

Analyse von Copula-Modellen bei Finanzmarktdaten mit Bindungen

Masterarbeit

Vorgelegt von

Kevin Tuz

aus 

Angefertigt am
Mathematischen Institut
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

29. November 2021

Betreuer: Prof. Dr. Axel Bücher

Grußbotschaft der Bewohner des Planeten Erde – Voyager Golden Record
Kurt Josef Waldheim – Generalsekretär der Vereinten Nationen (1977):

*Wir treten aus unserem Sonnensystem in das Universum
und sind auf der Suche nach Frieden und Freundschaft.*

*Zu lehren, wenn wir dazu berufen werden,
zu lernen, wenn wir das Privileg erhalten.*

*Wissend, dass unser Planet und all seine Bewohner nur
ein kleiner Teil dieses immensen Universums um uns sind,
unternehmen wir diesen Schritt in Demut und Hoffnung.*

* * *

Tausend Dank gebührt allen lieben Menschen, die mich in diesem Lebenskapitel unterstützt haben und mir das Privileg des Lernens gaben. Ohne die Unterstützung und Motivation wäre mir diese Möglichkeit des persönlichen Wachstums nie umsetzbar gewesen. Ich danke von ganzem Herzen:

- Prof. Dr. Axel Bücher: *Ihr Vertrauen und Ihre Unterstützung in der Forschung, Ihr Verständnis und die Freiheit, die Sie mir einräumten, war für mich die bestmögliche Grundlage für die mir so wichtige Arbeit. Ihnen verdanke ich dieses Kapitel meiner akademischen Laufbahn!*



Inhaltsverzeichnis

Abbildungsverzeichnis	II
Tabellenverzeichnis	III
Übersicht mathematischer Notationen	VI
Abkürzungsverzeichnis	VIII
1 Überblick und Einführung in die Forschungsfrage	1
2 Grundlegende Theorie	4
2.1 Mathematisch-statistische Grundlagen	4
2.2 Copulas und der Satz von Sklar	8
2.3 Copulafamilien	13
2.4 Weitergehende Definitionen für Copulas	15
2.5 Betrachtung bivariater Copula-Modelle	22
3 Bindungen und rangbasierte Copula-Methoden	24
3.1 Bindungen und Ränge	25
3.2 Die empirische Copula als rangbasiert nichtparametrischer Copula-Schätzer	29
3.3 Rangbasierte Copula-Methoden	32
4 Empirische und synthetische Datensätze	45
4.1 Rahmenbedingungen auf Finanzmärkten	45
4.2 Generierung synthetischer Datensätze	50
4.3 Sichtung empirischer Finanzmarktdaten	54
4.4 Auswertung der empirischen Daten	61
5 Methoden zur Handhabe von Bindungen	68
5.1 Naive Methoden	68
5.2 Randomisierungsmethoden	74

5.3	Intervall-Zensur nach Li et al. (2020)	77
5.4	Bootstrap-Methoden für Datensätze mit Bindungen nach Kojadinovic (2017)	81
6	Simulationsstudie	87
6.1	Rahmenbedingungen der Simulationen	87
6.2	Simulation zum Methodenvergleich bzgl. Radial-Symmetrie und Austauschbarkeit	91
6.3	Simulation zum Methodenvergleich bzgl. Modellanpassung und Anpassungsgüte	105
7	Zusammenfassung und Ausblick	116
7.1	Beantwortung der Forschungsfragen	116
7.2	Anwendungsempfehlungen	117
7.3	Kritik und Ausblick	120
A	Detaillierte Testresultate für die Tests auf Radial-Symmetrie und Austauschbarkeit	126
B	Detaillierte Testresultate zur MPLE-Modellanpassung	131
C	Detaillierte Testresultate zum Test der Anpassungsgüte	144
D	Algorithmen und Implementierungen	149
D.1	Allgemeine Funktionen	149
D.2	Grundlegende Tests ohne Modifikation	152
D.3	Methoden zur Anpassung der Daten und modifizierte Tests	154
E	Mathematischer Appendix	159
E.1	Verteilungsfunktionen	159
E.2	Beweise	160
	Erklärung	162

Abbildungsverzeichnis

2.1	Höhenlinienschema und dreidimensionale Visualisierung der Fréchet-Hoeffding-Schranken und der Unabhängigkeitscopula.	12
2.2	Streudiagramme der Gauss-, t-, Clayton- und Gumbel-Copula für verschiedene Parameter ϑ sowie der Frechét-Hoeffding-Schranken und der Unabhängigkeitscopula.	16
4.1	Visuelle Aufbereitung der Präzision und Bindungsanteile in Preis- und Renditedaten im zeitlichen Verlauf.	65
6.1	\mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie (1 von 4)	93
6.2	\mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie (2 von 4)	94
6.3	\mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie (3 von 4)	95
6.4	\mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie (4 von 4)	96
6.5	\mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit (1 von 4)	99
6.6	\mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit (2 von 4)	100
6.7	\mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit (3 von 4)	101
6.8	\mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit (4 von 4)	102
6.9	Bias und RMSE der Modellanpassungen (1 von 4)	110
6.10	Bias und RMSE der Modellanpassungen (2 von 4)	111
6.11	Bias und RMSE der Modellanpassungen (3 von 4)	112
6.12	Bias und RMSE der Modellanpassungen (4 von 4)	113
7.1	Schema zur Methodenauswahl für die Tests auf Radial-Symmetrie	118
7.2	Schema zur Methodenauswahl für die Tests auf Austauschbarkeit	118
7.3	Schema zur Methodenauswahl zur Modellanpassung	119
7.4	Schema zur Methodenauswahl für die Tests der Anpassungsgüte	119

Tabellenverzeichnis

4.1	Tickgrößen-Tabelle für Aktien, Zertifikate und börsengehandelte Fonds (ETFs)	46
4.2	Übersicht der Gruppenbezeichnungen, sortiert in Reihenfolge des Auftretens.	55
4.3	Auflistung der Finanzmarkt-Datensätze 1 bis 45 von 179 über die Regionen <i>EU</i> , <i>AMER</i> und <i>APAC</i>	57
4.4	Auflistung der Finanzmarkt-Datensätze 46 bis 90 von 179 über die Regionen <i>APAC</i> und <i>MEA</i> , und Sektoren <i>EG/UT</i> und <i>MT</i>	58
4.5	Auflistung der Finanzmarkt-Datensätze 91 bis 135 von 179 über die Sektoren <i>FN</i> , <i>ID</i> , <i>CN</i> , <i>LS</i> und <i>TC</i>	59
4.6	Auflistung der Finanzmarkt-Datensätze 136 bis 179 von 179 über den Sektor <i>TC</i> und die Gruppen <i>COMM</i> , <i>FX</i> und <i>CX</i>	60
4.7	Auswertung der Datensätze bzgl. der gesamten Datengrundlage und eingeschränkt auf Stichproben mit Bindungen in Preis- oder Renditedaten	61
4.8	Quellenbasierte Auswertung der Datensätze	63
4.9	Auswahl repräsentativer Datensätze	67
6.1	Übersicht der Methoden zur Mitigation von Bindungen bei den rangbasierten Tests auf Radial-Symmetrie und auf Austauschbarkeit zur Anwendung in der Simulationsstudie.	91
6.2	Übersicht der Methoden zur Mitigation von Bindungen bei den rangbasierten Tests auf Radial-Symmetrie und Austauschbarkeit zur Anwendung in der Simulationsstudie.	105
A.1	\mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie für die \mathcal{H}_0 -Copulas.	127
A.2	\mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie für die \mathcal{H}_1 -Copulas.	128
A.3	\mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit für die \mathcal{H}_0 -Copulas.	129

