt

$$C_{\Sigma}(\mathbf{u}) = \Phi_n(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)),$$

wobei Φ_n die n -variate Normalverteilungsfunktion darstellt und $\Phi = \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. Der Parameter Σ sei die Kovarianzmatrix, während $\boldsymbol{\mu}$ der als fest angenommene Erwartungswertvektor sei (Alternativ lässt sich statt Σ auch die aus ihr durch Transformation hervorgehende Korrelationsmatrix R betrachten). Für $n = 2$ können wir die Formel noch leicht konkretisieren, so dass sich die Korrelationsmatrix

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

als Parameter der elliptischen Copula nur noch auf einen Wert ρ reduziert und wir damit die zweidimensionale Gauß-Copula als folgendes numerisches Integral schreiben können:

$$C_{\rho}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \underbrace{\frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right)}_{= c_{\rho}(u,v)} dx dy. \quad (2.11)$$

Aufgrund dieser Gestalt lässt sich auch gleich die Copula-Dichte $c_{\rho}(u, v)$ aus dem Integral ablesen.

⁵[25] bietet eine ausführliche tabellarische Übersicht über wichtige Vertreter dieser Copula-Familie.

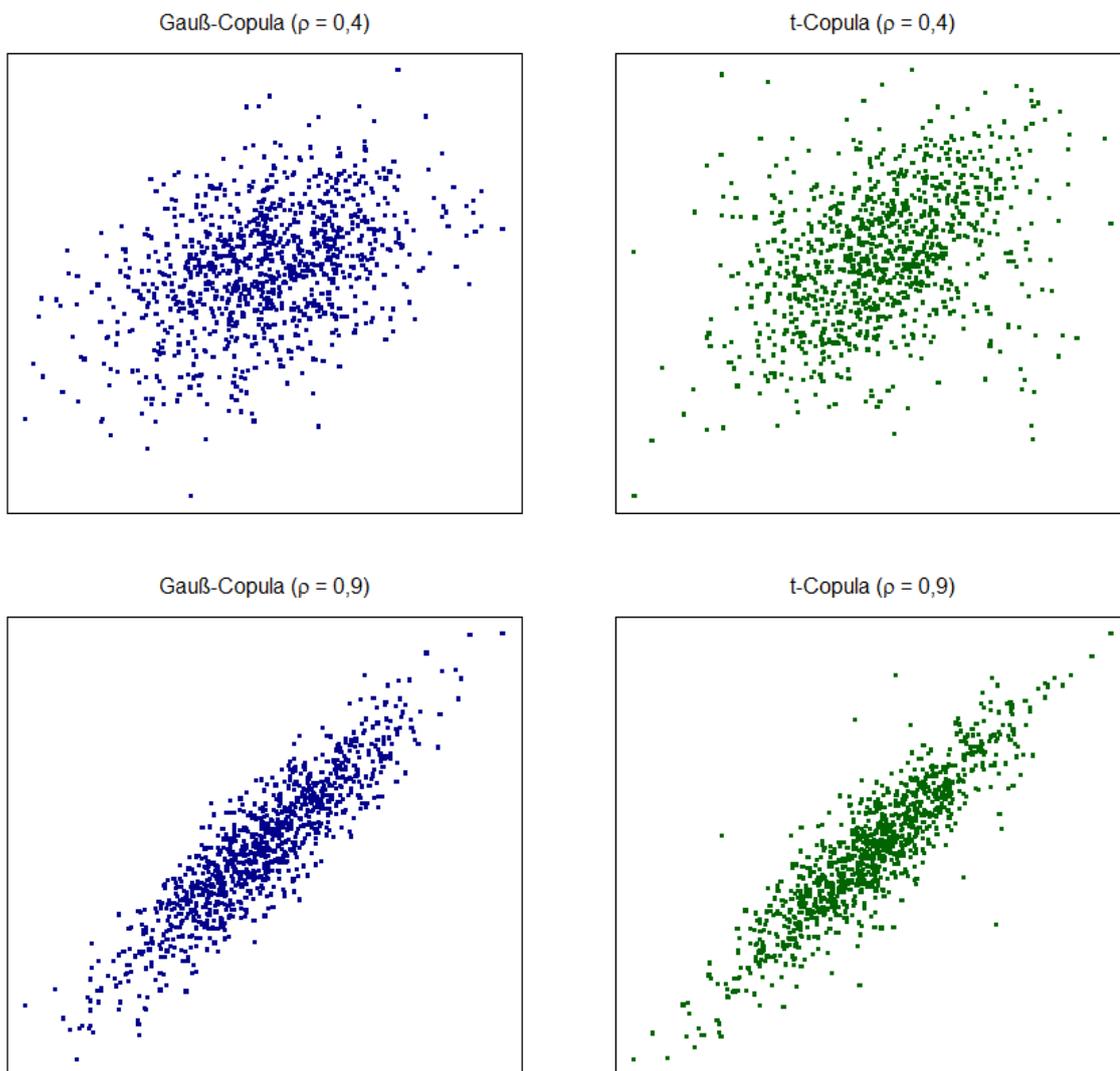


Abbildung 2.2.: Scatterplots der Gauß- beziehungsweise t-Copula

- *t-Copulas* besitzen die Form

$$C_{\Sigma, \nu}(\mathbf{u}) = T_{\nu}^n (T_{\nu}^{-1}(u_1), \dots, T_{\nu}^{-1}(u_n))$$

mit n -dimensionaler Student-t-Verteilung T_{ν}^n sowie den Inversen der eindimensionalen t-Verteilung T_{ν} . Dabei bezeichne ν die Anzahl der Freiheitsgrade, die also für alle Verteilungen gleich sein müssen. Für $n = 2$ erhalten wir eine ebenfalls noch übersichtliche Darstellung der t-Copula:

$$C_{\rho, \nu}(u, v) = \int_{-\infty}^{t_{\nu}^{-1}(u)} \int_{-\infty}^{t_{\nu}^{-1}(v)} \underbrace{\frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \left(1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}}}_{= c_{\rho, \nu}(u, v)} dx dy, \quad (2.12)$$

wobei $\rho \in [-1, 1]$ wie eben, $\nu > 0$ und $c_{\rho, \nu}(u, v)$ die Copula-Dichte sei.

In Abbildung 2.2 auf der vorigen Seite werden die Unterschiede zwischen je zwei Copulas der in obigem Beispiel vorgestellten Typen zu unterschiedlichen Parametern ρ veranschaulicht. Es handelt sich um zwei mal je zwei Plots von 1000 simulierten Daten aus einer Gauß-Copula (blau) und einer t-Copula (grün) zu den Parametern $\rho = 0,4$ und $\rho = 0,9$ mit jeweils den gleichen Marginalien (verschiedene Normalverteilungen).

2.3.2. Parametrische Schätzung

In der parametrischen Schätzung geht man im Allgemeinen von der Situation aus, dass man die Verteilungen jeweils bis auf zu schätzende Parameter bereits vollständig identifiziert hat.

Sei zunächst $\mathbf{X} := (X_1, \dots, X_n)^\top$ ein zeitkonstanter Zufallsvektor mit Marginalverteilungen $F_1(\cdot; \boldsymbol{\theta}), \dots, F_n(\cdot; \boldsymbol{\theta})$, parametrisiert durch den zu schätzenden Vektor $\boldsymbol{\theta}$. Wir bezeichnen mit $\mathbb{I} := \{1, \dots, n\}$ die Komponentenindexmenge. Die zugehörige Copula werde mit $C_{\boldsymbol{\theta}} = C(\cdot, \dots, \cdot; \boldsymbol{\theta})$ bezeichnet. Im Falle von Absolutstetigkeit der multivariaten Verteilung betrachtet man im Allgemeinen nicht die Darstellung $H(\mathbf{X}; \boldsymbol{\theta}) = C(F_1(X_1; \boldsymbol{\theta}), \dots, F_n(X_n; \boldsymbol{\theta}); \boldsymbol{\theta})$ für die gemeinsame multivariate Verteilung H , sondern bedient sich der Dichtefunktion h , die sich aus dem Produkt

$$\begin{aligned} h(\mathbf{X}; \boldsymbol{\theta}) &= \frac{\partial^n}{\partial X_1 \dots \partial X_n} H(\mathbf{X}; \boldsymbol{\theta}) \\ &= c(F_1(X_1; \boldsymbol{\theta}), \dots, F_n(X_n; \boldsymbol{\theta}); \boldsymbol{\theta}) \cdot \prod_{i=1}^n f_i(X_i; \boldsymbol{\theta}) \end{aligned}$$

der Randdichtefunktionen f_i , $i \in \mathbb{I}$, und der Copula-Dichte $c_{\boldsymbol{\theta}} := c(\cdot, \dots, \cdot; \boldsymbol{\theta})$ zusammensetzt. Aus diesem wird dann die Log-Likelihood-Funktion

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &:= \ln h(\mathbf{X}; \boldsymbol{\theta}) \\ &= \ln \left(c(F_1(X_1; \boldsymbol{\theta}), \dots, F_n(X_n; \boldsymbol{\theta}); \boldsymbol{\theta}) \cdot \prod_{i=1}^n f_i(X_i; \boldsymbol{\theta}) \right) \\ &= \sum_{i=1}^n \ln f_i(X_i; \boldsymbol{\theta}) + \ln c((F_1(X_1; \boldsymbol{\theta}), \dots, F_n(X_n; \boldsymbol{\theta}); \boldsymbol{\theta})) \end{aligned} \tag{2.13}$$

gebildet und durch Maximierung von \mathcal{L} erhalten wir den Schätzer $\hat{\boldsymbol{\theta}}$ des Parametervektors.

Wir verknüpfen nun in diesem Abschnitt die Betrachtungen der Copula-Parameterschätzung praktischerweise mit zeitabhängigen Zufallsvektoren, also multivariaten Zeitreihen, da wir diese Methoden in Kapitel 3 für GARCH-Zeitreihen direkt implementieren werden. Sei somit für die nachfolgende Illustration die Komponentenindexmenge wieder $\mathbb{I} := \{1, \dots, n\}$ und die Zeitindexmenge $\mathbb{T} := \{1, \dots, T\}$ für endliche Sample-Size T .

Nehmen wir demzufolge an, uns liegt jetzt eine beliebige n -dimensionale Zeitreihe $(\mathbf{z}_t)_{t \in \mathbb{T}}$ mit $\mathbf{z}_t = (z_{t,1}, \dots, z_{t,n})^\top$ und n -variater Verteilung H vor. Wir verwenden zunächst nur die Information aus den sich vollständig überlappenden Datenreihen, ohne Einschränkung sei somit \mathbb{T} der *gemeinsame* Beobachtungszeitraum *aller* Datenvektoren \mathbf{z}_t , $t \in \mathbb{T}$, unabhängig von $i \in \mathbb{I}$.⁶

Wir fassen wieder im Vektor $\boldsymbol{\theta}$ alle zu schätzenden Parameter für zuvor festgelegte (uni-variater) Marginalverteilungen sowie eine zuvor spezifizierte Copula-Familie zusammen.

Obige Dichte-Darstellung ergibt sich, erweitert auf die Zeitabhängigkeit, als

$$\begin{aligned} h(\mathbf{z}_t; \boldsymbol{\theta}) &= \frac{\partial^n}{\partial z_{t,1} \dots \partial z_{t,n}} H(\mathbf{z}_t; \boldsymbol{\theta}) \\ &= c(F_1(z_{t,1}; \boldsymbol{\theta}), \dots, F_n(z_{t,n}; \boldsymbol{\theta}); \boldsymbol{\theta}) \cdot \prod_{i=1}^n f_i(z_{t,i}; \boldsymbol{\theta}) \end{aligned}$$

beziehungsweise in der bedingten Form als

$$h_{|t-1}(\mathbf{z}_t; \boldsymbol{\theta}) = c_{|t-1}(F_{1|t-1}(z_{t,1}; \boldsymbol{\theta}), \dots, F_{n|t-1}(z_{t,n}; \boldsymbol{\theta}); \boldsymbol{\theta}) \cdot \prod_{i=1}^n f_{i|t-1}(z_{t,i}; \boldsymbol{\theta}),$$

gegeben alle Informationen aus der Vergangenheit *bis zum* Zeitpunkt $t - 1$. Die Informationsgesamtheit \mathcal{F}_{t-1} ist also die σ -Algebra $\sigma(\mathbf{z}_s : s \leq t - 1)$. Diese Darstellung benötigen wir für die nachfolgende „Gesamtdichte“:

Die „Gesamtdichte“ g *aller* Datenvektoren $\mathbf{z}_1, \dots, \mathbf{z}_T$ (zur Unterscheidung vom Begriff „gemeinsame Dichte“, des einzelnen Vektors \mathbf{z}_t) stellt sich im Allgemeinen aus dem Produkt

$$\begin{aligned} g(\mathbf{z}; \boldsymbol{\theta}) &= \prod_{t=1}^T h_{|t-1}(\mathbf{z}_t; \boldsymbol{\theta}) \\ &= \prod_{t=1}^T \left(c_{|t-1}(F_{1|t-1}(z_{t,1}; \boldsymbol{\theta}), \dots, F_{n|t-1}(z_{t,n}; \boldsymbol{\theta}); \boldsymbol{\theta}) \cdot \prod_{i=1}^n f_{i|t-1}(z_{t,i}; \boldsymbol{\theta}) \right) \quad (2.14) \\ &= \prod_{i=1}^n \prod_{t=1}^T f_{i|t-1}(z_{t,i}; \boldsymbol{\theta}) \cdot \prod_{t=1}^T c_{|t-1}(F_{1|t-1}(z_{t,1}; \boldsymbol{\theta}), \dots, F_{n|t-1}(z_{t,n}; \boldsymbol{\theta}); \boldsymbol{\theta}) \end{aligned}$$

der auf die Vergangenheit bedingten n -variater Dichten $h_{|t-1}$, $t \in \mathbb{T}$, zusammen.

⁶Später illustrieren wir eine Möglichkeit zur Verwendung verschiedener Datenlängen T_i für die einzelnen Komponenten $(z_{t,i})_{t \in \mathbb{T}_i}$ mit $\mathbb{T}_i := \{1, \dots, T_i\}$, $i \in \mathbb{I}$.

Nachdem wir Gleichung (2.14) logarithmiert haben, erhalten wir die im Argument $\boldsymbol{\theta}$ zu maximierende Log-Likelihood-Funktion

$$\begin{aligned}
 \mathcal{L}(\boldsymbol{\theta}) &:= \ln(g(\mathbf{z}; \boldsymbol{\theta})) \\
 &= \ln\left(\prod_{t=1}^T h_{|t-1}(\mathbf{z}_t; \boldsymbol{\theta})\right) \\
 &= \sum_{t=1}^T \ln(h_{|t-1}(\mathbf{z}_t; \boldsymbol{\theta})) \\
 &= \sum_{i=1}^n \sum_{t=1}^T \ln(f_{i|t-1}(z_{t,i}; \boldsymbol{\theta})) + \\
 &\quad + \sum_{t=1}^T \ln(c_{|t-1}(F_{1|t-1}(z_{t,1}; \boldsymbol{\theta}), \dots, F_{n|t-1}(z_{t,n}; \boldsymbol{\theta}); \boldsymbol{\theta})) \\
 &=: \sum_{i=1}^n \sum_{t=1}^T \ell_{t,i}(\boldsymbol{\theta}) + \sum_{t=1}^T \ell_{t,c}(\boldsymbol{\theta}) \\
 &=: \sum_{i=1}^n \mathcal{L}_i(\boldsymbol{\theta}) + \mathcal{L}_c(\boldsymbol{\theta}).
 \end{aligned} \tag{2.15}$$

Die Schätzung in der ML-Methode erfolgt jetzt, je nach Konstellation der Randparameter, durch Maximierung in einem oder mehreren Schritten. Beide Möglichkeiten laufen wie folgt ab.

Einschritt-Schätzung

Haben wir diese allgemeine Situation wie in Gleichung (2.15) vorliegen, so bleibt uns keine andere Möglichkeit, als den Schätzer $\hat{\boldsymbol{\theta}}$ mittels eines einzigen Schätzschrittes, formal durch die Maximierung

$$\hat{\boldsymbol{\theta}} := \operatorname{argmax}_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}) \tag{2.16}$$

der Log-Likelihood-Funktion, zu ermitteln. Wir nennen diesen Schätzer $\hat{\boldsymbol{\theta}}$ von nun an den *Einschritt-Maximum-Likelihood-Schätzer* (kurz *Einschritt-MLS* oder *Einschritt-Schätzer*).

Diese Optimierung ist sehr rechenaufwendig, zumal der Einschritt-Schätzer oftmals nur numerisch approximiert werden kann. Andererseits liefert dieses Verfahren ein zuverlässiges und sehr exaktes Resultat.

Wir wollen nun einige asymptotische Eigenschaften für den Einschritt-MLS $\hat{\boldsymbol{\theta}}$ anführen, ehe wir uns auf eine vereinfachte Situation der Trennbarkeit in einzelne Marginalparameter konzentrieren werden.

Satz 2.20 (Asymptotik des Einschnitt-MLS):

Sei $\boldsymbol{\theta}_0$ der wahre Parametervektor der zuvor beschriebenen Verteilungen. Maximiert man die Log-Likelihood-Funktion in Gleichung (2.15) simultan über alle Komponenten von $\boldsymbol{\theta}$, so ist der Einschnitt-Schätzer $\widehat{\boldsymbol{\theta}}$ unter gewissen Regularitätsbedingungen⁷

1. konsistent,
2. asymptotisch normalverteilt mit $\sqrt{T} \cdot (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{\theta}_0, I(\boldsymbol{\theta}_0)^{-1})$, wobei

$$I(\boldsymbol{\theta}_0) = \mathbb{E} \left[\left(\frac{\partial}{\partial \boldsymbol{\theta}_0} \mathcal{L}(\boldsymbol{\theta}_0) \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}_0} \mathcal{L}(\boldsymbol{\theta}_0) \right)^\top \right]$$

die Fisher-Informationsmatrix darstelle, und

3. asymptotisch effizient.

In der Praxis ist es oftmals zu aufwendig, alle Parameter gleichzeitig in einem Schritt zu schätzen. Insbesondere, wenn die Beobachtungsdimension n groß wird ($n \gg 2$) oder die Parameterdimension zunimmt, steigt der Rechenaufwand für die Optimierung immens.

Mehrschritt-Schätzung und flexible Datenlängen

Es kommt aber zu unseren Gunsten in einigen praktischen Situationen vor, dass die zugrundeliegende Information der Vergangenheit, enthalten in der σ -Algebra \mathcal{F}_{t-1} , nicht für alle Randvariablen relevant ist. So erreicht man beispielsweise, falls jede Marginalzeitreihe *nur* von ihren *eigenen* vergangenen Werten abhängt, eine Trennung des Vektors $\boldsymbol{\theta}$ in die einzelnen zu den Komponenten und zu der Copula gehörigen Parameter $\boldsymbol{\theta}_i$, $i \in \mathbb{I}$, beziehungsweise $\boldsymbol{\theta}_c$, so dass wir

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \dots, \boldsymbol{\theta}_n^\top, \boldsymbol{\theta}_c^\top)^\top$$

schreiben können. Dies ist dann der Fall, wenn die zugrundeliegende Informationsmenge $\mathcal{F}_{t-1,i}$ für eine Komponente nicht weniger relevante Information enthält als die Gesamtinformationsmenge.⁸ Daraus folgt außerdem für die Log-Likelihood-Summanden in Gleichung (2.15), dass $\mathcal{L}_i(\boldsymbol{\theta}) = \mathcal{L}_i(\boldsymbol{\theta}_i)$ für alle $i \in \mathbb{I}$ gilt; für \mathcal{L}_c ist dagegen das zu maximierende Argument nach wie vor der ganze Vektor $\boldsymbol{\theta}$, handelt es sich ja bei den Argumenten der bedingten Copula (mit Parameter $\boldsymbol{\theta}_c$) um die bedingten Randverteilungsfunktionen (mit Parametern $\boldsymbol{\theta}_i$, $i \in \mathbb{I}$).

⁷siehe [30] und [18].

⁸Formal spiegelt sich dies in der Bedingung $z_{t,i} | \mathcal{F}_{t-1,i} \stackrel{\mathcal{D}}{=} z_{t,i} | \mathcal{F}_{t-1}$ wider.

Aber was für Möglichkeiten bieten sich nun bei dieser Ausgangssituation? Betrachten wir dafür wieder die Log-Likelihood-Funktion \mathcal{L} aus Gleichung (2.15), dann ergibt sich die Darstellung

$$\mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_c) = \sum_{i=1}^n \mathcal{L}_i(\boldsymbol{\theta}_i) + \mathcal{L}_c(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_c). \quad (2.17)$$

Uns ist es nun möglich, die Maximierung von \mathcal{L} nicht simultan in allen Komponenten von $\boldsymbol{\theta}$ durchzuführen, sondern stattdessen in mehreren getrennten Schritten. Diese sogenannte *Mehrschrittschätzung*, bekannt geworden unter der Bezeichnung *IFM* (aus dem Englischen *inference functions for margins*) von [18], ist aufgeteilt in die zwei Hauptschritte

(1.) Schätzung der Marginalparameter durch

$$\tilde{\boldsymbol{\theta}}_i := \operatorname{argmax}_{\boldsymbol{\theta}_i} \mathcal{L}_i(\boldsymbol{\theta}_i), \quad i \in \mathbb{I},$$

unabhängig voneinander (dann vor allem *nacheinander*) und

(2.) Schätzung der Copula-Parameter durch

$$\tilde{\boldsymbol{\theta}}_c := \operatorname{argmax}_{\boldsymbol{\theta}_c} \mathcal{L}_c(\tilde{\boldsymbol{\theta}}_1, \dots, \tilde{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_c),$$

bedingt auf die Schätzer aus Schritt (1.).

Daraus erhalten wir insgesamt den alternativen *Mehrschritt-Maximum-Likelihood-Schätzer* (kurz Mehrschritt-MLS oder Mehrschritt-Schätzer)

$$\tilde{\boldsymbol{\theta}} := \left(\tilde{\boldsymbol{\theta}}_1^\top, \dots, \tilde{\boldsymbol{\theta}}_n^\top, \tilde{\boldsymbol{\theta}}_c^\top \right)^\top. \quad (2.18)$$

Eine weitere, äußerst sinnvolle Unterstützung bietet diese Ausgangssituation für die Datenlängen jeder einzelnen Komponentenzeitreihe $(z_{t,i})_{t \in \mathbb{T}_i}$, $i \in \mathbb{I}$. Wie bereits angekündigt, nutzen wir jetzt die separierte Summendarstellung aus Gleichung (2.17) aus, um flexible Datenlängen zu erlauben. Wir erhalten damit

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &= \sum_{i=1}^n \mathcal{L}_i(\boldsymbol{\theta}_i) + \mathcal{L}_c(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_c) \\ &= \sum_{i=1}^n \sum_{t=1}^{T_i} \ell_{t,i}(\boldsymbol{\theta}_i) + \sum_{t=1}^T \ell_{t,c}(\boldsymbol{\theta}) \end{aligned} \quad (2.19)$$

und können im ersten Optimierungsschritt auf alle Randdaten zurückgreifen, ohne diese auf die gemeinsame Beobachtungslänge „trimmen“ zu müssen. Nehmen wir an, dass in obiger Gleichung T_i die Sample-Size der i -ten beobachteten Zeitreihe $(z_{t,i})_{t \in \mathbb{T}_i}$ sei (alle beginnend mit $t = 1$), so können wir die einzelnen Zeitindexmengen durch $\mathbb{T}_i := \{1, \dots, T_i\}$ voneinander unterscheiden. Mit $T := \min\{T_i : i \in \mathbb{I}\}$ bezeichnen wir dann nur noch die

Länge des sich überlappenden Beobachtungszeitraums, ohne Einschränkung also $\mathbb{T} := \{1, \dots, T\}$. Da die Optimierung im ersten Schritt komponentenweise für jedes $i \in \mathbb{I}$ ohnehin unabhängig und nacheinander geschieht, ist diese Berücksichtigung verschiedener Sample-Sizes sehr einfach zu implementieren.

Aufgrund jener Konstellation und der damit ermöglichten „Nacheinander-Optimierung“ in den beiden Hauptschritten ist diese Vorgehensweise insgesamt um einiges weniger rechenintensiv als die simultane Einschnitt-Schätzung. Außerdem ist durch die Möglichkeit der Flexibilisierung der Sample-Size für jede Komponentenzeitreihe ein sehr realistischer Aspekt integrierbar, der unter anderen Umständen nicht berücksichtigt werden kann.

Doch wie effizient ist dieser Mehrschritt-Schätzer im Vergleich zum Einschnitt-MLS $\widehat{\boldsymbol{\theta}}$ aus Gleichung (2.16)? Wir wollen dieser Frage nun nachgehen und die Asymptotik der beiden Schätzer $\widetilde{\boldsymbol{\theta}}$ und $\widehat{\boldsymbol{\theta}}$ gegenüberstellen und besonders unter dem Gesichtspunkt der asymptotischen Effizienz miteinander vergleichen. Uns interessiert also vor allem das Verhalten des Mehrschritt-Schätzers, wenn der Stichprobenumfang T beliebig groß wird ($T \rightarrow \infty$).

Einschnitt- versus Mehrschritt-Schätzung

In den folgenden Darstellungen werden wir uns der Einfachheit halber wieder auf die Dimension $n = 2$ beschränken, da die Allgemeinheit der Formeln sonst auf Kosten der Übersicht geht. Sei dafür zunächst $\mathbb{T}_i := \{1, \dots, T_i\}$, $i \in \mathbb{I} := \{1, 2\}$. Ohne Einschränkung sollen also alle Zeitreihen mit $t = 1$ beginnen. Wir bezeichnen außerdem durch $T := T_1 \wedge T_2$ die Länge des Zeitraums, in dem beide Zeitreihen gleichzeitig beobachtet werden können. Die Log-Likelihood-Funktion wird in diesem Fall folglich zerlegt in die Summe $\mathcal{L}_1(\boldsymbol{\theta}_1) + \mathcal{L}_2(\boldsymbol{\theta}_2) + \mathcal{L}_c(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_c)$, wobei wir hier mit

$$(1.) \quad \widetilde{\boldsymbol{\theta}}_1 := \operatorname{argmax}_{\boldsymbol{\theta}_1} \mathcal{L}_1(\boldsymbol{\theta}_1) := \operatorname{argmax}_{\boldsymbol{\theta}_1} \frac{1}{T_1} \sum_{t=1}^{T_1} \ell_{t,1}(\boldsymbol{\theta}_1),$$

$$\widetilde{\boldsymbol{\theta}}_2 := \operatorname{argmax}_{\boldsymbol{\theta}_2} \mathcal{L}_2(\boldsymbol{\theta}_2) := \operatorname{argmax}_{\boldsymbol{\theta}_2} \frac{1}{T_2} \sum_{t=1}^{T_2} \ell_{t,2}(\boldsymbol{\theta}_2),$$

$$(2.) \quad \widetilde{\boldsymbol{\theta}}_C := \operatorname{argmax}_{\boldsymbol{\theta}_c} \mathcal{L}_c(\widetilde{\boldsymbol{\theta}}_1, \widetilde{\boldsymbol{\theta}}_2, \boldsymbol{\theta}_c) := \operatorname{argmax}_{\boldsymbol{\theta}_c} \frac{1}{T} \sum_{t=1}^T \ell_{t,c}(\widetilde{\boldsymbol{\theta}}_1, \widetilde{\boldsymbol{\theta}}_2, \boldsymbol{\theta}_c)$$

die entsprechenden Schätzer als maximierende Argumente der (zusätzlich noch „standardisierten“) Log-Likelihood-Summanden definieren wollen. Die Parameterdimensionen bezeichnen wir mit p , q und r , so dass $\boldsymbol{\theta}_1 \in \mathbb{R}^p$, $\boldsymbol{\theta}_2 \in \mathbb{R}^q$, $\boldsymbol{\theta}_c \in \mathbb{R}^r$ und damit $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top, \boldsymbol{\theta}_c^\top)^\top \in \mathbb{R}^s$ mit $s := p + q + r$.

In [30] werden nun Regularitätsbedingungen an die Funktionen \mathcal{L}_1 , \mathcal{L}_2 und \mathcal{L}_c gegeben, unter denen der Mehrschritt-Schätzer $\widetilde{\boldsymbol{\theta}}$ (sogar unter Annahmen unterschiedlicher Startzeitpunkte) asymptotisch normalverteilt ist. In folgendem Satz halten wir dessen Resultat

fest, angepasst an einen gemeinsamen Startzeitpunkt $t = 1$ bei unterschiedlichen Sample-Sizes T_1 und T_2 :

Satz 2.21 (Asymptotik des Mehrschritt-Schätzers):

Sei $\boldsymbol{\theta}_0$ der wahre Parametervektor. Seien $T_1 \geq T_2 \geq T$ derart, dass für $T_1 \rightarrow \infty$ die Grenzwertbetrachtungen

$$\frac{T_2}{T_1} \rightarrow \lambda_{2,1} \in (0, 1] \quad \text{beziehungsweise} \quad \frac{T}{T_1} \rightarrow \lambda_{c,1} \in (0, 1]$$

existieren. Unter gewissen Regularitätsbedingungen⁹ an die Erwartungswert-Matrizen $A := \mathbb{E}[H(\tilde{\boldsymbol{\theta}})]$ und $B := \mathbb{E}[P(\tilde{\boldsymbol{\theta}})]$ ist der Mehrschritt-Schätzer

1. konsistent und
2. gemäß folgender schwacher Konvergenz

$$B^{-1/2} \cdot \mathcal{T}^{1/2} \cdot A \cdot (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, I_s), \quad T_1 \rightarrow \infty,$$

asymptotisch normalverteilt.

3. Des Weiteren ist der Schätzer *nicht* asymptotisch effizient:

Normieren wir \mathcal{T} zur Matrix $\mathcal{T}^* := \frac{1}{T}\mathcal{T}$, so erhalten wir mit dem Grenzwert

$$\mathcal{T}_\infty^* := \lim_{T_1 \rightarrow \infty} \mathcal{T}^*$$

und dem Matrixprodukt

$$V(\tilde{\boldsymbol{\theta}}) = A^{-1} \cdot (\mathcal{T}_\infty^*)^{-1/2} \cdot B \cdot (\mathcal{T}_\infty^*)^{-1/2} \cdot (A^{-1})^\top \quad (2.20)$$

die asymptotische Kovarianzmatrix des Mehrschritt-Schätzers, welche die untere Schranke der Cramér-Rao-Ungleichung niemals erreicht.¹⁰

Dabei seien I_k die $k \times k$ -Einheitsmatrix und

$$\mathcal{T} := \begin{pmatrix} T_1 \cdot I_p & 0 & 0 \\ 0 & T_2 \cdot I_q & 0 \\ 0 & 0 & T \cdot I_r \end{pmatrix}, \quad (2.21)$$

⁹Vgl. [30], S. 152f.