A.4	\mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit für die \mathcal{H}_1 -Copulas.	130
B.1	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gauß'sche Copula mit $\tau = 0.25$	132
B.2	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gauß'sche Copula mit $\tau = 0.50$	133
B.3	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gauß'sche Copula mit $\tau = 0.75$	134
B.4	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die t-Copula mit $\tau = 0.25$	135
B.5	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die t-Copula mit $\tau = 0.50$	136
B.6	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die t-Copula mit $\tau = 0.75$	137
B.7	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Clayton-Copula mit $\tau = 0.25$	138
B.8	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Clayton-Copula mit $\tau = 0.50$	139
B.9	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Clayton-Copula mit $\tau = 0.75$	140
B.10	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gumbel-Copula mit $\tau = 0.25$	141
B.11	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gumbel-Copula mit $\tau = 0.50$	142
B.12	Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gumbel-Copula mit $\tau = 0.75$	143
C.1	\mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der Gauß'schen Copula	145
C.2	\mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der t-Copula	146
C.3	\mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der Clayton-Copula	147
C.4	\mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der Gumbel-Copula	148

Übersicht mathematischer Notationen

Aufbauend auf den Relationen “ $<$ ” und “ \leq ” für Skalare, wobei hier \leq definiert ist als $a \leq b : \Leftrightarrow a < b \vee a = b$, seien für Vektoren $x, y \in \mathbb{R}^d$ mit Komponenten $x = (x_1, \dots, x_d)$ und $y = (y_1, \dots, y_d)$ die drei Relationen “ $<$ ”, “ \leq ” und “ \leqq ” definiert als:

$$\begin{aligned} x < y &\quad \forall k \in \{1, \dots, d\}: x_k < y_k, \\ x \leq y &\quad \forall k \in \{1, \dots, d\}: x_k \leq y_k, \\ x \leqq y &\quad \forall k \in \{1, \dots, d\}: x_k \leqq y_k \text{ und } \exists k \in \{1, \dots, d\}: x_k < y_k. \end{aligned}$$

Unter dieser Definition ist der Ausdruck $a \leq b$ für Skalare gleichbedeutend zu $a < b$; zum einfacheren Verständnis wird der Ausdruck $a \leq b$ für Skalare vermieden, falls möglich. Die Erweiterung und gewisse Teilmengen der reellen Zahlen werden mit

$$\begin{aligned} \overline{\mathbb{R}} &\quad \text{Erweiterte reelle Zahlen, d. h. } \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}, \\ \mathbb{R}_>^d, \mathbb{R}_\geq^d, \mathbb{R}_\leq^d &\quad \text{Positive } \mathbb{R}^d\text{-Teilräume, definiert als } \mathbb{R}_>^d := \{x \in \mathbb{R}^d : x > 0\}, \\ &\quad \mathbb{R}_\geq^d := \{x \in \mathbb{R}^d : x \geq 0\} \text{ und } \mathbb{R}_\leq^d := \{x \in \mathbb{R}^d : x \leq 0\} \end{aligned}$$

notiert.

Für (komponentenweise) Minima und Maxima von Vektoren $x, y \in \mathbb{R}^d$ werden neben den bereits bekannten Schreibweisen $\min\{x, y\}$, $\max\{x, y\}$ auch

$$\begin{aligned} x \wedge y &\quad \text{Minimum von } x \text{ und } y, \\ x \vee y &\quad \text{Maximum von } x \text{ und } y \end{aligned}$$

als abkürzende Notationen verwendet; diese sind an die mengentheoretischen Notationen \cup und \cap angelehnt.

Für Funktionen $f : X \rightarrow Y$ bedeuten

$$\begin{aligned} \text{dom}(f) &\quad \text{Definitionsbereich der Funktion } f, \text{ dom}(f) = X, \\ \text{ran}(f) &\quad \text{Bildbereich der Funktion } f, \text{ ran}(f) \subseteq Y. \end{aligned}$$

Für maßtheoretische und mathematisch-statistische Aussagen werden die folgenden Notationen für Räume und Funktionen genutzt:

$(\Omega, \mathcal{A}, \mu)$	Maßraum mit Grundmenge Ω , zugehöriger σ -Algebra \mathcal{A} und darauf definiertem Maß $\mu : \mathcal{A} \rightarrow \mathbb{R}_{\geq}$,
$(\Omega, \mathcal{A}, \mathbb{P})$	Wahrscheinlichkeitsraum, d. h. Maßraum $(\Omega, \mathcal{A}, \mu = \mathbb{P})$ mit dem Wahrscheinlichkeitsmaß $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$,
$L^p(\Omega)$	L^p -Räume: Lebesgue'sche Räume der p -fach integrierbaren Funktionen $f : \Omega \rightarrow \mathbb{R}$, $(1 \leq p \leq \infty)$,
$\mathbb{1}_A(x)$	Indikatorkfunktion einer Menge oder Aussage A ,
$\lambda(A)$	Lebesguemaß einer Menge A ,
$ A $	Kardinalität einer abzählbaren Menge A .

Für Verteilungs- und Konvergenzaussagen von Zufallsvariablen und -vektoren werden folgende Notationen genutzt:

$X \sim F$	die Zufallsvariable X folgt der Verteilung F ,
$X \perp\!\!\!\perp Y$	die Zufallsvariablen X und Y sind stochastisch unabhängig,
$(X_k)_k \stackrel{i.i.d.}{\sim} F$	die Zufallsvariablen $(X_k)_k$ sind unabhängig identisch verteilt (i.i.d.) bzgl. einer Verteilung F , d. h. es sind $X_k \sim F$ für alle k , und $X_k \perp\!\!\!\perp X_l$ für alle $k \neq l$,
$X_k \xrightarrow{d} X$	X_k konvergiert in Verteilung X bzw. das induzierte Maß \mathbb{P}_{X_k} konvergiert schwach gegen \mathbb{P}_X für $k \rightarrow \infty$,
$X_k \xrightarrow{\mathbb{P}} X$	X_k konvergiert stochastisch gegen X für $k \rightarrow \infty$.

Abkürzungsverzeichnis

ADNT	Average Daily Number of Transactions – Durchschnittliche Transaktionsanzahl pro Tag
AMER	Americas – Wirtschaftsraum Nord-, Mittel- und Südamerika
APAC	Asia-Pacific – Wirtschaftsraum Asien und (West-)Pazifik
Bias	Verzerrung, Abweichung
Bin _{n,p}	Binomialverteilung über $n \in \mathbb{N}$ Versuche mit Erfolgswahrscheinlichkeit $p \in [0, 1]$
càdlàg	Continue à gauche et limite à droite – Rechtsseitig stetig und linksseitig existierende Grenzwerte
CDF	Cumulative Probability Density Function – Verteilungsfunktion. Siehe auch: PDF
CN	Consumer Goods – Konsumgüter- und Handelsunternehmen
COMM	Commodity market – Rohstoffhandel
CX	Cryptocurrency – Kryptowährung
EG	Energy – Energiewirtschaft
ETF	Exchange Traded Fund – Börsengehandelter Fond
EU	Europäische Union oder Wirtschaftsraum Europa inklusive Großbritannien und Russland
FN	Financials – Finanzdienstleistungen
FX	Foreign Exchange – Wechselkurse und Währungen
ID	Industrials – Industrieunternehmen
i.i.d.	Independent and identically distributed – unabhängig und identisch verteilt

ITS	Inverse Transform Sampling – Inversionsmethode bzw. Quantiltransformation. Siehe auch: PIT
IQR	Interquartilsabstand – Intervall zwischen dem 25-Prozent- und dem 75-Prozent-Quantil einer Stichprobe
LS	Life Science & Healthcare – Gesundheitswesen
MEA	Middle East & Africas – Wirtschaftsräume Naher bzw. Mittlerer Osten und Afrika
MME	Momenten-Methode-Schätzer
MPLE	Maximum-Pseudo-Likelihood-Schätzer über Pseudostichproben.
MSE	Mean Squared Error – Mittlere Quadratische Abweichung
MT	Materials, Materialien / Grundstoffe 55
OTC	Over The Counter, Außerbörslicher Handel
PDF	Probability Density Function (Dichtefunktion, Wahrscheinlichkeitsdichte). Siehe auch: CDF
PIT	Probability Integral Transformation (Transformation in Wahrscheinlichkeit). Siehe auch: ITS
Tick	Notierungssprung, Einheit für Kursänderungen
TC	Technology, Technologieunternehmen
Unif $_{\Omega}$	Gleichverteilung (Uniform Distribution) auf Ω , z. B. ist $\text{Unif}_{[0,1]}$ die Gleichverteilung auf dem Null-Eins-Intervall
UMPU	Uniformly Most Powerful Unbiased (Test) – gleichmäßig bester unverfälschter Test. Siehe auch: UMVU
UMVU	Uniformly Minimum Variance Unbiased (Estimator) – gleichmäßig bester unverzerrter Schätzer. Siehe auch: UMPU
UT	Utilities, Versorgungsunternehmen