¹⁰Dies ist eine hinreichende Bedingung an die asymptotische Kovarianzmatrix für asymptotische Effizienz des entsprechenden Schätzers.

$$H(\tilde{\boldsymbol{\theta}}) := \begin{pmatrix} \frac{1}{T_1} \sum_{t=1}^{T_1} \frac{\partial^2}{\partial \boldsymbol{\theta}_1^2} \ell_{t,1}(\tilde{\boldsymbol{\theta}}_1) & 0 & 0 \\ 0 & \frac{1}{T_2} \sum_{t=1}^{T_2} \frac{\partial^2}{\partial \boldsymbol{\theta}_2^2} \ell_{t,2}(\tilde{\boldsymbol{\theta}}_2) & 0 \\ \frac{1}{T} \sum_{t=1}^T \frac{\partial^2}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_c} \ell_{t,c}(\tilde{\boldsymbol{\theta}}) & \frac{1}{T} \sum_{t=1}^T \frac{\partial^2}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_c} \ell_{t,c}(\tilde{\boldsymbol{\theta}}) & \frac{1}{T} \sum_{t=1}^T \frac{\partial^2}{\partial \boldsymbol{\theta}_c^2} \ell_{t,c}(\tilde{\boldsymbol{\theta}}) \end{pmatrix}, \quad (2.22)$$

$$P(\tilde{\boldsymbol{\theta}}) := \begin{pmatrix} \frac{1}{T_1} \sum_{t=1}^{T_1} s_{t,1} s_{t,1}^\top & \frac{1}{\sqrt{T_1 T_2}} \sum_{t=1}^{T_2} s_{t,1} s_{t,2}^\top & \frac{1}{\sqrt{T_1 T}} \sum_{t=1}^T s_{t,1} s_{t,c}^\top \\ \frac{1}{\sqrt{T_2 T_1}} \sum_{t=1}^{T_2} s_{t,2} s_{t,1}^\top & \frac{1}{T_2} \sum_{t=1}^{T_2} s_{t,2} s_{t,2}^\top & \frac{1}{\sqrt{T_2 T}} \sum_{t=1}^T s_{t,2} s_{t,c}^\top \\ \frac{1}{\sqrt{T T_1}} \sum_{t=1}^T s_{t,c} s_{t,1}^\top & \frac{1}{\sqrt{T T_2}} \sum_{t=1}^T s_{t,c} s_{t,2}^\top & \frac{1}{T} \sum_{t=1}^T s_{t,c} s_{t,c}^\top \end{pmatrix} \quad (2.23)$$

weitere Matrizen mit den Gradienten $s_{t,i} := \frac{\partial}{\partial \boldsymbol{\theta}_i} \ell_{t,i}(\tilde{\boldsymbol{\theta}}_i)$, $i \in \mathbb{I}$, beziehungsweise $s_{t,c} := \frac{\partial}{\partial \boldsymbol{\theta}_c} \ell_{t,c}(\tilde{\boldsymbol{\theta}})$, $t \in \mathbb{T}$.

Da eine asymptotische Effizienz also nicht erreicht wird, schwächen wir hier unsere Ambitionen etwas ab und vergleichen nur den *Grad der Effizienz* zwischen dem Einschnitt- und dem Mehrschritt-Schätzer (oder die *relative Effizienz*, gemessen am Einschnitt-MLS). Dieser Vergleich wird durch die „Differenz der Abstände“ zur unteren Schranke der Cramér-Rao-Ungleichung und somit durch die Differenz der asymptotischen Kovarianzmatrizen unserer Schätzer quantifiziert: Ist die Differenzmatrix positiv semi-definit, so ist der zweite Parameter asymptotisch effizienter als der erste; im negativ semi-definiten Fall ist es entsprechend umgekehrt.

In der Situation $T = T_1 = T_2$ ist der Einschnitt-Schätzer $\hat{\boldsymbol{\theta}}$ stets der effizienteste Schätzer, da mit diesem exakt die Identität in der Cramér-Rao-Ungleichung erreicht wird (Satz 2.20), wohingegen die Kovarianzmatrix des Mehrschritt-Schätzers $\tilde{\boldsymbol{\theta}}$ immer über der unteren Schranke liegt (Satz 2.21). Im leicht abweichenden Falle von zumindest asymptotisch vernachlässigbaren Unterschieden zwischen den einzelnen Sample-Sizes gilt ähnlich: $\hat{\boldsymbol{\theta}}$ ist effizienter als $\tilde{\boldsymbol{\theta}}$. Bei nicht vernachlässigbaren Unterschieden zwischen T , T_1 und T_2 kann man dagegen hinreichende Bedingungen dafür konstruieren, dass der Mehrschritt-Schätzer zumindest *nicht weniger effizient* ist als der Einschnitt-Schätzer.

Satz 2.22 (Bedingungen für vergleichbaren Effizienzgrad des Mehrschritt-Schätzers):
Gilt für die Grenzwerte

$$\begin{aligned} \frac{T}{T_1} &\rightarrow \lambda_{c,1} < \frac{\widehat{V}_{jj}}{M_{jj}} < 1, \quad j \in \{1, \dots, p\} \quad \text{oder} \\ \frac{T}{T_2} &\rightarrow \lambda_{c,2} < \frac{\widehat{V}_{kk}}{M_{kk}} < 1, \quad k \in \{p+1, \dots, q\} \end{aligned}$$

mit asymptotischer Kovarianzmatrix \widehat{V} von $\widehat{\boldsymbol{\theta}}$ und $M = A^{-1} \cdot B \cdot (A^{-1})^\top$, dann ist der Mehrschritt-Schätzer $\widetilde{\boldsymbol{\theta}}$ nicht weniger asymptotisch effizient als der Einschnitt-Schätzer $\widehat{\boldsymbol{\theta}}$.¹¹

Es gibt sogar spezielle Situationen, in denen keiner der Schätzer effizienter ist als der andere. Dann vergleicht man den Effizienzgrad für die Marginal- und für die Copulaparameter separat. Nehmen wir zum Beispiel den Fall, dass Phasen des Nicht-Überlappens der beobachteten Daten deutlich länger sind als Zeiträume mit Überschneidungen. Dann ist der Mehrschritt-MLS für die Marginalparameter effizienter, für die Copula-Parameter ist dagegen der Einschnitt-MLS der effizientere Schätzer.

2.4. Ausblick

Dieses Kapitel diene vorwiegend einem knappen Einblick in die Theorie der Copulas und die vielen Facetten der Parameterschätzung, welche sehr ausführlich beschrieben wurde. Wir haben zunächst zentrale Aussagen über Copula-Funktionen knapp festgehalten, auf denen die praktischen Anwendungsmöglichkeiten in der Statistik beruhen. So wurden unter anderem Verfahren zur Simulation von abhängigen Zufallspaaren mit Hilfe von Copula-Funktionen veranschaulicht sowie nützliche Klassen von Copula-Funktionen jeweils mit ihren Vorzügen und Nachteilen vorgestellt. Schließlich haben wir (mit Blick auf die inhaltliche Verbindung zu Kapitel 1 im nachfolgenden Kapitel) einige detaillierte Vorbereitungen auf die parametrische Schätzung von copula-basierten Zeitreihenmodellen geboten.

Einige weitere Anwendungsgebiete von Copulas wie die copula-basierte Analyse von Abhängigkeitsstrukturen beispielsweise für die Typen *monotone Abhängigkeit* oder *Abhängigkeit in den Flanken* (englisch *tail dependence*) werden in eingangs erwähnten Einstiegswerken vorgestellt. Das Ziel dieser Arbeit stets im Hinterkopf, haben wir bewusst auf Illustrationen dieser Konzepte verzichtet. Daher schließen wir dieses Kapitel mit einem nur sehr kleinen Ausblick über Aktuelles aus dem Themengebiet der Copula-Funktionen.

¹¹Vgl. [30], S. 153.

Viele finanzwissenschaftliche Zweige der Mathematik wie das Risikomanagement, zum Beispiel in der Kreditrisikoanalyse, haben die Theorie und Anwendung von Copulas inzwischen vollständig verinnerlicht und deren Nützlichkeit ist mittlerweile allgemein anerkannt. Doch trotz der Vielzahl umfassender Entwicklungen gibt es meist keine allgemeine Erweiterung von bivariaten Modellen auf höhere Dimensionen. Erschwerend kommt hinzu, dass man die Stetigkeit der eindimensionalen Verteilungen benötigt, um aus diesen eine eindeutige multivariate Verteilungsfunktion zu konstruieren. So gibt es etwa noch wenig Fortschritt hinsichtlich der Analyse der Abhängigkeitsstruktur für nicht-stetige Verteilungsfunktionen. Da die Idee hinter den Copula-Funktionen ein noch relativ neues Konzept ist und die Kluft zwischen Theorie und Praxis vor allem in der Ökonometrie für viele nicht schmal genug sein kann, trifft sie nicht überall auf volle Zustimmung. Heftige Kritik am „Hype“ um dieses Thema übt etwa beispielsweise [24], welchen wir mit folgendem Satz zitieren wollen:

It is contradictory that in risk management, where one observes a lot of dependence through time, copulas are applied most frequently.

Dass man an diesem Schwachpunkt jedoch mehrfach anzusetzen versucht hat, zeigen wir nun im nächsten Kapitel dieser Arbeit, in dem wir bei unserer Modellierung einen zeitvariierenden Abhängigkeitsparameter für eine Copula zugrundelegen werden. Dort wird ein copula-basiertes Zeitreihenmodell vorgestellt, entwickelt auf dem Grundgedanken der Copula-Theorie, univariate Modelle (nämlich die Marginalmodelle) mit einer Copula-Funktion zu einem multivariaten Modell zu verknüpfen.

Kapitel 3

Copula-GARCH-Modelle

Stellen wir uns als motivierendes Beispiel zwei verschiedene Währungskurse vor. Gehen wir davon aus, dass jeweils eine Student-t-Verteilung die bedingten univariaten Verteilungsfunktionen der Devisen am adäquatesten beschreibt. Will man nun die gemeinsame Verteilung modellieren, hat man nur die bivariate t-Verteilung als passende Erweiterung zur Verfügung. Allerdings haben wir auch schon Kapitel in 2 (Abschnitt 2.3) erfahren, dass die Randverteilungen einer multivariaten elliptischen Verteilung, zu denen die t-Verteilung bekanntlich gehört, exakt vom gleichen Typ sein müssen. Doch wäre dies ziemlich restriktiv, wenn zwei ungleiche t-verteilte Devisenkurse stets identische Freiheitsgrad-Parameter hätten.

Die Schwierigkeit in dieser Situation wäre bei vernachlässigbar kleinen Unterschieden zwischen den Randverteilungen noch einigermaßen zu ignorieren. Wir möchten aber auch in der Lage sein, zwei völlig voneinander verschiedene Zeitreihen gemeinsam zu modellieren, beispielsweise einen Währungskurs und den Kurs einer Aktienrendite, so dass die Randverteilungen überhaupt nicht mehr vergleichbaren Typs sein müssen und wir somit nicht ohne Weiteres auf eine gegebene bivariate Verteilung zurückgreifen können.

In diesem Kapitel machen wir uns deswegen die Aussagen und Folgerungen aus Sklars Theorem (Satz 2.5) aus Kapitel 2 zunutze, die uns erlauben, die n -dimensionale gemeinsame Verteilungsfunktion in ihre Randverteilungen und in eine Copula zu zerlegen, mit der man vollständige Informationen über die Abhängigkeitsstruktur zur Verfügung hat.

Eine der Implikationen gab uns Folgerung 2.6: Nun können wir nach den in Kapitel 1 kennengelernten Regeln die univariaten Finanzzeitreihen separat modellieren – zu völlig inhomogenen Modellannahmen (wie etwa an die Verteilungsparameter oder gar die Verteilungsklasse, an die Sample-Size und sogar an das univariate Modell selbst) – und schließlich deren univariate Verteilungen als Randverteilungen eines höherdimensionalen Modells auffassen, das es in einem zweiten Schritt mit Hilfe einer Copula-Funktion anzupassen gilt.

Der Nutzen dieser Vorgehensweise ist enorm, gerade da die vorhandene Literatur zur univariaten Modellierung eine wesentlich höhere Auswahl an flexiblen parametrischen Verteilungsfamilien als zu multivariaten Verteilungen zu bieten hat. Außerdem sind meist auch multivariate GARCH-Erweiterungen aufgrund der komplexen Kovolatilitätsgleichungen

nur sehr mühsam und umständlich zu implementieren sowie wegen des erhöhten Gradbedarfs mit zu hohem Rechenaufwand verbunden, da die Dimension der zu schätzenden Parameter mit wachsender Beobachtungsdimension n rapide ansteigen kann (vergleiche etwa Kapitel 1, Abschnitt 1.3).

Auch haben wir in Kapitel 2 erkannt, dass man die Theorie der (unbedingten) Copula-Funktionen in die Sprache der bedingten Copulas übersetzen kann. Daraus folgt für uns, dass wir (bedingte) Copulas in jedem Fall in den finanzökonometrischen Kontext der Volatilitätsmodellierung einbauen können.

Dieses Kapitel stellt die inhaltlichen Entwicklungen der ersten beiden Kapitel in einen sinnvollen Zusammenhang. Zuerst werden wir in Abschnitt 3.1 copula-basierte Volatilitätsmodelle als eine eigene multivariate Modellklasse klassifizieren und von anderen Modellklassen abgrenzen. In Abschnitt 3.2 wird ein solches copula-basiertes multivariates GARCH-Modell dann beispielhaft präsentiert und in einer Simulationsstudie auf deren Tauglichkeit hin geprüft. Hier werden wir den grundlegenden Modellaufbau sowie die Simulation und Schätzung im Einzelnen näher betrachten, wobei wir auf die in den Kapiteln 1 (Abschnitt 1.2.2) und 2 (Abschnitt 2.3.2) behandelte Schätztheorie (für GARCH-Modelle beziehungsweise copula-basierte Modelle) zurückgreifen werden.

3.1. Copula-basierte Volatilitätsmodelle

Sei für diesen Abschnitt zunächst die diskrete Zeitindexmenge \mathbb{T} nicht näher spezifiziert. Die Komponentenindexmenge \mathbb{I} sei definiert durch $\mathbb{I} := \{1, \dots, n\}$ für eine beliebige, feste Dimension $n \geq 2$.

Copula-basierte Volatilitätsmodelle finden sich in der aktuellen Forschung seit wenigen Jahren vergleichsweise häufig, doch versucht meist jeder Autor auf seine eigene Weise eine formale Eingrenzung der Modellklasse. Aufgrund des vielfachen praktischen Erfolgs der GARCH-Modelle für grundlegende univariate Volatilitätsuntersuchungen ist bei diesen dann fast immer von Copula-GARCH-Modellen die Rede.

Zwar läuft dieses Kapitel ebenso darauf hinaus, ein solches sehr vereinfachtes Modell zu präsentieren, hauptsächlich jedoch zu dem Zweck die Einfachheit der Implementierung zu veranschaulichen. Wir versuchen hier in der Tat eine konventionelle Definition der Modellklasse zu liefern – auch unter probabilistischen Aspekten hinsichtlich der zugrundeliegenden stochastischen Prozesse. Letztere Intention bedient sich vorwiegend der in den Kapiteln 1 und 2 bereits in Grundzügen geschaffenen wahrscheinlichkeitstheoretischen Grundlagen.

In [20] werden nachfolgend definierte *copula-basierte Volatilitätsmodelle* als sogenannte *CMD-Modelle* beschrieben (aus dem Englischen *copula based multivariate dynamic*). Wir werden dieses praktische Kürzel aus Gründen der Einfachheit adaptieren.

3.1.1. Probabilistische Definition

Gehen wir ab jetzt nur noch von stetigen Verteilungsfunktionen für alle nachfolgend behandelten Zufallsvariablen aus, so lässt sich eine erste Definition von CMD-Modellen formulieren:

Definition 3.1 (Copula-basiertes Volatilitätsmodell, CMD-Prozess):

Sei $(\mathbf{r}_t)_{t \in \mathbb{T}}$ mit $\mathbf{r}_t = (r_{t,1}, \dots, r_{t,n})^\top$ eine Renditezeitreihe, die in nicht näher definierter, deterministischer Weise vom Zufallshock $(\mathbf{z}_t)_{t \in \mathbb{T}}$ mit $\mathbf{z}_t = (z_{t,1}, \dots, z_{t,n})^\top$ abhängt. Mit $\mathcal{F}_{t-1} := \sigma(\mathbf{r}_s : s \leq t-1)$ sei die Informationsgesamtheit aller bereits vergangener Kursverläufe symbolisiert. Wir treffen weiterhin folgende Annahmen:

- (a) Der Zufallshock \mathbf{z}_t habe zu jedem Zeitpunkt $t \in \mathbb{T}$ die univariaten Randverteilungsfunktionen F_1, \dots, F_n .
- (b) Der Vektor $\mathbf{u}_t = (u_{t,1}, \dots, u_{t,n})^\top$ habe zum jeweiligen Zeitpunkt $t \in \mathbb{T}$ die bedingte n -Copula $C_t(\mathbf{u}_t | \mathcal{F}_{t-1})$ als (im Allgemeinen *zeitabhängige*) multivariate bedingte Verteilungsfunktion.
- (c) Für den Zusammenhang zwischen \mathbf{z}_t und \mathbf{u}_t gelte schließlich $u_{t,i} = F_i(z_{t,i} | \mathcal{F}_{t-1})$, für alle $t \in \mathbb{T}$ und $i \in \mathbb{I}$.

Sind diese Bedingungen für die Abhängigkeit der Zufallshock-Komponente erfüllt, ausgedrückt mit Hilfe der bedingten Copula C_t , dann nennen wir das n -variate Volatilitätsmodell für $(\mathbf{r}_t)_{t \in \mathbb{T}}$ *copula-basiert* oder ein *CMD-Modell*. Den n -dimensionalen Renditeprozess $(\mathbf{r}_t)_{t \in \mathbb{T}}$ wollen wir in diesem Kontext mit *CMD-Prozess* bezeichnen.

Bemerkungen:

- In dieser Definition können wir nun mit

$$H_t(\mathbf{z}_t | \mathcal{F}_{t-1}) := C_t(F_1(z_{t,1} | \mathcal{F}_{t-1}), \dots, F_n(z_{t,n} | \mathcal{F}_{t-1}) | \mathcal{F}_{t-1}) = C_t(u_{t,1}, \dots, u_{t,n} | \mathcal{F}_{t-1})$$

eine eindeutige multivariate (gegebenenfalls zeitlich variierende) bedingte Verteilung für das CMD-Modell festlegen.

- Im allgemeinen Fall haben wir bewusst die Copula C_t als zeitabhängig eingeführt. Dies bedeutet einfach, dass der zeitabhängige Vektor \mathbf{u}_t nicht für alle Zeiten $t \in \mathbb{T}$ die gleiche Copula-Funktion als multivariate Verteilungsfunktion besitzen muss. Insbesondere handelt es sich aber bei den Komponenten von \mathbf{u}_t stets um uniform

verteilte Zufallsvariablen, wie auch der Zusammenhang aus Punkt (c) verdeutlicht (vergleiche auch Satz 2.2). Für weiterführende Ansätze zu zeitabhängigen Copulas siehe [10].

- Die Zeitabhängigkeit einer Copula wird im parametrischen Fall, von dem wir in der gesamten Arbeit ausgehen, durch zeitlich veränderliche Parameter ausgedrückt. Das heißt außerdem, dass wir innerhalb des Zeitverlaufs in unserem Modellkontext nicht zwischen Copula-Familien wechseln, sondern uns nur innerhalb einer vorher festgelegten Familie bewegen werden und lediglich den Parameter variieren lassen. In Unterabschnitt 3.2.1 werden wir solche Spezifikationen der Copula-Parameter vornehmen und näher beschreiben.

Für alle nachfolgenden Betrachtungen werden wir eine bestimmte CMD-Unterklasse näher betrachten, nämlich die *Copula-GARCH-Klasse*.

Definition 3.2 (Copula-GARCH-Modell, Copula-GARCH-Prozess):

Sei $\mathbf{p} = (p_1, \dots, p_n)^\top$ beziehungsweise $\mathbf{q} = (q_1, \dots, q_n)^\top$. Ein CMD-Modell heißt *Copula-GARCH*(\mathbf{p}, \mathbf{q})-Modell, wenn es sich bei allen univariaten Marginalmodellen um GARCH(p_i, q_i)-Modelle, $i \in \mathbb{I}$, handelt. Das heißt für die Renditezeitreihe $(\mathbf{r}_t)_{t \in \mathbb{T}}$ mit Volatilität $(\sigma_t^2)_{t \in \mathbb{T}}$ gilt der Zusammenhang

$$r_{t,i} = \sigma_{t,i} \cdot z_{t,i} \quad (3.1)$$

$$\sigma_{t,i}^2 = \omega^{(i)} + \sum_{k=1}^{p_i} \alpha_k^{(i)} \cdot r_{t-k,i}^2 + \sum_{j=1}^{q_i} \beta_j^{(i)} \cdot \sigma_{t-j,i}^2 \quad (3.2)$$

für $t \in \mathbb{T}$ und $i \in \mathbb{I}$. Die (multivariate) Zeitreihe $(\mathbf{r}_t)_{t \in \mathbb{T}}$ wird dann *Copula-GARCH-Prozess* genannt, wobei $\alpha_{p_i}^{(i)} \neq 0 \neq \beta_{q_i}^{(i)}$ für alle $i \in \mathbb{I}$ zur Eindeutigkeit der Grade p und q gesetzt wird.

Copula-GARCH-Modelle sind also CMD-Modelle, bei denen der deterministische Zusammenhang zwischen Rendite und Zufallskomponente Rand für Rand durch univariate GARCH-Modelle konkretisiert sind.

Bemerkungen:

- Es sei angemerkt, dass als Lockerung der global homogenen Zeitindexmenge \mathbb{T} in dieser Modellkonstellation für jedes Randmodell auch wieder unterschiedliche Beobachtungsumfänge und eine Schätzung in mehreren Stufen möglich sind, denn die Parameter des einen Marginalmodells hängen nicht von den Parametern der übrigen

Komponenten ab (für die schätztheoretische Beschreibung siehe auch Unterabschnitt 2.3.2).

- Man macht sich leicht klar, dass bestimmte in Kapitel 1, Unterabschnitt 1.3.2, vorgestellte multivariate GARCH-Modelle zu der Copula-GARCH-Klasse gehören, andere wiederum nicht: Mit standardnormalverteilten GARCH-Residuen $(\mathbf{z}_t)_{t \in \mathbb{T}}$ ist beispielsweise das CCC-Modell äquivalent zu einem Copula-GARCH-Modell mit entsprechenden univariaten GARCH(p, q)-Marginalprozessen (p und q unabhängig von $i \in \mathbb{I}$) und einer sie verbindenden Gauß-Copula. Dagegen zählen alle allgemeinen Typen der VEC-GARCH-Modelle (auch BEKK) nicht dazu.

Bevor wir uns über die praktischen Konsequenzen der gerade behandelten Modellklasse Gedanken machen, ist es sinnvoll und auch nötig zu klären, ob ein (streng) stationärer Copula-GARCH-Prozess überhaupt existiert. Bei einem Copula-GARCH-Modell mit einer *zeitkonstanten* Copula C genügt es zu verlangen, dass für jede Komponentenzreihe $(r_{t,i})_{t \in \mathbb{T}}$, $i \in \mathbb{I}$, eines Copula-GARCH-Prozesses die entsprechenden Kriterien für Stationarität an die GARCH-Parameter gleichzeitig gelten müssen. Da wir dann mittels Sklars Theorem (Satz 2.5) eine *eindeutige* zeitkonstante multivariate (bedingte) Verteilung für die GARCH-Residuen $(\mathbf{z}_t)_{t \in \mathbb{T}}$, und damit für $(\mathbf{r}_t)_{t \in \mathbb{T}}$, festlegen können, bestimmen alle Stationaritätskriterien für jeden Randprozess gemeinsam die Stationarität des gesamten Vektorprozesses $(\mathbf{r}_t)_{t \in \mathbb{T}}$. Bei einer *zeitvariierenden* Copula allerdings, so [29], gibt es aktuell keine mit den univariaten Stationaritätskriterien vergleichbaren Ergebnisse in der Literatur. Geeignete Restriktionen für jede univariate Randzeitreihe sowie die Copula (beispielsweise an den zeitvariierenden Copula-Parameter für eine feste Copula-Familie) könnten im Zweifelsfall nicht ausreichen die gewünschte Stationarität zu garantieren, sondern es müssten gegebenenfalls komplexe komponentenübergreifende Bedingungen gestellt werden. Wir setzen daher der Einfachheit halber die Stationarität als Eigenschaft einer CMD-Zeitreihe a priori voraus.

Folgerung 3.3:

Die Konstellation für einen Copula-GARCH(\mathbf{p}, \mathbf{q})-Prozess $(\mathbf{r}_t)_{t \in \mathbb{T}}$ aus Definition 3.2 erlaubt im Falle einer parametrischen Spezifikation eine Schätzung der Parameter gemäß der mehrstufigen Maximum-Likelihood-Schätzung aus Kapitel 2 (Unterabschnitt 2.3.2). Bezeichnen wir mit $\boldsymbol{\theta}_i = (\omega^{(i)}, \alpha_1^{(i)}, \dots, \alpha_{p_i}^{(i)}, \beta_1^{(i)}, \dots, \beta_{q_i}^{(i)})^\top$, $i \in \mathbb{I}$, den Vektor der GARCH-Parameter für jedes individuelle Marginalmodell, und mit $\boldsymbol{\theta}_c$ den Copula-Parametervektor, so können wir nun durch

$$(1.) \quad \tilde{\boldsymbol{\theta}}_i := \operatorname{argmax}_{\boldsymbol{\theta}_i} \mathcal{L}_i(\boldsymbol{\theta}_i), \quad i \in \mathbb{I}, \text{ und}$$

$$(2.) \quad \tilde{\boldsymbol{\theta}}_c := \operatorname{argmax}_{\boldsymbol{\theta}_c} \mathcal{L}_c(\tilde{\boldsymbol{\theta}}_1, \dots, \tilde{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_i),$$

formal also durch die separate Maximierung der Funktionen $\mathcal{L}_1, \dots, \mathcal{L}_n$ und \mathcal{L}_c , den Mehrschritt-Schätzer $\tilde{\boldsymbol{\theta}} := (\tilde{\boldsymbol{\theta}}_1^\top, \dots, \tilde{\boldsymbol{\theta}}_n^\top, \tilde{\boldsymbol{\theta}}_c^\top)^\top$ für den Gesamt-Parametervektor ermitteln. Dabei hat die Log-Likelihood-Funktion die Gestalt

$$\begin{aligned}
 \mathcal{L}(\boldsymbol{\theta}) &:= \sum_{i=1}^n \mathcal{L}_i(\boldsymbol{\theta}_i) + \mathcal{L}_c(\boldsymbol{\theta}) \\
 &:= \sum_{i=1}^n \sum_{t=1}^T \ell_{t,i}(\boldsymbol{\theta}_i) + \sum_{t=1}^T \ell_{t,c}(\boldsymbol{\theta}) \\
 &:= \sum_{i=1}^n \sum_{t=1}^T \ln(f_{i|t-1}(z_{t,i}; \boldsymbol{\theta}_i)) + \\
 &\quad + \sum_{t=1}^T \ln(c_{|t-1}(F_{1|t-1}(z_{t,1}; \boldsymbol{\theta}), \dots, F_{n|t-1}(z_{t,n}; \boldsymbol{\theta}); \boldsymbol{\theta})),
 \end{aligned} \tag{3.3}$$

die wir bereits aus Kapitel 2 (Unterabschnitt 2.3.2) kennen. Im kommenden Unterabschnitt, der Simulationsstudie, werden wir nun die Dimension, die Sample-Size, die Copula-Familien und schließlich auch die Log-Likelihood-Funktionen konkreter angeben.

3.2. Simulationsstudie

Nun möchten wir im Rahmen einer Simulationsstudie herausfinden, ob die eben angedeutete Parameterschätzung in mehreren Schritten tatsächlich bei mehrfacher Durchführung des gleichen Simulationsexperiments die beschriebene Theorie über die Asymptotik des Mehrschritt-Schätzers $\tilde{\boldsymbol{\theta}}$ (vergleiche Kapitel 2, Unterabschnitt 2.3.2) bestätigt.

In der Simulationsstudie haben wir uns für eine einheitliche Sample-Size von $T := 2500$ entschieden, so dass $\mathbb{T} := \{1, \dots, 2500\}$ für jede der insgesamt $n = 2$ Zeitreihen $(r_{t,1})_{t \in \mathbb{T}}$ und $(r_{t,2})_{t \in \mathbb{T}}$ gelte. Die Simulationsstudie wurde insgesamt $N = 1000$ Mal durchgeführt.

Im nachfolgenden Abschnitt stellen wir zunächst den verwendeten Modellaufbau für die Renditezeitreihe $(\mathbf{r}_t)_{t \in \mathbb{T}}$ mit $\mathbf{r}_t = (r_{t,1}, r_{t,2})^\top$ inklusive der getroffenen Verteilungsannahmen für die GARCH-Residuen $\mathbf{z}_t = (z_{t,1}, z_{t,2})^\top$ vor. Dies umfasst zwei GARCH(1,1)-Prozesse mit leicht unterschiedlichen Parametern für die Simulation. Anschließend behandeln wir insgesamt vier Kombinationen von verwendeten Copula-Spezifikationen (zwei zeitkonstante, zwei zeitvariierende). Die letzten beiden Abschnitte befassen sich dann mit einer groben Formulierung der verwendeten Algorithmen sowie der Präsentation der Schätzergebnisse im Rahmen der Simulationsstudie.