Kapitel 1

Überblick und Einführung in die Forschungsfrage

Mathematisch-statistische Modellierungen von Abhängigkeitsstrukturen sind in der wirtschaftlichen Anwendung oft von Notwendigkeit oder Relevanz: Im quantitativen Risikomanagement werden beispielsweise Risikobewertungen auf Basis statistischer Simulationen durchgeführt. Risikomodelle können dabei aus einer Vielzahl von Risikofaktoren und externen Einflüssen basieren, die nicht unabhängig nebeneinander stehen; entsprechend sind multivariate Modelle notwendig (vgl. McNeil et al. (2015): S. 173).

Eine 1:1-Abbildung der einzelnen Prozesse hinter diesen Faktoren ist in der Praxis oft nicht durchführbar, weswegen die Abhängigkeiten über Korrelationen oder andere Konzepte zusammengefasst werden müssen. Die Anwendung von linearen Korrelationen ist für Verteilungen abseits von der Normalverteilung und elliptischer Verteilungen nur bedingt sinnvoll. Gleichzeitig wird eine flexible Wahl und Kombination der Verteilungen im Sachzusammenhang gefordert, so dass ein allgemeineres Konzept für die multivariate Modellierung gewünscht ist (vgl. McNeil et al. (2015): S. 220). Diese Verallgemeinerung wird mit den sogenannten Copulas geliefert.

Die Nutzung von Copulas in der Praxis und Forschung ist gut dokumentiert. So wird in McNeil et al. (2015), genauer in den Kapiteln 7 und 15.2, und in Patton (2013) die Anwendung von Copulas im (Finanz-)Risikomanagement und in der Ökonometrie eingeordnet; Hofert et al. (2018) motiviert die Implementierung von Copula-Modellen in der Programmiersprache *R* über die Nutzung in der angewandten Statistik für Finanzmärkte und Ingenieurwissenschaften. Hlawatsch und Reichling (2010) behandeln die finanzwirtschaftliche Nutzung von Copulas im Portfoliomangement, Risikomanagement und der Optionspreisbewertung. Die Veröffentlichungen Pappadà et al. (2016) und Lü et al. (2020) nutzen Copulas zum Beispiel für umweltwissenschaftli-

che Modelle in der Hydrologie und der Analyse von Boden-Messwerten.

Für statistische Methoden und Tests liegen gewisse Anforderungen und Annahmen zugrunde; diese sind im Praxisumfeld bzw. in der Anwendung auf empirischen Daten nicht notwendigerweise erfüllt. Für Copula-Methoden sind insbesondere die Stetigkeit der Marginale und die Bindungsfreiheit angenommen, wobei die letztere Annahme leicht zu verletzen ist (vgl. Kojadinovic (2017): S. 2). In Genest et al. (2011) wird zum Beispiel der Effekt von Bindungen für rangbasierte Copula-Methoden im Versicherungswesen behandelt und weitergehend auch die Auswirkungen von Anpassungen der Methoden diskutiert.

Entsprechend ist für Copula-Methoden auf Finanzmarktdaten zu prüfen, ob dort die jeweiligen Annahmen verletzt werden und falls ja, ist die Beantwortung folgender Forschungsfragen von Interesse:

- ① Welche Eigenschaften, Strukturen, Rahmenbedingungen, Datenfehler etc. in empirischen Finanzmarktdaten führen zu Verletzungen der Annahmen? Welcher Anteil der Daten ist davon betroffen?
- ② Welche Auswirkungen (bzgl. Bias, Anwendbarkeit, Robustheit oder ähnlichen Kriterien) haben Annahmeverstöße auf die Methoden?

Auch die Betrachtung von möglichen Anpassungen der Methoden ist sinnvoll. So wurden zum Beispiel in Kojadinovic (2017), Pappadà et al. (2016) und Li et al. (2020) verschiedene Copula-Methoden für die Anwendung auf Daten mit Bindungen modifiziert. Es ist hervorzuheben, dass Modifikationen möglicherweise die Ergebnisse verfälschen können.

Wie in der Medizinethik ist hier also bei der Anwendung von Modifikationen der Grundsatz *Primum non nocere – Erstens: Tun Sie keinen Schaden.* – anzuwenden. Entsprechend sind weitere Forschungsfragen zu formulieren:

- ③ Wie groß ist der Nutzen einer Modifikation im Vergleich zu den anderen Methoden für Datensätze mit Bindungen?
- ④ Falls mehrere Methoden für einen Sachverhalt vorliegen: Welche der Methoden ist zur Anwendung zu empfehlen? Welche Methode ist nur unter Einschränkungen oder nur mit der Bereitstellung von zusätzlichen Informationen nutzbar?

Mit Hinsicht auf die Beantwortung der Fragen ① bis ④ wird folgende Strukturierung für diese Arbeit gewählt:

Kapitel 2: Die Einführung in diese Arbeit findet im Kapitel 2 statt; dort werden die wahrscheinlichkeitstheoretischen Grundlagen sowie darauf aufbauend

die wichtigsten Definitionen und Aussagen zu Copulas vorgestellt. Es werden dann die für diese Arbeit wichtigen Copulafamilien vorgestellt.

Kapitel 3: Das dritte Kapitel behandelt Copula-Methoden und statistische Tests, die insbesondere im Umfeld der Finanzmarktmodellierung von Interesse sind. Hierbei werden auch die zugrundeliegenden Annahmen besprochen sowie die Begriffe Bindung und Rang eingeführt.

Kapitel 4: In diesem Kapitel liegt der Schwerpunkt auf der Beantwortung der Frage ①: Im Abschnitt 4.1 werden die Rahmenbedingungen auf Finanzmärkten qualitativ gesichtet und mit dem möglichen Auftreten in Beziehung gesetzt. Diese Rahmenbedingungen werden im Abschnitt 4.2 genutzt, um einem Algorithmus für die Generierung synthetischer Datensätze zu entwerfen. Eine Sichtung empirischer Finanzmarktdaten wird im Abschnitt 4.1 durchgeführt; diese werden dann mit Hinblick auf die Parameterwahl des datengenerierenden Algorithmus im Abschnitt 4.4 analysiert.

Kapitel 5: Das fünfte Kapitel nimmt erneut Bezug auf die Copula-Methoden aus Kapitel 3, nun unter Berücksichtigung von Bindungen in den Datensätzen. Es werden bereits existierende Modifikationen aus der Praxis und Forschung inklusive der Motivation bzw. des zugrundeliegenden Konzepts vorgestellt.

Kapitel 6: Die Simulationsstudie in Kapitel 6 dient der Beantwortung der Fragen ② und ③. Zur Modellierung der Studie werden die Erkenntnisse aus Kapitel 4 und die Methoden aus Kapitel 5 genutzt. Es wird auf Basis der vorigen Ergebnisse ein Methodenvergleich durchgeführt. Zielführend ist hier die Beantwortung von ④, woraus – falls möglich – konkrete Anwendungsempfehlungen hergeleitet werden sollen.

Kapitel 7: Im siebten Kapitel werden die Ergebnisse zusammengefasst und Entscheidungshilfen für die Anwendung der Methoden bereitgestellt. Auch werden ein Résumé über diese Arbeit, offene Fragen und mögliche weitere Forschungsansätze angegeben.