3.2.1. Modellaufbau

Marginalmodelle

Wie erwähnt, handelt es sich bei den Randmodellen unserer Simulationsstudie um zwei univariate Copula-GARCH $((1, 1)^\top, (1, 1)^\top)$ -Modelle mit je standardnormalverteilten Residuen, das heißt

$$r_{t,1} = \sigma_{t,1} \cdot z_{t,1} \quad (3.4)$$

$$\sigma_{t,1}^2 = \omega^{(1)} + \alpha_1^{(1)} \cdot r_{t-1,1}^2 + \beta_1^{(1)} \cdot \sigma_{t-1,1}^2 \quad (3.5)$$

$$z_{t,1} \sim \mathcal{N}(0, 1) =: \Phi, \quad u_{t,1} := \Phi(z_{t,1} | \mathcal{F}_{t-1}) \quad (3.6)$$

sowie

$$r_{t,2} = \sigma_{t,2} \cdot z_{t,2} \quad (3.7)$$

$$\sigma_{t,2}^2 = \omega^{(2)} + \alpha_1^{(2)} \cdot r_{t-1,2}^2 + \beta_1^{(2)} \cdot \sigma_{t-1,2}^2 \quad (3.8)$$

$$z_{t,2} \sim \mathcal{N}(0, 1) =: \Phi, \quad u_{t,2} := \Phi(z_{t,2} | \mathcal{F}_{t-1}). \quad (3.9)$$

Als Parameter für die Simulation haben wir in unserer Studie

$$(\omega^{(1)}, \omega^{(2)})^\top = (1,00 \cdot 10^{-6}, 1,00 \cdot 10^{-6})^\top,$$

$$(\alpha_1^{(1)}, \alpha_1^{(2)})^\top = (0,19, 0,15)^\top,$$

$$(\beta_1^{(1)}, \beta_1^{(2)})^\top = (0,80, 0,83)^\top$$

gewählt (siehe auch Anhang A).

Dabei wird die gegenseitige Abhängigkeit durch die bedingte 2-Copula

$$C_t(\mathbf{u}_t | \mathcal{F}_{t-1}) = C_t(\Phi(z_{t,1} | \mathcal{F}_{t-1}), \Phi(z_{t,2} | \mathcal{F}_{t-1}) | \mathcal{F}_{t-1})$$

als bedingte Verteilung des Zufallsvektors $\mathbf{u}_t := (u_{t,1}, u_{t,2})^\top$ bestimmt. Diese spezifizieren wir als Nächstes.

Copula-Spezifikation

Wir möchten nun einige mögliche Spezifizierungen der verwendeten Copula-Funktionen angeben: je zwei Vertreter der elliptischen Copula-Klasse mit jeweils zeitkonstantem und zeitvariierendem Korrelationsparameter.

(i) Zeitkonstante Copula

Zur Wahl stehen die beiden elliptischen Copulas, die bivariate Gauß-Copula C_ρ und die bivariate t-Copula $C_{\rho,\nu}$. Diese haben jeweils die Copula-Dichten

(a) Gauß-Copula:

$$c(u_{t,1}, u_{t,2}; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right), \quad (3.10)$$

(b) t-Copula:

$$c(u_{t,1}, u_{t,2}; \rho, \nu) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \left(1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}}, \quad (3.11)$$

unabhängig für alle $t \in \mathbb{T}$, siehe auch Gleichungen (2.11) beziehungsweise (2.12) in Kapitel 2 (aus dem Beispiel auf Seite 39), wobei ρ der lineare Korrelationskoeffizient und ν der Freiheitsgrad-Parameter für die t-Copula sei. Obige Copulas (beziehungsweise Copula-Dichten) sind als zeitkonstant angegeben. Hier sind konkret die gewählten Parameter $\theta = \rho = 0,7$ beziehungsweise $\theta = (\rho, \nu)^\top = (0,7; 6)^\top$ zu schätzen gewesen.

(ii) Zeitabhängige Copula: Time Varying Correlation (TVC)

Einer Spezifikation von [37] vergleichbar, definieren wir den zeitvariierenden Korrelationsparameter ρ_t der Copula-Dichten (3.10) respektive (3.11) mittels einer Gleichung der allgemeinen Gestalt

$$\rho_t = (1 - a - b)\rho_{\text{TVC}} + a\xi_{t-1} + b\rho_{t-1}, \quad (3.12)$$

wobei die Einflussvariable

$$\xi_t := \frac{\sum_{k=0}^{m-1} z_{t-k,1} \cdot z_{t-k,2}}{\sqrt{(\sum_{k=0}^{m-1} z_{t-k,1}^2) \cdot (\sum_{k=0}^{m-1} z_{t-k,2}^2)}}, \quad m \geq n,$$

die (empirische) Korrelation zwischen den Residuen der vergangenen $m - 1$ Perioden darstelle. Mit dem zeitvariierenden Copula-Parameter ρ_t erhalten wir nun für die obigen Copula-Dichten die zeitabhängigen Versionen:

(c) Gauß-Copula:

$$c(u_{t,1}, u_{t,2}; \rho_{\text{TVC}}, a, b) = \frac{1}{2\pi\sqrt{1-\rho_t^2}} \cdot \exp\left(-\frac{x^2 - 2\rho_t xy + y^2}{2(1-\rho_t^2)}\right), \quad t \in \mathbb{T}, \quad (3.13)$$

(d) t-Copula:

$$c(u_{t,1}, u_{t,2}; \rho_{\text{TVC}}, a, b, \nu) = \frac{1}{2\pi\sqrt{1-\rho_t^2}} \cdot \left(1 + \frac{x^2 - 2\rho_t xy + y^2}{\nu(1-\rho_t^2)}\right)^{-\frac{\nu+2}{2}}, \quad t \in \mathbb{T}. \quad (3.14)$$

Als später zu schätzende Parameter haben wir $\rho_{\text{TVC}} = 0,7$, $a = 0,2$ und $b = 0,5$ gewählt, für die gesamte Simulationsstudie setzten wir $m = 5$ als festen und *nicht zu schätzenden* Parameter.

Der Freiheitsgrad-Parameter ν der t-Copula aus Gleichung (3.14) wurde der Einfachheit halber in jedem Fall als zeitkonstant angenommen, als Wert in der Simulationsstudie wurde $\nu = 6$ gewählt.

Bleiben wir weiterhin bei der Bezeichnung $\boldsymbol{\theta}_c$ für den Parametervektor der gewählten Copula, so erhalten wir jetzt insgesamt folgende vier Kombinationen für die Spezifikationen:

$$(a) \quad \text{Gauß-Copula ohne TVC:} \quad \boldsymbol{\theta}_c = \rho \quad (3.15)$$

$$(b) \quad \text{t-Copula ohne TVC:} \quad \boldsymbol{\theta}_c = (\rho, \nu)^\top \quad (3.16)$$

$$(c) \quad \text{Gauß-Copula mit TVC:} \quad \boldsymbol{\theta}_c = (\rho_{\text{TVC}}, a, b)^\top \quad (3.17)$$

$$(d) \quad \text{t-Copula mit TVC:} \quad \boldsymbol{\theta}_c = (\rho_{\text{TVC}}, a, b, \nu)^\top \quad (3.18)$$

Auf die Buchstaben (a) bis (d) wird im folgenden Kontext weiterhin verwiesen.

3.2.2. Algorithmen zu Simulation und Schätzung

Mittels der Inversionsmethode (Kapitel 2, Algorithmus 2.13) erhalten wir stets zwei Zufallszahlen zu gegebener Copula, die mit oder ohne TVC spezifiziert werden kann.

Doch wie sieht die Simulation Schritt für Schritt aus? Der Modellaufbau im letzten Abschnitt ist bewusst in der Reihenfolge der mehrstufigen Schätzung formuliert worden, die Simulation erfolgt dagegen in umgekehrtem Ablauf. Dabei unterscheiden wir folgende beiden Algorithmen, je nach Zeitvariation oder Zeitkonstanz der Copula-Parameter.

Algorithmus 3.4 (Simulation mit beliebiger Copula, ohne TVC):

Sei C die zuvor festgelegte Copula (Gauß- oder t-Copula).

- (1.) Generiere zum festen Parameter ρ (und ν) eine Folge von Realisationen der Zeitreihe $\mathbf{u}_t = (u_{t,1}, u_{t,2})^\top$ so, dass $\mathbf{u}_t | \mathcal{F}_{t-1} \sim C$ für $t = 1, \dots, 2500$.
- (2.) Mittels $z_{t,i} := \Phi^{-1}(u_{t,i})$, $i = 1, 2$, transformiere jede der beiden Komponenten für $t = 1, \dots, 2500$ (damit erhalten wir alle normalverteilten GARCH-Residuen).
- (3.) Entwickle nun die beiden GARCH(1,1)-Prozesse zu gegebenen Parametern: Erzeuge dazu nacheinander alle Volatilitäten $\sigma_{t,i}^2$, $i = 1, 2$, gemäß (3.5) beziehungsweise (3.8) für $t = 1, \dots, 2500$ und ermittle anschließend die simulierte Renditezeitreihe durch (3.4) beziehungsweise (3.7).

Algorithmus 3.5 (Simulation mit beliebiger Copula, mit TVC):

Sei C die zuvor festgelegte Copula (Gauß- oder t-Copula).

- (1.) Generiere zum festen Parameter ρ_{TVC} (und ν) zuerst $m = 5$ Initialwerte von Realisationen der Zeitreihe $\mathbf{u}_t^{\text{Start}} = (u_{t,1}^{\text{Start}}, u_{t,2}^{\text{Start}})^\top$ so, dass $\mathbf{u}_t^{\text{Start}} | \mathcal{F}_{t-1} \sim C$ für $t = 1, \dots, 5$.
- (2.) Mittels $z_{t,i} := \Phi^{-1}(u_{t,i}^{\text{Start}})$, $i = 1, 2$, transformiere jede der beiden Komponenten für $t = 1, \dots, 5$.
- (3.) Für jeden übrigen Zeitpunkt $t = 6, \dots, 2500$ erzeuge dann jeweils abhängig von den Initialwerten aus Schritt (1.) und den restlichen Parametern a und b abwechselnd
 - (i) einmal die Korrelation ρ_t gemäß (3.12) sowie
 - (ii) einmal die restlichen Paare $\mathbf{u}_t^{\text{Rest}} = (u_{t,1}^{\text{Rest}}, u_{t,2}^{\text{Rest}})^\top$ mit $\mathbf{u}_t^{\text{Rest}} | \mathcal{F}_{t-1} \sim C$ zum eben berechneten Copula-Parameter ρ_t .

Anschließend führe stets auch die Transformation $z_{t,i} := \Phi^{-1}(u_{t,i})$, $i = 1, 2$ durch (damit erhalten wir alle normalverteilten GARCH-Residuen).

- (4.) Entwickle den GARCH(1,1)-Prozess zu gegebenen Parametern: Erzeuge dazu nacheinander die Volatilitäten $\sigma_{t,i}^2$, $i = 1, 2$, gemäß (3.5) beziehungsweise (3.8) für $t = 1, \dots, 2500$ und erhalte anschließend die simulierte Renditezeitreihe durch (3.4) beziehungsweise (3.7).

Die parametrische Schätzung wurde bei unserer Simulationsstudie stets mit dem Mehrschritt-Maximum-Likelihood-Verfahren durchgeführt. Dazu wurden in den vier Optionen (a) bis (d) die folgenden Log-Likelihood-Funktionen (gemäß Gleichung 3.3) verwendet:

$$(a) \quad \mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}_1(\boldsymbol{\theta}_1) + \mathcal{L}_2(\boldsymbol{\theta}_2) + \mathcal{L}_c(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho)$$

$$(b) \quad \mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}_1(\boldsymbol{\theta}_1) + \mathcal{L}_2(\boldsymbol{\theta}_2) + \mathcal{L}_c(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho, \nu)$$

$$(c) \quad \mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}_1(\boldsymbol{\theta}_1) + \mathcal{L}_2(\boldsymbol{\theta}_2) + \mathcal{L}_c(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{\text{TVC}}, a, b)$$

$$(d) \quad \mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}_1(\boldsymbol{\theta}_1) + \mathcal{L}_2(\boldsymbol{\theta}_2) + \mathcal{L}_c(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{\text{TVC}}, a, b, \nu)$$

Dabei verläuft der erste Schritt der mehrstufigen Schätzung, die Schätzung der GARCH-Parameter, jeweils identisch, nämlich nach der Maximum-Likelihood-Methode aus Unterabschnitt 1.2.2 in Kapitel 1. Der zweite Hauptschritt, die Schätzung der Copula-Parameter bedingt auf die GARCH-Schätzer aus dem vorigen Schritt, erfolgt dagegen je nach Fall leicht unterschiedlich: sind bei (a) und (b) (ohne TVC) jeweils nur die zeitkonstanten Copula-Parameter zu schätzen, so müssen bei (c) und (d) zusätzlich die TVC-Parameter

geschätzt werden, wobei diese die Spezifikation der zeitabhängigen Copula-Parameter bestimmen.

3.2.3. Ergebnisse der Simulationsstudie

In Anhang A sind die mittleren Schätzwerte und Standardabweichungen sowie die Ergebnisse des Tests auf Normalität aus der Simulationsstudie mit Optionen (a) bis (d), bei $N = 750$ Durchläufen zunächst übersichtlich in Tabellenform aufgeführt. Anschließend sind alle Dichte-Plots für jeden einzelnen Schätzer angefügt. Die Theorie aus Satz 2.21, Kapitel 2, sagt eine asymptotische Normalität für den Mehrschritt-Schätzer voraus. Um dies zu testen, haben wir jeweils das Verfahren nach Shapiro verwendet. Im Folgenden diskutieren wir kurz die Ergebnisse.

Die Optionen (a) und (b), beide ohne die TVC-Spezifikation, liefern trotz der geringen Sample-Size von $T = 2500$ recht genaue Ergebnisse bei der Parameterschätzung, wie man an den Standardfehlern erkennen kann (siehe Tabelle A.2). Der Normalitätstest deutet jeweils sehr stark auf eine Normalverteilung hin, mit Ausnahme der sehr kleinen GARCH-Parameter $\omega^{(i)}$, $i = 1, 2$. Dies kann man auch schon auf den ersten Blick über die Dichte-Plots bei fast allen Parametern außer diesen erkennen. Des Weiteren rufen die TVC-Parameter a und b recht niedrige p-Werte der entsprechenden Teststatistik hervor, ebenso wie in den t-Copula-Varianten die Marginalparameter $\alpha_1^{(2)}$ (Option (b)) beziehungsweise $\beta_1^{(2)}$ (Option (d)) und jeweils der Freiheitsgrad-Parameter ν .

Bei den TVC-Ansätzen in den Optionen (c) und (d) sind zusätzlich noch je zwei Parameter a und b zu schätzen gewesen (vergleiche Gleichungen (3.17) und (3.18)). Hier war die Schätzung im Vergleich zu den zeitkonstanten Copula-Parametern wesentlich ungenauer (siehe Tabelle A.2). Ebenso dauerte die Simulation der Zeitreihe und Optimierung der jeweiligen Log-Likelihood-Funktion einiges länger, da für jede Zeit $t = 5, \dots, 2500$ eine eigene Copula-Dichte (zum Parameter ρ_t gemäß Gleichung (3.12)) ausgewertet werden musste.