Kapitel 2

Grundlegende Theorie

Für die Motivation und Nutzung der Copula-Theorie sind die notwendigen mathematisch-statistischen Grundlagen zu behandeln. Diese umfassen uni- und multivariate Verteilungen, Transformationen zwischen Verteilungen sowie ausgewählte Konvergenzdefinitionen. Darauf aufbauend werden im Abschnitt 2.2 Copulas und zentrale Eigenschaften definiert sowie im Abschnitt 2.3 vier Copulafamilien vorgestellt. Im Abschnitt 2.4 werden weitergehende Definitionen und Konstruktionen sowie Ansätze zur Verallgemeinerung von bivariaten Modellen angegeben.

2.1 Mathematisch-statistische Grundlagen

Zu Beginn werden einige für diese Arbeit relevante Definitionen, Sätze und Methoden vorgestellt. Hierbei sei angenommen, dass im Folgenden, sofern nicht anders definiert, $d \in \mathbb{N}$ die Dimensionalität von Verteilungsfunktionen, Zufallsvektoren, etc., und $n \in \mathbb{N}$ die Größe von Stichproben notiere.

Zielführend für die Einführung des Copula-Begriffs ist die Behandlung von multivariaten Verteilungsfunktionen, da Copula-Funktionen einen Spezialfall dieser darstellen. Diese Grundlagen werden in verkürzter Form bereitgestellt, da nicht alle Facetten für den Inhalt dieser Arbeit notwendig sind. Weitergehende Informationen sind beispielsweise in Klenke (2008), Rüschenhoff (2016) oder anderer Literatur zur Wahrscheinlichkeitstheorie zu finden.

2.1.1 Definition (Wahrscheinlichkeitsdichte, Wahrscheinlichkeitsmaß und Verteilungsfunktion). *Sei $\Omega \subseteq \mathbb{R}^d$ Borel'sch und $(\Omega, \mathcal{B}_\Omega^d, \lambda_\Omega^d)$ der zugehörige Maßraum mit dem Lebesgue-Maß $\lambda_\Omega^d := \lambda^d|_\Omega$ auf Ω und der Borel'schen σ -Algebra $\mathcal{B}_\Omega^d := \mathcal{B}(\Omega)$. Eine Funktion $f : \Omega \rightarrow \mathbb{R}$, für die die vier Eigenschaften*

(i.) *Messbarkeit: Die Funktion $f : (\Omega, \mathcal{B}_\Omega^d) \rightarrow (\mathbb{R}, \mathcal{B})$ ist $(\mathcal{B}_\Omega^d, \mathcal{B})$ -messbar.*

- (ii.) Nicht-Negativität: Für das Bild $\text{ran}(f)$ von f gilt $\text{ran}(f) \subseteq \mathbb{R}_0^+$.
- (iii.) Integrierbarkeit: Es ist $f \in L^1(\Omega, \mathcal{B}_\Omega^d, \lambda_\Omega^d)$, d. h. es gilt $\int_{\Omega} f d\lambda_\Omega^d < \infty$.
- (iv.) Eins-Normiertheit: Die Integrierbarkeit ist erfüllt mit $\int_{\Omega} f d\lambda_\Omega^d = 1$.

erfüllt sind, konstruiert ein fast-sicher eindeutiges Wahrscheinlichkeitsmaß $\mathbb{P} : \mathcal{B}_\Omega^d \rightarrow [0, 1]$ auf dem Messraum $(\Omega, \mathcal{B}_\Omega^d)$ über

$$\mathbb{P}(E) := \int_E f d\lambda_\Omega^d \quad (E \in \mathcal{B}_\Omega^d)$$

und wird Wahrscheinlichkeitsdichte bzw. Dichtefunktion von \mathbb{P} bezeichnet.
Eine Funktion $F : \mathbb{R}^d \rightarrow [0, 1]$, welche die drei Eigenschaften

- (v.) Càdlàg: Die Funktion $F : \mathbb{R}^d \rightarrow [0, 1]$ ist càdlàg, d. h. sie ist rechtsseitig stetig und es existieren die linksseitigen Limiten.
- (vi.) Monotes Wachstum: Die Funktion F ist monoton wachsend, d. h. für $x, y \in \mathbb{R}^d$ mit $x \leq y$ gilt $F(x) \leq F(y)$.
- (vii.) Null-Eins-Grenzwerte: Es gelten die beiden Konvergenzaussagen

$$\lim_{\max_i x_i \rightarrow -\infty} F(x) = 0 \quad \text{und} \quad \lim_{\min_i x_i \rightarrow \infty} F(x) = 1,$$

wobei $x = (x_1, \dots, x_d)$ ein Vektor sei.

erfüllt, wird Verteilungsfunktion genannt.

Die zur Dichtefunktion f bzw. zum Wahrscheinlichkeitsmaß \mathbb{P} gehörende Verteilungsfunktion $F : \mathbb{R}^d \rightarrow [0, 1]$ ist definiert als

$$F(x) := \mathbb{P}((-\infty, x] \cap \Omega) \quad (x \in \Omega),$$

wobei $(-\infty, x] := \{y \in \mathbb{R}^d : y \leq x\} = (-\infty, x_1] \times \dots \times (-\infty, x_d]$ für mehrdimensionale $x = (x_1, \dots, x_d)$ sei.

2.1.2 Definition (Generalisierte Inverse / Quantilfunktion). Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion. Die (linksstetige) generalisierte Inverse bzw. die Quantilfunktion $F^\leftarrow : [0, 1] \rightarrow \overline{\mathbb{R}}$ zu F ist definiert als

$$F^\leftarrow(\alpha) := \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}, \quad (\alpha \in [0, 1]) \quad (2-1)$$

wobei die Konvention $\inf \emptyset := \infty$ gültig sei. $F^\leftarrow(\alpha)$ heißt α -Quantil von F .

2.1.3 Bemerkung. In der Literatur werden auch alternative Definitionen der generalisierten Inversen formuliert und diskutiert:

- (i.) Für die in Definition 2.1.2 formulierten Inversen gilt stets $F^\leftarrow(0) = -\infty$; für gewisse Verteilungsfunktionen F existieren Punkte $x \in \mathbb{R}$, die $F(x) = 0$ erfüllen. Um diesen Fall abzufangen, wird in mancher Literatur die generalisierte Inverse daher alternativ in der Form

$$F^\leftarrow(\alpha) := \begin{cases} \sup\{x \in \mathbb{R} : F(x) = 0\}, & \alpha = 0, \\ \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}, & \alpha \in (0, 1], \end{cases} \quad (2-2)$$

definiert (vgl. Witting (1985): S. 20, Definition 1.16). Die Existenz von x mit $F(x) = 0$ wird über den Fall $\sup\{x : F(x) = 0\}$ abgefangen. Falls kein solches x existiert, gelte die Konvention “ $\sup \emptyset := -\infty$ ”.

- (ii.) Embrechts und Hofert (2013) sprechen sich gegen die Nutzung einer solchen alternativen Definition (2-2) aus, da das Null-Quantil $F^\leftarrow(0)$ in der Praxis kaum bzw. keine Relevanz innehaltet und gleichzeitig die Komplexität der Handhabung und Beweisführung höher ist.

Wegen Bemerkung 2.1.3 (ii.) wird in dieser Arbeit die einfache Formel (2-1) für die generalisierten Inversen genutzt. Diese Definition stellt keine Einschränkung für empirische Quantilfunktionen dar; die empirische Quantilfunktion \hat{F}_n kann je nach Definition auf die Intervalle $[0, 1]$ oder $(0, 1)$ abbilden. Diese Funktion wird im Kapitel 3 tiefergehend behandelt.

Im folgenden Lemma werden einige wichtige Eigenschaften der Quantilfunktion über deren Wohldefiniertheit und Abschätzbarkeit präsentiert.

2.1.4 Lemma (Eigenschaften der Quantilfunktion). Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion mit Quantilfunktion $F^\leftarrow : [0, 1] \rightarrow \overline{\mathbb{R}}$. Seien $\alpha \in [0, 1]$ und $x \in \overline{\mathbb{R}}$ mit $x \geq F^\leftarrow(0)$. Es gelten dann die folgenden Eigenschaften:

(i.) Isotonie der Quantilfunktion: Die Quantilfunktion F^\leftarrow wächst monoton.

(ii.) Direkte Beziehung von F und F^\leftarrow : Es gilt die Äquivalenz

$$\alpha \leqq F(x) \iff F^\leftarrow(\alpha) \leqq x.$$

(iii.) Abschätzung der Position: Für alle $x \in \mathbb{R}$ mit $F(x) \in (0, 1)$ überschätzt das $F(x)$ -Quantil die ursprüngliche Position x nicht. Es gilt also

$$F^\leftarrow(F(x)) \leqq x \quad (x \in \mathbb{R} : F(x) \in (0, 1)),$$

wobei hier die Gleichheit auf dem Inneren oder dem oberen Rand eines Plateaus ungültig ist; genauer gilt

$$x \neq \inf\{y \in \mathbb{R} : F(y) = F(x)\} \iff F^\leftarrow(F(x)) < x.$$

- (iv.) Abschätzung der Wahrscheinlichkeit: *Das Wahrscheinlichkeitsniveau $\alpha \in (0, 1)$ wird durch das α -Quantil $F^\leftarrow(\alpha)$ nicht unterschätzt; es gilt*

$$F(F^\leftarrow(\alpha)) \geq \alpha \quad (\alpha \in (0, 1)).$$

Diese Ungleichung gilt nur in Unstetigkeitsstellen strikt und es gilt folgende Äquivalenzaussage:

$$\alpha \in \text{ran}(F) \iff F(F^\leftarrow(\alpha)) = \alpha.$$

Der Beweis ist im Appendix E.2 ausgeführt. Die Eigenschaften aus vorherigem Lemma sind auch auf Zufallsvariablen anwendbar. Insbesondere die Äquivalenzaussage im Lemma 2.1.4 (ii.) führt zu Transformationen von Zufallsvariablen, die über die Eigenschaften in 2.1.4 (iii.) und (iv.) zu zwei grundlegenden Transformationsmethoden fortgeführt werden. Darauf aufbauend wird im Folgeabschnitt die Copula-Funktion motiviert.

2.1.5 Lemma (Transformation in Wahrscheinlichkeit). *Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine beliebige stetige Verteilungsfunktion und $X \sim F$ eine über diese Verteilung generierte Zufallsvariable.*

Dann ist für X die Transformation in Wahrscheinlichkeit, auch Probability Integral Transformation (PIT) genannt, definiert als Zufallsvariable $F(X)$. Hierbei gilt die Transformationseigenschaft

$$X \sim F \implies F(X) \sim \text{Unif}_{[0,1]}. \quad (2-3)$$

Beweis. Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine stetige Verteilungsfunktion und zugehörig $X \sim F$ eine Zufallsvariable. Dann gilt für die Verteilungsfunktion F_U zu $U := F(X)$ die folgende Gleichung:

$$\begin{aligned} F_U(u) &:= \mathbb{P}(U \leq u) := \mathbb{P}(F(X) \leq u) \\ &\stackrel{\textcircled{1}}{=} \mathbb{P}(X \leq F^\leftarrow(u)) =: F(F^\leftarrow(u)) \stackrel{\textcircled{2}}{=} u, \end{aligned}$$

wobei in ① Lemma 2.1.4 (ii.) und in ② 2.1.4 (iv.) unter der Stetigkeit von F genutzt wurde. Somit ist $F_U(u) = u$, also $U := F(X) \sim \text{Unif}_{[0,1]}$. \square

Zusammenfassend sind also Zufallsvariablen mit beliebiger stetiger Verteilung in $\text{Unif}_{[0,1]}$ -verteilte Zufallsvariablen transformierbar. Die nun folgende Methode erlaubt die Gegenrichtung:

2.1.6 Lemma (Inversionsmethode / Quantiltransformation). *Seien eine beliebige Verteilungsfunktion $F : \mathbb{R} \rightarrow [0, 1]$ und die zugehörige Quantilfunktion $F^\leftarrow : [0, 1] \rightarrow \overline{\mathbb{R}}$ gegeben. Weiter betrachte die Zufallsvariable $U \sim \text{Unif}_{[0,1]}$.*

Die Inversionsmethode bzw. Quantiltransformation, auch Inverse Transform Sampling (ITS) genannt, bildet U auf die Zufallsvariable $F^\leftarrow(U)$ ab. Es gilt dabei die folgende Transformationseigenschaft

$$U \sim \text{Unif}_{[0,1]} \implies F^\leftarrow(U) \sim F. \quad (2-4)$$

Beweis. Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion und F^\leftarrow die zugehörige Quantilfunktion. Seien weiter $U \sim \text{Unif}_{[0,1]}$ und $X := F^\leftarrow(U)$ Zufallsvariablen. Es gilt dann für die Verteilungsfunktion F_X zu X die Gleichung

$$\begin{aligned} F_X(x) &:= \mathbb{P}(X \leq x) := \mathbb{P}(F^\leftarrow(U) \leq x) \\ &\stackrel{\textcircled{1}}{=} \mathbb{P}(U \leq F(x)) =: F_U(F(x)) \stackrel{\textcircled{2}}{=} F(x), \end{aligned}$$

wobei in ① Lemma 2.1.4 (ii.) und in ② die Definition der $\text{Unif}_{[0,1]}$ -Verteilung $F_U(u) = u$ für $u \in [0, 1]$ angewandt wurde. Somit folgt $X := F^\leftarrow(U) \sim F$. \square

Diese Methoden erlauben – im univariaten Fall – die Transformation zwischen genügend stetigen Verteilungen F und G und die durch diese Verteilungen generierte Zufallsvariablen. Für eine Funktion $H : [0, 1]^1 \rightarrow [0, 1]$, $H(u) := G(F^\leftarrow(u))$ gilt somit, dass G die Darstellung $G(x) = H(F(x))$ bzgl. $F(x)$ besitzt. H enthält also alle Informationen zur *Übersetzung* der Verteilung F in die Verteilung G . Da multivariate Verteilungen sowohl die zugehörigen Randverteilungen als auch die Abhängigkeitsstrukturen dazwischen modellieren, stellt sich die Frage, ob diese zwei Teile getrennt modelliert werden können. Die Marginale F_X, F_Y einer Verteilung $F_{X,Y}$ sind in der Regel bereits bekannt oder herleitbar. Eine mehrdimensionale Verallgemeinerung der Funktion H , zum Beispiel ein $H : [0, 1]^2 \rightarrow [0, 1]$ derart, dass $F_{X,Y}(x, y) = H(F_X(x), F_Y(y))$ darstellbar ist, würde eine direkte Modellierung oder zumindest eine Formulierung der Abhängigkeitsstruktur zwischen den Marginalen ermöglichen. Falls dies möglich ist: Ist diese Funktion über PIT / ITS aus den Verteilungen $F_{X,Y}, F_X, F_Y$ konstruierbar?

Diese Fragen über einer Verallgemeinerung für den multivariaten Fall und die Nutzung zur Modellierung von Abhängigkeiten sind die naheliegenden nächsten Schritte. Hiermit wird die Definition von Copulas motiviert.

2.2 Copulas und der Satz von Sklar

Zu Beginn dieses Abschnitts werden zwei äquivalente Definitionen für Copula-Funktionen formuliert. Der darauffolgende *Satz von Sklar* stellt zentrale Aussagen zur Existenz von Copulas und zur Modellierung von Abhängigkeitsstrukturen zwischen univariaten Verteilungen dar. Hierbei wird auch die

Parallele einer impliziten Copula-Konstruktion zur PIT und ITS erkenntlich. Abschließend werden zentrale Eigenschaften und die in dieser Arbeit verwendeten Copulas dargestellt.