3.2.4. Anwendung auf reale Daten

In unserer praktischen Anwendung haben wir Zeitreihen von zwei DAX-Unternehmen, der Commerzbank und der Deutschen Bank, miteinander verglichen. Die Originaldaten lagen als Preisindizes vor, weshalb beide Zeitreihen zunächst zu Renditezeitreihen transformiert wurden. Anschließend unterwarfen wir diese Zeitreihen dem gleichen mehrstufigen Schätzverfahren wie zuvor die simulierten Daten in der Studie.

Im Gegensatz zur Simulation wissen wir bei realen Daten nicht, ob wir bei den Optionen (a) bis (d) jeweils eine richtige Wahl getroffen haben. Eine gängige Art und Weise, die

Richtigkeit der Spezifikation zu testen, wäre beispielsweise die jeweiligen Log-Likelihood-Maxima zu vergleichen. Diese betragen bei den vier Optionen:

(a) $\mathcal{L}(\tilde{\theta}) = 813,94$

(b) $\mathcal{L}(\tilde{\theta}) = 870,52$

(c) $\mathcal{L}(\tilde{\theta}) = 845,76$

(d) $\mathcal{L}(\tilde{\theta}) = \underline{912,42}$

Insgesamt sind also die t-Copula-Optionen den Gauß-Copula-Optionen (Optionen (b) und (d) gegen (a) beziehungsweise (c)) vorzuziehen. Die Spezifikation mit der zeitvariierenden t-Copula (Option (d)) würde vor allen anderen bevorzugt, denn bei dieser wird der höchste Log-Likelihood-Wert erreicht.

Das gesamte Schätzergebnis ist in Tabelle B.1 in Anhang B aufgeführt, ebenso wie die Plots der Preis- und Renditezeitreihen sowie des zeitabhängigen Korrelationskoeffizienten ρ_t aus den TVC-Spezifikationen (c) und (d). Bei Betrachtung des Zeitverlaufs stellt man fest, dass die Korrelation über die Jahre hinweg stets recht hoch war (die Durchschnitte liegen bei etwa 0,6), allerdings auch eine relativ große Schwankung aufwies (etwa zwischen 0,3 und 0,8, vergleiche auch Abbildung B.2).

Wir wollen nun zum Abschluss die Vor- und Nachteile der neu entwickelten Copula-GARCH-Klasse auf kritische Weise betrachten und gegenüberstellen.

3.3. Diskussion: Vorzüge und Schwachpunkte

Wir haben in diesem Kapitel eine neue Klasse von multivariaten Volatilitätsmodellen kennengelernt, bei der die univariaten Randmodelle mit Hilfe des Copula-Konzepts zu einem Gesamtmodell vereinigt werden. Dabei haben wir uns der Grundlagen und -begriffe bedient, die zuvor in den ersten beiden Kapiteln eingeführt wurden. Insbesondere galt dies für die Parameterschätzung, welche in in Kapitel 1 zunächst für GARCH-Modelle im Univariaten ausführlich behandelt wurde, bevor in Kapitel 2 die entsprechenden Schätzverfahren für beliebige copula-basierte Modelle erläutert wurden.

Anschließend haben wir unsere Simulations- und Schätzmethoden in einer Simulationsstudie implementiert und auf ihre Tauglichkeit hin geprüft, die Theorie über das asymptotische Verhalten der einzelnen Schätzer zu belegen. Dies ist tatsächlich zu einem sehr großen Teil gelungen, wie sich herausgestellt hat.

Es stellt sich allerdings die Frage, wie sinnvoll die vereinfachten Modellannahmen im Vergleich zur Realität sind. Mit dieser Problemstellung haben sich bereits mehrere Autoren

befasst, so ist unter anderem in [16], [31] oder [1] von möglichen erweiterten Modellbedingungen die Rede, die wir kurz auch selbst auffristen wollen:

- Ähnlich wie in Kapitel 1 erwähnt, könnte man, um den Leverage-Effekt zu berücksichtigen, von einer Aufteilung des Zufallsschocks in unterschiedlich gewichtete positive und negative Komponenten ausgehen.
- Des Weiteren ist eine alternative Verteilung wie zum Beispiel eine Student-t-Verteilung oder die sogenannte *schiefe t-Verteilung* (kurz ST-Verteilung) nach [14] statt der Normalitätsannahme durchaus vorstellbar und sogar realitätsnäher, wie wir uns in Kapitel 1 bereits klar gemacht haben.
- Für die verwendeten Copula-Funktionen gilt Vergleichbares: es sind auch Vertreter aus der archimedischen Familie möglich, jedoch beachte man hierbei, dass die Parameter nicht mehr den Korrelationskoeffizienten entsprechen. Möchte man hier gleichzeitig die Abhängigkeitsstruktur zeitabhängig modellieren, macht die TVC-Spezifikation keinen Sinn mehr. Dagegen kann man auf Rangkorrelationskoeffizienten wie *Kendall's Tau* oder *Spearman's Rho* zurückgreifen, für die es reine copula-basierte Darstellungen gibt (vergleiche zum Beispiel [25] oder [19]), und diese dann zeitlich variieren lassen.

Wie in Kapitel 2 bereits erwähnt, bleibt auch noch bei der copula-basierten Volatilitätsmodellierung eines der aktuellen Hauptprobleme erhalten: trotz der Vielzahl umfassender Entwicklungen gibt es meist keine allgemeine Erweiterung von bivariaten Modellen auf höhere Dimensionen.

Es sei ferner angemerkt, dass man neben der parametrischen Schätzung auch sehr häufig semi- oder nicht-parametrische Schätzverfahren gebraucht, um beispielsweise überhaupt erst einmal die jeweiligen Verteilungsfamilien der Randmodelle zu spezifizieren. Zwar geht dieser Ansatz auf Kosten der Genauigkeit, dafür werden aber Falsch-Spezifikationen von vornherein vermieden. Entscheidet man sich dennoch für die parametrische Schätzung, so sind weitergehende Tests (allen voran Güte-Tests) von hoher Bedeutung, um gegebenenfalls falschen Spezifikationen auf die Spur zu kommen.

Wir schließen nun im nachfolgenden Kapitel die Arbeit mit einer Zusammenfassung und einem Fazit ab.

Schlusswort

Zusammenfassung

In dieser Bachelorarbeit wurden zunächst Einblicke in zwei von einander isolierte Themengebiete gewährt, um den Leser inhaltlich auf das Hauptthema, die Einführung von *Copula-GARCH-Modellen*, vorzubereiten.

Im ersten Kapitel befassten wir uns mit der Thematik der finanzökonometrischen Modellierung. Dort wurden zunächst Grundbegriffe eingeführt, allgemeine, empirisch belegte Sachverhalte rund um Finanzzeitreihen (Stylized Facts) illustriert und daran anknüpfend die sogenannten bedingt-heteroskedastischen Modelle (oder Volatilitätsmodelle) vorgestellt. Unser Hauptaugenmerk lag dabei im Univariaten auf den (G)ARCH-Modellen von Engle (1982) beziehungsweise Bollerslev (1986), die wir kurz veranschaulicht haben und bei denen wir ausführlich auf die Parameterschätzung eingegangen sind. Ferner befassten wir uns mit einem Überblick über die multivariaten Erweiterungen dieser Volatilitätsmodelle und gaben als Abschluss des Kapitels einen Ausblick auf weitere Erweiterungsmöglichkeiten, die die realen finanzwirtschaftlichen Phänomene noch besser nachzubilden versuchen.

Das anschließende zweite Kapitel behandelte die Einführung in die Welt der Copula-Funktionen. Viele finanzwirtschaftliche Zweige der Mathematik haben mittlerweile das Konzept vollständig verinnerlicht, da es sehr flexible Implikationen für die Praxis liefert. Grundlegend können wir die Information, die eine multivariate Verteilungsfunktion eines Zufallsvektors trägt, in zwei Hauptinformationen aufsplitten: Diejenige der univariaten Randverteilungen und in die bloße Abhängigkeitsstruktur. Letztere repräsentieren nun die Copula-Funktionen: Wählen wir eine willkürliche Copula, können wir entweder zusammen mit beliebigen univariaten Randverteilungen eine neue multivariate Verteilungsfunktion erschaffen oder zu gegebener multivariater Verteilung und univariaten Rändern die Copula vollständig extrahieren. Je nach Zweck hat beides seine praktischen Vorzüge. In der zweiten Hälfte des Kapitels haben wir uns deswegen auf die möglichen Anwendungsgebiete, Simulation von abhängigen Zufallspaaren bei bivariaten Copulas sowie der parametrischen Schätzung – allgemein für zeitkonstante und für zeitabhängige Zufallsvektoren (multivariate Zeitreihen) – aus Schätztheoriesicht konzentriert.

Die Grundlagen, die diese beiden Kapitel für uns geschaffen haben, ermöglichten uns im abschließenden dritten Kapitel eine neue Sichtweise auf das Konstruieren von multivariaten Volatilitätsmodellen: Mit den nun erlangten Werkzeugen der Copulas können wir jetzt beliebige (völlig inhomogene) univariate Modelle parallel, aber getrennt, beobach-

ten und deren Abhängigkeit mittels einer Copula-Funktion modellieren. Dazu bedienten wir uns der mehrstufigen Parameterschätzung (Mehrschritt-Schätzung), die wir zuvor in Kapitel 2 ausführlich beschrieben haben. Nachdem wir zuerst die Modellklasse der copula-basierten Volatilitätsmodelle vorgestellt hatten, führten wir danach ein eigenes sogenanntes Copula-GARCH-Modell ein, bei dem alle Marginalmodelle dem GARCH-Konzept folgten. Als Illustration konstruierten wir anschließend ein eigenes bivariates Copula-GARCH-Modell, das wir in einer Simulationsstudie umfassend getestet haben. Die Mehrschritt-Schätzung wurde hierbei per Hand implementiert und auf die vier Optionen (a) zeitkonstante Gauß-Copula, (b) zeitkonstante t-Copula sowie (c) zeitabhängige Gauß-Copula und (d) zeitabhängige t-Copula angewandt. Dabei hatten wir als einzigen zeitvariierenden Copula-Parameter gerade den linearen Korrelationskoeffizienten gewählt, den wir einer Time-Varying-Correlation-Spezifikation zuordneten. Alle Schätzungen lieferten zuverlässige und recht genaue Ergebnisse und auch die jeweiligen Mehrschritt-Schätzer wiesen bereits Andeutungen der Asymptotik auf, die in Kapitel 2 theoretisch vorhergesagt wurde.

Auf den gleichen Schätzoptionen basierend wurden auch reale Daten verwendet, um die implementierte Methodik in der Praxis auszuprobieren: Hierbei entschieden wir uns für Werktagesswerte von Renditezeitreihen der beiden DAX-Unternehmen Commerzbank und Deutsche Bank aus einem Zeitraum von jeweils knapp 15 Jahren. Dabei stellte sich heraus, dass die Korrelation als Copula-Parameter recht hoch, aber auch sehr schwankend über die Zeit hinweg war, was auf jeden Fall für eine TVC-Spezifikation sprach. Außerdem schien die t-Copula die geeignetere und der Gauß-Copula vorzuziehende Alternative zu sein. Die Ergebnisse haben wir zur Ergänzung als Grafiken und Tabellen im Anhang verfügbar gemacht.

Aktuelles und Ausblick

Momentan ist A. Patton der wohl ambitionierteste unter den Autoren, welche die Welt der Copulas und GARCH-Modellierung in Einklang bringen möchten. So widmete er etwa seine gesamte Dissertation ([28]) diesem Themenbereich, aus der mitunter auch der lesenswerte Artikel [29] mit einer Anwendung auf Wechselkurse stammt. Auch die Publikation [30], an die sich der schätztheoretische Teil unseres Hauptthemas größtenteils anlehnte, ist in dieser Abhandlung zu finden. Letztere Arbeit wird außerdem in [16] als Grundlage für ein eigenes Copula-GARCH-Modell mit einer ST-Verteilung verwendet, aus der unsere Idee mit der zeitabhängigen Korrelation (TVC) herrührte. Ein weiteres Essay von Patton ([31]) gibt schließlich einen Überblick über copula-basierte Modelle für Finanzzeitreihen.

Interessante Aspekte der Copulas in Verbindung mit GARCH-Modellierung werden des Weiteren im Aufsatz [1] aufgezeigt, in dem eine effiziente Bayes-Schätzung beschrieben

und durchgeführt wird (auch mit archimedischen statt elliptischen Copula-Funktionen); ebenfalls informativ ist [13] mit einem recht aktuellen Einblick (Stand 2009) in dynamische Copula-Modelle. Weiterführende Anwendung von Copula-GARCH-Modellen auf Deutsch findet man beispielsweise in [38] mit Blick auf die Schätzung von Risikomaßen wie etwa dem Value-at-Risk für ein Risikoportfolio. Allgemeines über Copulas und auch ihre Anwendung in der Praxis jenseits der Zeitreihenanalyse liefern in der deutschsprachigen Literatur unter anderem die Arbeiten [27] und [32].

Abschließend bleibt festzuhalten, dass wir die copula-basierte GARCH-Modellierung als sehr mächtiges und höchstinteressantes multivariates Konzept kennengelernt haben, das in dieser Bachelorarbeit nur sehr vereinfacht dargestellt werden konnte. In der Ökonometrie genießen Copulas aktuell ohnehin überaus hohe Aufmerksamkeit, da sie sehr vielseitig und flexibel einsetzbar und mit ihnen komplexe Modelle konstruierbar sind. Noch dazu machen gerade diese Vorzüge die Implementierung von dynamischen Copula-Modellen so einfach. Es ist zwar abzusehen und sehr wahrscheinlich, dass es in naher Zukunft noch viel über die aktuellen Entwicklungen (beispielsweise von Patton) zu lesen geben wird, doch ebenso könnte der allgemeine Forschungstrend bald verflachen, wenn beispielsweise das Interesse an der bivariaten Modellierung ausgeschöpft ist und die Forschung nicht schnell genug neue Entwicklungen zu höheren Dimensionen nachführen kann. An einigen Universitäten mit Arbeitsgebieten in der statistischen Datenanalyse, die sich auf dem aktuellen Forschungsstand befinden (oder sich sogar selbst mit eigener Forschungsarbeit am Fortschritt beteiligen), wird das Lehrangebot seit einigen Jahren um diese neuen Themengebiete erweitert. So werden zwar schon länger Lehrveranstaltungen wie Vorlesungen oder Seminare zu Copulas angeboten, doch es bleibt mit Spannung abzuwarten, ob sich demnächst auch Schnittstellen zur klassischen (und Finanz-)Zeitreihenanalyse herausbilden werden, die den Studenten höherer Semester einen interessanten Zugang zur Thematik der copula-gestützten Finanzzeitreihenmodellierung gewähren könnten. Die Zahl adäquater (auch als Lehrbücher tauglicher) Werke wie etwa [23], worin sowohl Zeitreihenanalyse und das Copula-Konzept in einer Fassung behandelt werden, ist jedenfalls noch überschaubar.