2.2.1 Definition (Copula). *Eine Funktion $C : [0, 1]^d \rightarrow [0, 1]$ wird Copula genannt, falls die folgenden Eigenschaften erfüllt sind:*

- (i.) Verteilungsfunktion: *C ist eine multivariate Verteilungsfunktion bzw. die Einschränkung einer solchen auf dem Einheits-Hyperwürfel $[0, 1]^d$.*
- (ii.) Unif_[0,1]-Marginale: *Die Marginale C_1, \dots, C_d der Copula C beschreiben Gleichverteilungen auf dem Intervall $[0, 1]$. Es ist also*

$$C_j(u_j) := C(1, \dots, 1, u_j, 1, \dots, 1) = u_j \quad (u_j \in [0, 1])$$

gültig für jedes $j \in \{1, \dots, d\}$.

Neben dieser intuitiven, wahrscheinlichkeitstheoretischen Definition über die multivariate Verteilung und Marginale ist ebenso eine Charakterisierung über analytische Eigenschaften formulierbar. Die folgenden Definitionen für Copulas und den zugehörig notwendigen Eigenschaften basieren auf den in Kapitel 2.1 und 2.2 sowie 2.10 in Nelsen (2006) vorliegenden Formulierungen.

2.2.2 Definition (d -steigende Funktionen). *Seien $F : [0, 1]^d \rightarrow [0, 1]$ eine Funktion und $R := [x_1, y_1] \times \dots \times [x_d, y_d] \subseteq [0, 1]^d$ ein Hyperrechteck.*

Das F -Volumen von R ist definiert als

$$\text{Vol}|_F(R) := \sum_{\substack{z = (z_1, \dots, z_d) \\ \text{mit } z_i \in \{x_i, y_i\}}} F(z) \cdot \prod_{i=1}^d (\mathbb{1}_{\{z_i=y_i\}} - \mathbb{1}_{\{z_i=x_i\}});$$

es werden hierbei die Ecken von R bzgl. F ausgewertet und mit wechselndem Vorzeichen aggregiert. Die Funktion F heißt d -steigend bzw. d -wachsend, falls

$$\forall R := \bigtimes_{i=1}^d [x_i, y_i] \subseteq [0, 1]^d : \bigtimes_{i=1}^d \{x_i, y_i\} \subseteq \text{dom}(F) \implies \text{Vol}|_F(R) \geq 0$$

gilt.

2.2.3 Beispiel (2-steigende Funktionen). *Im zweidimensionalen Fall ist für beliebiges $F : [0, 1]^2 \rightarrow [0, 1]$ das zugehörige F -Volumen für Hyperrechtecke $R := [x_1, y_1] \times [x_2, y_2] \subseteq [0, 1]^2$ definiert als*

$$\text{Vol}|_F(R) := F(x_1, x_2) + F(y_1, y_2) - F(y_1, x_2) - F(x_1, y_2).$$

Die zu den Funktionen $\min, \max : [0, 1]^2 \rightarrow [0, 1]$ zugehörigen min- und max-Volumen sind konstruiert als

$$\begin{aligned}\text{Vol}|_{\min}(R) &:= (x_1 \vee x_2) + (y_1 \vee y_2) - (y_1 \vee x_2) - (x_1 \vee y_2), \\ \text{Vol}|_{\max}(R) &:= (x_1 \wedge x_2) + (y_1 \wedge y_2) - (y_1 \wedge x_2) - (x_1 \wedge y_2).\end{aligned}$$

Da $x_i \leqq y_i$ nach Konstruktion von R gilt, gelten für die betrachteten x_i, y_i die Abschätzungen

$$(x_1 \vee x_2) \leqq (y_1 \vee x_2), (x_1 \vee y_2) \leqq (y_1 \vee y_2).$$

Sei nun o. E. $x_1 \leqq x_2$. Dann gilt die Auswertung $x_1 = (x_1 \vee x_2) = (x_1 \vee y_2)$ und somit dann auch $(x_1 \vee x_2) - (x_1 \vee y_2) = 0$. Für die restlichen Terme gilt $(y_1 \vee y_2) - (y_1 \vee x_2) \geqq 0$ und damit $\text{Vol}|_{\min}(R) \geqq 0$. Also ist \min 2-wachsend. Für $R = [0, 1]^2$ ist jedoch $\text{Vol}|_{\max}(R) = -1$, d. h. \max ist nicht 2-steigend.

Mit der nun eingeführten Eigenschaft *d-steigend*, einer sogenannten *Erdungseigenschaft*, die ad hoc eingeführt wird, und der bereits bekannten Anforderung an die Marginale lässt sich nun eine alternative, aber äquivalente Copula-Definition formulieren.

2.2.4 Definition (Copula – Analytische Definition). Sei $C : [0, 1]^d \rightarrow [0, 1]$ eine Funktion. Sind die Eigenschaften

- (i.) *d-steigend*: Für alle Hyperrechtecke $R \subseteq [0, 1]^d$ gilt $\text{Vol}|_C(R) \geqq 0$.
- (ii.) *Erdung*: Für $u = (u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) \in [0, 1]^d$ gilt $C(u) = 0$.
- (iii.) *Marginale*: $C(1, \dots, 1, u_j, 1, \dots, 1) = u_j$ für $u_j \in [0, 1]$, $j \in \{1, \dots, d\}$. erfüllt, so wird C Copula genannt (vgl. Nelsen (2006): S. 45, Def. 2.10.6).

Neben einer Definition sind auch die Existenz, die Eindeutigkeit und Ansätze zur Konstruktion von Copulas von Interesse. Diese Aspekte werden mit dem folgenden Satz von Sklar behandelt.

2.2.5 Satz (Satz von Sklar (1959)). Sei $F : \mathbb{R}^d \rightarrow [0, 1]$ eine multivariate Verteilungsfunktion mit den Marginale $F_1, \dots, F_d : \mathbb{R} \rightarrow [0, 1]$.

Es existiert dann eine Copula $C : [0, 1]^d \rightarrow [0, 1]$ derart, dass F bezüglich C und F_1, \dots, F_d eine Darstellung der Form

$$F(x) = C(F_1(x_1), \dots, F_d(x_d)) \quad (x = (x_1, \dots, x_d) \in \overline{\mathbb{R}}^d)$$

besitzt, wobei die Copula C auf $\text{ran}(F_1) \times \dots \times \text{ran}(F_d) \subseteq [0, 1]^d$ eindeutig definiert ist.

Gilt zusätzlich die Stetigkeit der Marginale F_1, \dots, F_d , so ist die Copula C eindeutig auf $[0, 1]^d$ bestimmt als

$$C(u) = F(F_1^\leftarrow(u_1), \dots, F_d^\leftarrow(u_d)) \quad (u = (u_1, \dots, u_d) \in [0, 1]^d) \quad (2-5)$$

Beweis. Der Satz von Sklar ist im Original in Sklar (1959) und in alternativer Beweisführung in Rüschen (2009) bewiesen. \square

Mit der Konstruktion in Formel (2–5) werden die Parallelen zur Anwendung der ITS und PIT ($F(F^{-1}(U)) \sim \text{Unif}_{[0,1]}$ für $U \sim \text{Unif}_{[0,1]}$) ersichtlich. Da Copulas Abhängigkeitsstrukturen modellieren, ist es sinnvoll die intuitiven Kandidaten für Abhängigkeit zu betrachten: *Unabhängigkeit* sowie *Totale Positive* und *Totale Negative Abhängigkeit* im Sinne einer positiven und negativen Korrelation von ± 1 . Erster Kandidat wird von der Definition der stochastischen Unabhängigkeit

$$X \perp\!\!\!\perp Y \iff F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

motiviert, wobei hier $F_{X,Y}$ die gemeinsame Verteilungsfunktion zu den Zufallsvariablen X und Y , und F_X, F_Y die entsprechenden Marginale notieren.

2.2.6 Beispiel (Unabhängigkeits- bzw. Produktcopula). *Die Unabhängigkeits- bzw. Produktcopula $\Pi : [0, 1]^d \rightarrow [0, 1]$ ist definiert als Funktion*

$$\Pi(u) := \prod_{j=1}^d u_j \quad (u = (u_1, \dots, u_d) \in [0, 1]^d). \quad (2-6)$$

Diese Funktion ist trivialerweise eine Verteilungsfunktion auf $[0, 1]^d$. Wegen der Gleichung $\Pi(1, \dots, 1, u_j, 1, \dots, 1) = 1 \cdot \dots \cdot 1 \cdot u_j \cdot 1 \cdot \dots \cdot 1 = u_j$ liegen $\text{Unif}_{[0,1]}$ -Marginale vor, so dass Π eine Copula ist.

Die in Beispiel 2.2.3 gezeigten Funktionen werden zur Definition der sogenannten *Fréchet-Hoeffding-Schranken* herangezogen. Da die Maximumsfunktion nicht d -steigend ist, ist ohne Anpassung keine Copula dazu definierbar. Unabhängig dessen, ob nach Anpassung Copulas vorliegen, werden mit diesen allgemeine Abschätzbarkeitsaussagen formuliert.

2.2.7 Definition (Fréchet–Hoeffding-Schranken). *Sei $d \in \mathbb{N}, d \geq 2$. Die untere Fréchet-Hoeffding-Schranke $W^d : [0, 1]^d \rightarrow [0, 1]$ und die obere Fréchet-Hoeffding-Schranke $M^d : [0, 1]^d \rightarrow [0, 1]$ sind definiert als*

$$W^d(u) := \max \left\{ \left(\sum_{j=1}^d u_j - (d-1) \right), 0 \right\}, \quad (2-7)$$

$$M^d(u) := \min \{u_1, \dots, u_d\}, \quad (2-8)$$

für alle $u = (u_1, \dots, u_d) \in [0, 1]^d$.

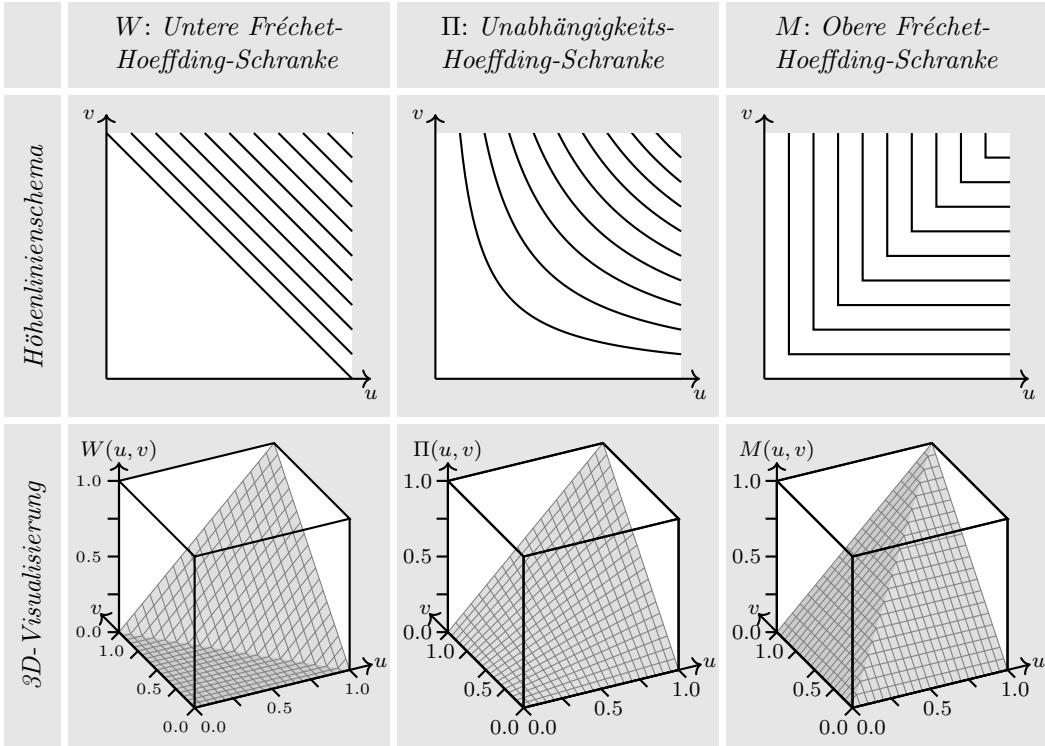


Abbildung 2.1: Höhenlinienschema und dreidimensionale Visualisierung der Fréchet-Hoeffding-Schranken und der Unabhängigkeitscopula. Quelle der Grafik: Eigene Erstellung, angelehnt an Hofert et al. (2018): S. 11 Fig. 2.1, S. 21 Fig. 2.8,

Die beiden Fréchet-Hoeffding-Schranken und die Unabhängigkeitscopula werden in Abbildung 2.1 skizziert. Die zentralen Eigenschaften und die Anwendung dieser Schranken werden im folgenden Lemma und Satz behandelt.

2.2.8 Lemma. *Für alle $d \in \mathbb{N}$, $d \geq 2$ ist die obere Fréchet-Hoeffding-Schranke M^d eine Copula. Im Fall $d = 2$ ist die untere Fréchet-Hoeffding-Schranke W^d auch eine Copula; dies gilt nicht für $d \in \mathbb{N}$, $d > 2$.*

In diesem Fall gilt die folgende schwächere Aussage:

$$\forall u \in [0, 1]^d \exists C_u : [0, 1]^d \rightarrow [0, 1] \text{ Copula} : C_u(u) = W^d(u).$$

Beweis. Der erste Teil wird in Nelsen (2006) als Bemerkung zum Theorem 2.10.11 formuliert. Der zweite Teil entspricht Theorem 2.10.13 ebenda. \square

Die beiden Copulas M^d und W^d der Fréchet-Hoeffding-Schranken werden in englischsprachiger Literatur als *Komonotonie-Copula* (*comonotonicity copula*) und *Gegenmonotonie-Copula* (*countermonotonicity copula*) bezeichnet. Zur Vereinfachung werden die Fréchet-Hoeffding-Schranken im zweidimensionalen Fall mit $W = W^2$ und $M = M^2$ notiert.

2.2.9 Satz (Fréchet-Hoeffding-Schranken). *Sei $C : [0, 1]^d \rightarrow [0, 1]$ eine beliebige Copula. Dann wird diese Copula beschränkt durch die Fréchet-Hoeffding-Schranken, genauer gilt*

$$W^d(u) \leqq C(u) \leqq M^d(u) \quad (u \in [0, 1]^d).$$

Beweis. Die Abschätzung wird in Nelsen (2006) in Thm. 2.10.12 als direkte Folgerung von Thm. 2.10.3 und 2.10.4 formuliert. \square

2.3 Copulafamilien

In diesem Abschnitt werden vier Copulafamilien vorgestellt. Die Familien werden implizit aus bekannten Verteilungsfunktionen konstruiert oder explizit formuliert. Die bereits vorgestellte Unabhängigkeitscopula und die Fréchet-Hoeffding-Schranken werden mit diesen Familien – falls möglich – in Beziehung gesetzt. Die Beispiele und Definitionen sind aus den Abschnitten 7.1 und 7.4 in McNeil et al. (2015), dem Abschnitt 2.1 in Hofert et al. (2018), und Kapitel 4 in Nelsen (2006) bezogen.

Ein naheliegender Kandidat für die *Übersetzung* zu einer Copula ist die Abhängigkeit innerhalb der multivariaten Normalverteilung $\mathcal{N}_d(\mu, \Sigma)$, die über die aus der Kovarianzmatrix Σ hergeleiteten Korrelationen modelliert wird. Da sowohl die multivariate Verteilung $\mathcal{N}_d(\mu, \Sigma)$ als auch dessen Marginalen $\mathcal{N}_1(\mu_j, \sigma_j^2)$, $j \in \{1, \dots, d\}$ die einzelnen Varianzen $\sigma_j^2 := \sigma_{jj}$ zur Modellierung beziehen, reicht hier eine Betrachtung der zu Σ zugehörigen Korrelationsmatrix P_Σ aus. Diese Matrix P_Σ wird über

$$P_\Sigma := D^{-1} \Sigma D^{-1}, \quad D := \text{diag}(\sigma_{11}^{1/2}, \dots, \sigma_{dd}^{1/2}) = (\text{diag}(\Sigma))^{1/2} \quad (2-9)$$

konstruiert (vgl. McNeil et al. (2015): S. 176).