Anhang A

Ergebnisse der Simulationsstudie

Rahmenbedingungen		Parameter	
Durchlaufzahl N	750	$\omega^{(1)}$	1,00 e-06
Sample-Size T	2500	$\alpha_1^{(1)}$	0,19
Dimension n	2	$\beta_1^{(1)}$	0,80
Fixer Lag m (TVC)	5	$\omega^{(2)}$	1,00 e-06
Marginalmodelle	GARCH(1,1)	$\alpha_1^{(2)}$	0,15
verwendete Copulas	Gauß- und t-	$\beta_1^{(2)}$	0,83
		ρ_{TVC}	0,70
		a	0,20
		b	0,50
		ν	6,00

Tabelle A.1.: Überblick über die ursprünglichen Voreinstellungen und gewählten Parameter der Simulationsstudie

Erläuterungen zu den Abbildungen A.1 bis A.4:

- rote Kurve: Empirische Dichte der Daten
- blaue Kurve: rekonstruierte Normaldichte mit Mittelwert und Standardabweichung aus den Daten
- rote vertikale Linie: Lage des ursprünglich gewählten Parameters
- blaue vertikale Linie: Lage des Schätzer-Mittelwertes

Parameter	Mittelwert	Standardfehler	p-Wert	Parameter	Mittelwert	Standardfehler	p-Wert
(a) <i>Gauß-Copula ohne TVC</i>							
$\omega^{(1)}$	1,068846 e-06	2,446547 e-07	1,461 e-09	$\omega^{(1)}$	1,093134 e-06	2,507368 e-07	1,529 e-10
$\alpha_1^{(1)}$	0,1937549	0,01936174	0,9372	$\alpha_1^{(1)}$	0,1932757	0,01926746	0,3501
$\beta_1^{(1)}$	0,7945263	0,01798111	0,1883	$\beta_1^{(1)}$	0,7941838	0,01766066	0,3621
$\omega^{(2)}$	1,074907 e-06	2,604992 e-07	6,749 e-13	$\omega^{(2)}$	1,087447 e-06	2,534144 e-07	3,225 e-08
$\alpha_1^{(2)}$	0,1531858	0,01726883	0,1728	$\alpha_1^{(2)}$	0,1525832	0,01709565	0,0009753
$\beta_1^{(2)}$	0,8245987	0,01884800	0,0003280	$\beta_1^{(2)}$	0,8248152	0,01810497	0,1014
ρ	0,7005906	0,01044730	0,2530	ρ	0,7006816	0,01193121	0,2113
ν				ν	6,1427460	0,95006850	2,756 e-15
(b) <i>t-Copula ohne TVC</i>							
(c) <i>Gauß-Copula mit TVC</i>							
$\omega^{(1)}$	1,076985 e-06	2,486302 e-07	1,279 e-10	$\omega^{(1)}$	1,079848 e-06	2,454298 e-07	2,046 e-09
$\alpha_1^{(1)}$	0,1935273	0,02026660	0,3031	$\alpha_1^{(1)}$	0,1957637	0,01922242	0,8848
$\beta_1^{(1)}$	0,7940024	0,01937054	0,7611	$\beta_1^{(1)}$	0,7925950	0,01779566	0,2063
$\omega^{(2)}$	1,085872 e-06	2,543160 e-07	7,174 e-08	$\omega^{(2)}$	1,080872 e-06	2,697988 e-07	9,46 e-14
$\alpha_1^{(2)}$	0,1527042	0,01811741	0,6042	$\alpha_1^{(2)}$	0,1546314	0,01769570	0,2813
$\beta_1^{(2)}$	0,8245413	0,01875309	0,6107	$\beta_1^{(2)}$	0,8232625	0,01873354	0,04694
ρ_{TVC}	0,6992754	0,01665089	0,1735	ρ_{TVC}	0,6997186	0,01865507	0,008602
a	0,2035578	0,03424519	0,05197	a	0,2021585	0,036603089	4,709 e-05
b	0,4876829	0,08651934	5,673 e-07	b	0,4885435	0,09119870	0,0001203
ν				ν	6,2252680	0,95738350	1,057 e-14
(d) <i>t-Copula mit TVC</i>							

Tabelle A.2.: Mittlere Schätzergebnisse und Standardfehler für die Parameter aus Option (a) bis (d) sowie p-Werte des Shapiro-Tests auf Normalität

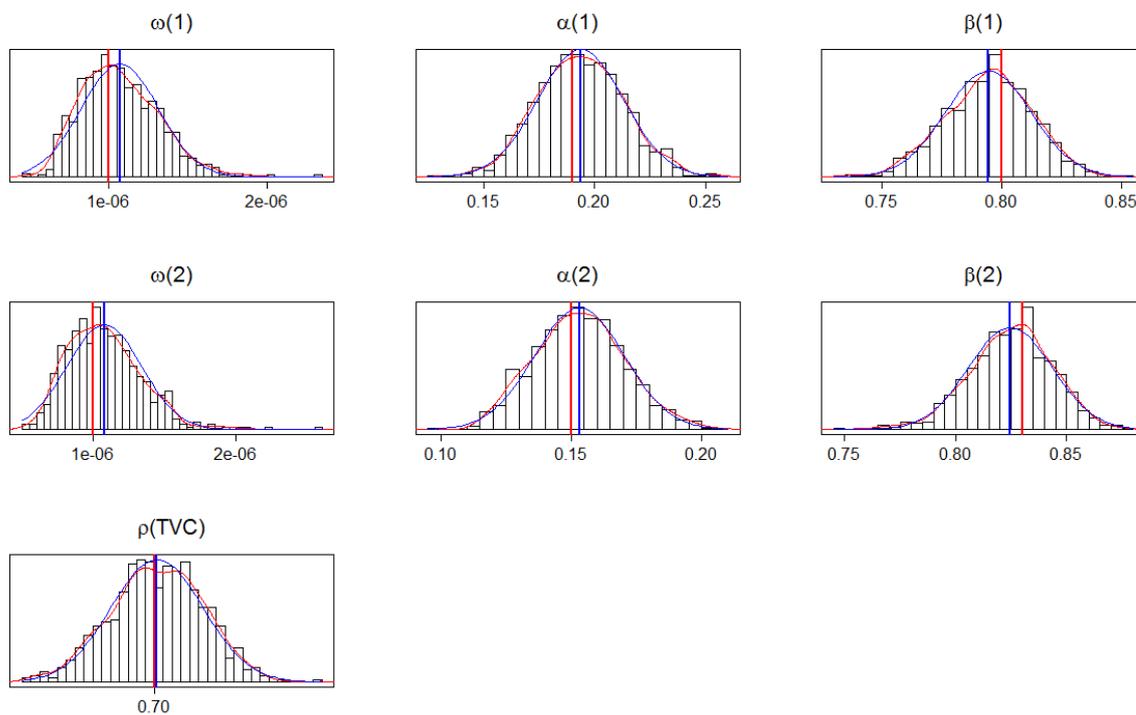


Abbildung A.1.: Dichteplots der Schätzer, Option (a) *Gauß-Copula ohne TVC*

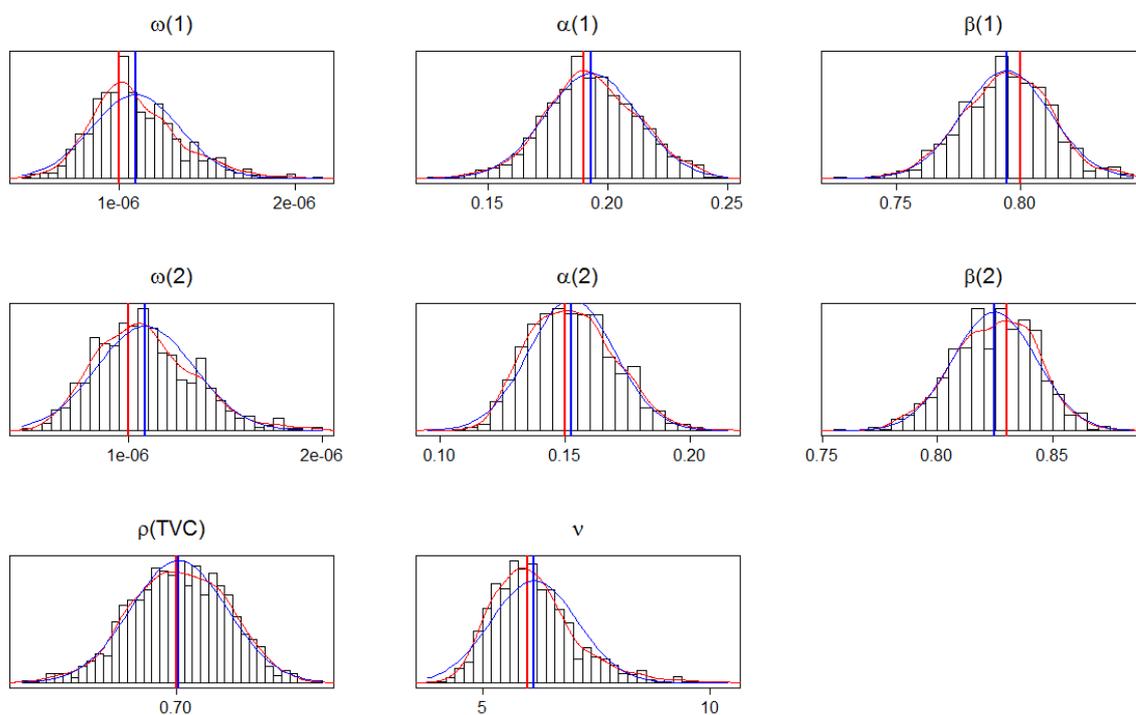


Abbildung A.2.: Dichteplots der Schätzer, Option (b) *t-Copula ohne TVC*

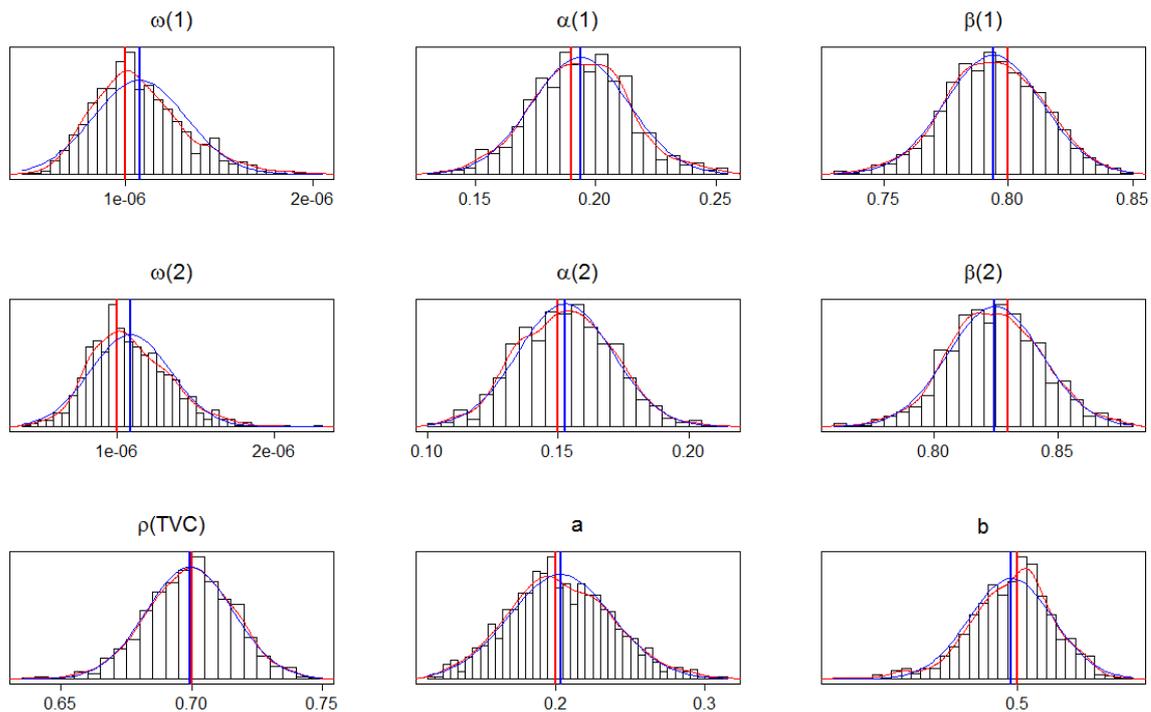


Abbildung A.3.: Dichteplots der Schätzer, Option (c) *Gauß-Copula mit TVC*

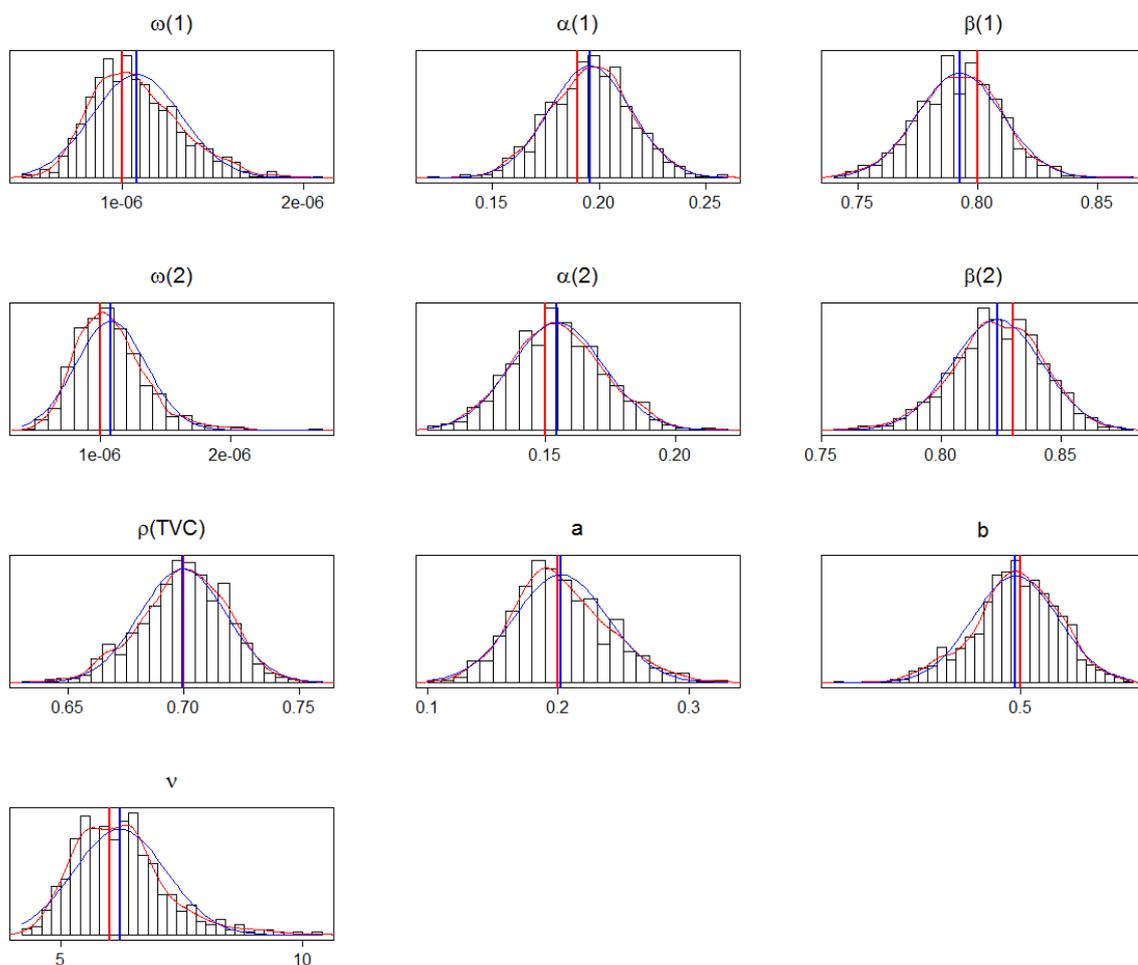


Abbildung A.4.: Dichteplots der Schätzer, Option (d) *t-Copula mit TVC*

Anhang B

Ergebnisse der Anwendung

Zugrundeliegende Daten:

- DAX-Preisindizes der Commerzbank und Deutsche Bank
- Zeitraum: 02.01.1995 – 26.01.2009 (insgesamt je zwei mal 3562 Tagesdaten)

Marginalparameter			
GARCH ($i = 1$)		GARCH ($i = 2$)	
$\omega^{(1)}$	0,000004	$\omega^{(2)}$	0,000004
$\alpha_1^{(1)}$	0,101815	$\alpha_1^{(2)}$	0,096470
$\beta_1^{(1)}$	0,896120	$\beta_1^{(2)}$	0,901548

Copula-Parameter				
	Option (a)	Option (b)	Option (c)	Option (d)
ρ_{TVC}	0,613157	0,644690	0,649894	0,765956
a	–	–	0,012828	0,014746
b	–	–	0,975388	0,977826
ν	–	10,184534	–	9,977657

Tabelle B.1.: Schätzergebnisse für die Parameter aus den Optionen (a) bis (d) bei Anwendung auf reale Daten (Commerzbank- und Deutsche-Bank-Aktien)

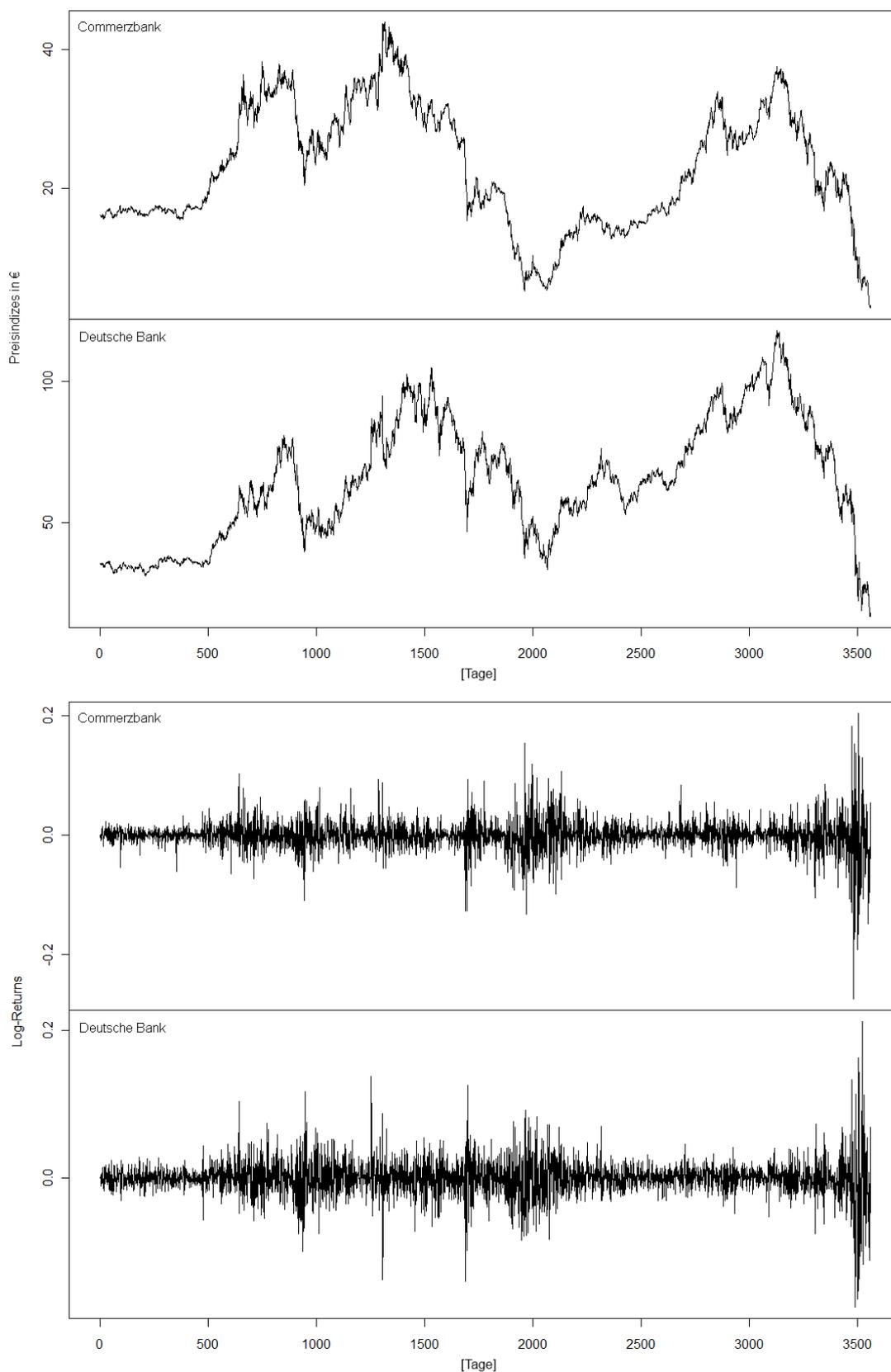


Abbildung B.1.: Preisverläufe [oben] und Log>Returns [unten] der Commerzbank- und Deutsche-Bank-Aktien im Vergleich (3562 Werktagessdaten zwischen 02.01.1995 und 26.01.2009)

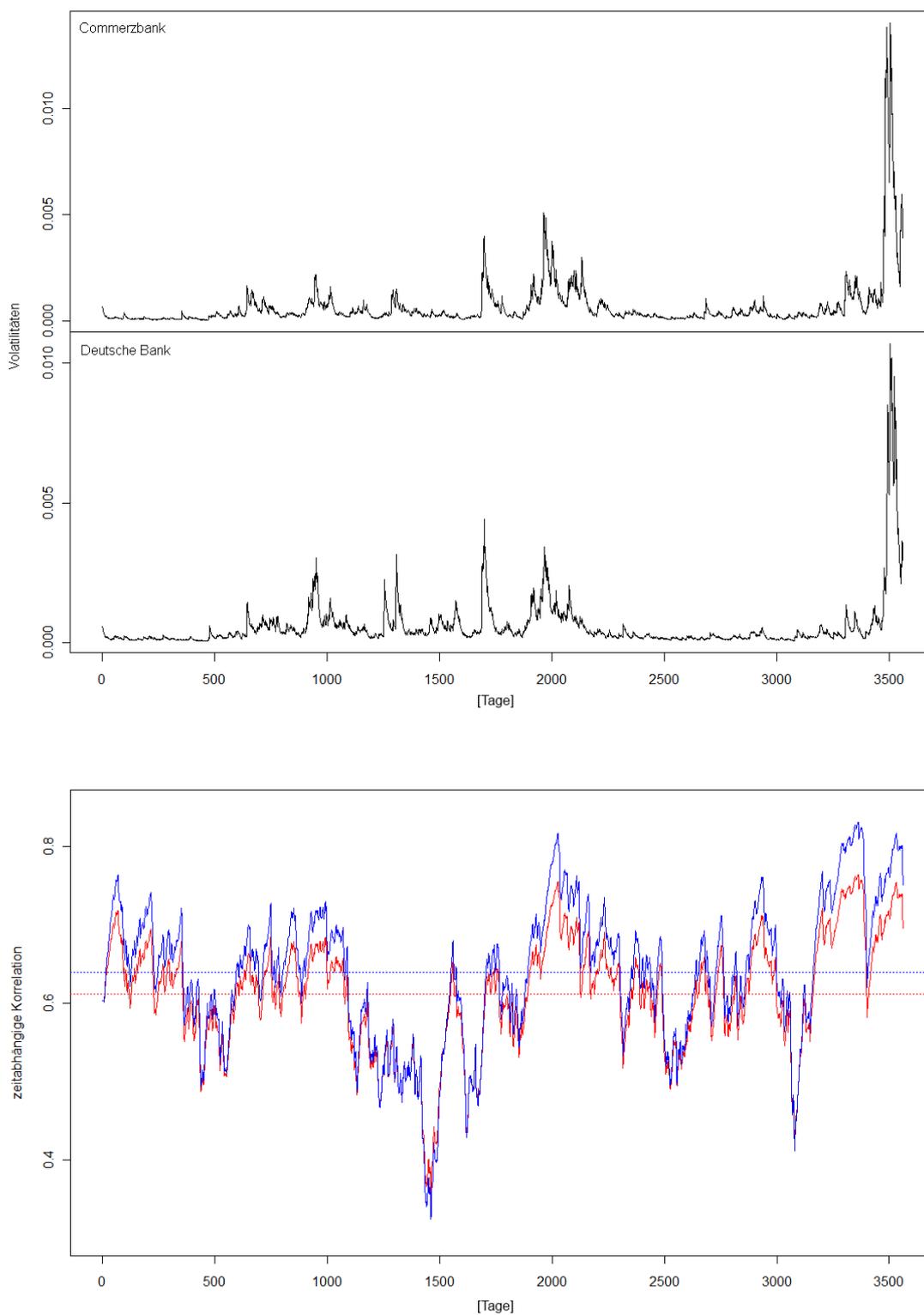


Abbildung B.2.: Volatilitätsverlauf der Commerzbank- und Deutsche-Bank-Aktien im Vergleich [oben] und zeitabhängige Korrelation mit den Optionen (c) *Gauß-Copula mit TVC* [rot] und (d) *t-Copula mit TVC* [blau] und jeweils mittlere Korrelationen (horizontale Linien) [unten] (3562 Werktagesdaten zwischen 02.01.1995 und 26.01.2009)

Literaturverzeichnis

- [1] Almeida, Czado (2010), *Efficient Bayesian inference for stochastic time-varying copula models*, Preprint submitted to Computational Statistics and Data Analysis
- [2] Bachelier (1900), *Théorie de la Spéculation*, Annales scientifiques de l'École Normale Supérieure Sér. 3
- [3] Diebold, Gunther, Tay (1998), *Evaluating Density Forecasts with Applications to Financial Risk Management*, International Economic Review
- [4] Embrechts, Lindskog, McNeil (2001), *Modelling Dependence with Copulas and Applications to Risk Management*, in Rachev (2003) *Handbook of heavy tailed distributions in finance*
- [5] Embrechts, McNeil, Straumann (2002), *Correlation and Dependence Properties in Risk Management: Properties and Pitfalls*, Cambridge University Press
- [6] Engle (1982), *Auto-regressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.*, Econometrica 50
- [7] Engle, Sheppard (2001), *Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH*, NBER Working Papers 8554, National Bureau of Economic Research, Inc.
- [8] Fama (1965), *The Behaviour of Stock Market Prices*, Journal of Business 38
- [9] Ferguson (1996), *A Course in Large Sample Theory*, Chapman and Hall
- [10] Fermanian, Wegkamp (2004), *Time dependent copulas*, Preprint INSEE, Paris
- [11] Fisher (1932), *Statistical Methods for Research Workers*
- [12] Franke, Härdle, Hafner (2004), *Einführung in die Statistik der Finanzmärkte*, Springer-Verlag
- [13] Hafner, Manner (2009), *Dynamic stochastic copula models: Estimation, inference and applications*
- [14] Hansen (1994), *Autoregressive conditional density estimation*, International Economic Review
- [15] Hassler (2007), *Stochastische Integration und Zeitreihenmodellierung: Eine Einführung mit Anwendungen aus Finanzierung und Ökonometrie (Statistik und ihre Anwendungen)*, Springer-Verlag

-
- [16] Jondeau, Rockinger (2006), *The Copula-GARCH model of conditional dependencies: An international stock market application*, Journal of International Money and Finance
- [17] Jondeau, Poon, Rockinger (2007), *Financial Modeling Under Non-Gaussian Distributions*, Springer Finance
- [18] Joe, Xu (1996), *The Estimation Method of Inference Functions for Margins for Multivariate Models*, Department of Statistics, University of British Columbia
- [19] Joe (1997), *Multivariate Models and Dependence Concepts*, Monographs in Statistics and Probability 73, Chapman and Hall, London
- [20] Klein, Köck, Tinkl (2009), *Spatial-serial dependency in multivariate GARCH models and dynamic copulas: a simulation study*, Diskussionspapier, Friedrich-Alexander-Universität Erlangen-Nürnberg
- [21] Kreiß, Neuhaus (2006), *Einführung in die Zeitreihenanalyse*, Springer-Verlag
- [22] Mandelbrot (1963), *The Variation of certain speculative Prices*, Journal of Business 35
- [23] McNeil, Frey, Embrechts (2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press
- [24] Mikosch (2006), *Copulas: Tales and Facts*, University of Copenhagen
- [25] Nelsen (2006), *An Introduction to Copulas*, Springer Series in Statistics
- [26] Neusser (2009), *Zeitreihenanalyse in den Wirtschaftswissenschaften*, Vieweg+Teubner-Verlag
- [27] Niehof (2009), *Modellierung von stochastischen Abhängigkeiten mittels Copulas*, Diplomarbeit, Westfälische Wilhelms-Universität Münster
- [28] Patton (2002), *Applications of Copula Theory in Financial Econometrics*, Unveröffentlichte Dissertation, University of California, San Diego
- [29] Patton (2006a), *Modelling Asymmetric Exchange Rate Dependence Using the Conditional Copula*, International Economic Review
- [30] Patton (2006b), *Estimation of Multivariate Models for Time Series of Possibly Different Lengths*, Journal of Applied Economics
- [31] Patton (2007), *Copula-Based Models for Financial Time Series*, in Davis, Kreiß, Mikosch *Handbook of Financial Time Series*, Springer-Verlag

-
- [32] Scharlo (2009), *Quantifizierung von Marktrisiken mittels Copula-Funktionen*, Diplomarbeit, Leibniz-Universität Hannover
- [33] Schmitz (2003), *Copulas and Stochastic Processes*, Diplomarbeit, RWTH Aachen
- [34] Silvennoinen, Teräsvirta (2008), *Multivariate GARCH-Models*, in Davis, Kreiß, Mikosch *Handbook of Financial Time Series*, Springer-Verlag
- [35] Sklar (1959), *Fonctions de repartition a n dimensions et leurs marges*, Publications de l'Institut Statistique de l'Université de Paris
- [36] Tsay (2005), *Analysis of Financial Time Series*, Wiley Series in Probability and Statistics
- [37] Tse, Tsui (2002), *A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model With Time-Varying Correlations*, Journal of Business and Economic Statistics
- [38] Weiß (2010), *Über die Verwendung von Copula-Funktionen im quantitativen Risikomanagement und in der Untersuchung von Bankenrisiken*, Inauguraldissertation, Ruhr-Universität Bochum
- [39] White (1994), *Estimation, Inference and Specification Analysis*, Econometric Society Monographs 22, Cambridge University Press, Cambridge, UK

Erklärung

Marburg, 6. September 2010

Hiermit versichere ich, dass ich diese Bachelorarbeit eigenständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Stellen meiner Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, habe ich in jedem Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht. Selbiges gilt sinngemäß für Tabellen und Abbildungen. Diese Arbeit hat in dieser oder einer ähnlichen Form noch nicht im Rahmen einer anderen Prüfung vorgelegen.

(Márton Eifert, Verfasser)