2.3.1 Beispiel (Gauß'sche Copula / Normal-Copula). *Sei $\Phi_{\mu, \Sigma} : \mathbb{R}^d \rightarrow [0, 1]$ die Verteilungsfunktion der multivariaten Normalverteilung $\mathcal{N}_d(\mu, \Sigma)$ mit Erwartungsvektor $\mu \in \mathbb{R}^d$ und Kovarianzmatrix $\Sigma \in \mathbb{R}^{d \times d}$. Sei $\Phi : \mathbb{R} \rightarrow [0, 1]$ die Verteilungsfunktion der Standard-Normalverteilung $\mathcal{N}_1(0, 1)$.*

Unter Anwendung des Satzes von Sklar wird die Gauß'sche Copula bzw. Normal-Copula $C_P^{\text{Gauß}} : [0, 1]^d \rightarrow [0, 1]$ implizit definiert als

$$C_P^{\text{Gauß}}(u) := \Phi_{0, P}(\Phi^\leftarrow(u_1), \dots, \Phi^\leftarrow(u_d)) \quad (u \in [0, 1]^d) \quad (2-10)$$

und über die Korrelationsmatrix P parametrisiert. Im bivariaten Fall wird bei der Notation $C_\vartheta^{\text{Gauß}}$ die Korrelationsmatrix $P_\vartheta = \begin{pmatrix} 1 & \vartheta \\ \vartheta & 1 \end{pmatrix}$ zur Korrelation $\vartheta \in [-1, +1]$ betrachtet.

Diese Familie beinhaltet die Unabhängigkeitscopula Π sowie die Copulas der Fréchet-Hoeffding-Schranken M^d und W^2 . Diese werden über die Einheitsmatrix I_d , Eins-Matrix J_d und der Matrix $P_{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ modelliert (vgl. McNeil et al. (2015): S. 226-227).

Mit dem gleichen Ansatz lässt sich die Abhängigkeitsstruktur der multivariaten t -Verteilung $t_d(\mu, \Sigma, \nu)$ als implizite Copula formulieren.

2.3.2 Beispiel (t -Copula). Die t -Copula $C_{P,\nu}^t : [0, 1]^d \rightarrow [0, 1]$ mit Freiheitsgrad $\nu > 2$ und Korrelationsmatrix P modelliert die Abhängigkeitsstruktur einer t -Verteilung $t_d(\cdot, P, \nu)$. Die Copula ist unter Anwendung des Satzes von Sklar definiert als

$$C_{P,\nu}^t(u) := t_{0,P,\nu}(t_\nu^\leftarrow(u_1), \dots, t_\nu^\leftarrow(u_d)) \quad (u \in [0, 1]^d), \quad (2-11)$$

wobei $t_{0,P,\nu} : \mathbb{R}^d \rightarrow [0, 1]$ die Verteilungsfunktion zur $t_d(0, P, \nu)$ -Verteilung ist. Die Verteilungsfunktion $t_\nu : \mathbb{R} \rightarrow [0, 1]$ gehört zur univariaten t -Verteilung mit ν Freiheitsgraden. Im bivariaten Fallkürzt die Schreibweise $C_{\vartheta,\nu}^t$ die t -Copula zur Korrelationsmatrix $P_\vartheta = \begin{pmatrix} 1 & \vartheta \\ \vartheta & 1 \end{pmatrix}$ ab.

Analog zur Gauss-Copula lassen sich für die t -Copula die Fréchet-Hoeffding-Schranken M^d und W^2 herleiten. Die Unabhängigkeits-Copula ist nicht über die t -Copula modellierbar, da im Fall der t -Verteilung Unkorreliertheit nicht Unabhängigkeit impliziert (vgl. McNeil et al. (2015): S. 228).

Es gibt Copulafamilien, die explizit formuliert werden können. Zu diesen gehören beispielsweise die in den nächsten Beispielen gezeigten Familien der Clayton- und Gumbel-(Hougaard-)Copulas.

2.3.3 Beispiel (Clayton-Copula). Die Clayton-Copula $C_\vartheta^{\text{Cl}} : [0, 1]^d \rightarrow [0, 1]$ zum Parameter $\vartheta \in \Theta = \mathbb{R}_>$ ist definiert als

$$C_\vartheta^{\text{Cl}}(u) := \left(\sum_{j=1}^d \frac{1}{u_j^\vartheta} - (d-1) \right)^{-1/\vartheta} \quad (u \in [0, 1]^d). \quad (2-12)$$

Im bivariaten Fall ist die Clayton-Copula auch für den größeren Parameterraum $\tilde{\Theta} = [-1, \infty) \setminus \{0\}$ definiert; diese Clayton-Copula ist gegeben als

$$C_\vartheta^{\text{Cl}}(u) := \left(\left(\frac{1}{u_1^\vartheta} + \frac{1}{u_2^\vartheta} - 1 \right) \vee 0 \right)^{-1/\vartheta} \quad (u \in [0, 1]^2). \quad (2-13)$$

Die drei Copulas Π , M^d und W^2 können über die Clayton-Copula hergeleitet werden. In d Dimensionen gelten die Grenzwertbetrachtungen

$$C_0^{\text{Cl}}(u) := \lim_{\vartheta \rightarrow 0} C_\vartheta^{\text{Cl}}(u) = \Pi(u),$$

$$C_\infty^{\text{Cl}}(u) := \lim_{\vartheta \rightarrow \infty} C_\vartheta^{\text{Cl}}(u) = M^d(u),$$

wobei $u \in [0, 1]^d$ ist. Der erweiterte bivariate Fall nimmt für $\vartheta = -1$ die untere Fréchet-Hoeffding-Schranke an, da die Gleichung

$$C_{-1}^{\text{Cl}}(u) = ((u_1 + u_2 - 1) \vee 0) = W^2(u)$$

für alle $u \in [0, 1]^2$ erfüllt ist (vgl. McNeil et al. (2015): S. 229; Hofert et al. (2018): S. 16, Example 2.1.5).

2.3.4 Beispiel (Gumbel-Hougaard-Copula). Die Gumbel-(Hougaard-)Copula $C_\vartheta^{\text{GH}} : [0, 1]^d \rightarrow [0, 1]$ zum Parameter $\vartheta \in \Theta = [1, \infty)$ ist definiert als

$$C_\vartheta^{\text{GH}}(u) := \exp\left(-\left(\sum_{j=1}^d (-\ln u_j)^\vartheta\right)^{1/\vartheta}\right) \quad (u \in [0, 1]^d). \quad (2-14)$$

Die Gumbel-Copula für $\vartheta = 1$ stimmt mit der Unabhängigkeitscopula Π überein; dies folgt direkt aus der Gleichung

$$C_1^{\text{GH}}(u) = \exp\left(-\sum_{j=1}^d (-\ln u_j)\right) = \exp\left(\ln\left(\prod_{j=1}^d u_j\right)\right) = \prod_{j=1}^d u_j = \Pi(u),$$

welche für alle $u \in [0, 1]^d$ erfüllt ist. Weiter erlaubt die Familie der Gumbel-Copulas das Herleiten von M^d . Die Grenzwertbetrachtung

$$C_\infty^{\text{GH}}(u) := \lim_{\vartheta \rightarrow \infty} C_\vartheta^{\text{GH}}(u) = M^d(u) \quad (u \in [0, 1])$$

ist für den bivariaten Fall gültig (vgl. McNeil et al. (2015): S. 228), wobei der bivariate Fall ohne Einschränkungen auf den multivariaten Fall verallgemeinert werden kann (vgl. Nelsen (2006): S. 153, Example 4.25).

Die in den Beispielen 2.3.1 bis 2.3.4 vorgestellten Copulas sind in Abbildung 2.2 für den bivariaten Fall dargestellt.

2.4 Weitergehende Definitionen für Copulas

Für die statistischen Tests und Methoden in Kapitel 3 werden weitere Definitionen an Abhängigkeitsmaße, Copula-Dichten und Konstruktionen benötigt. Diese werden in den folgenden drei Teilabschnitten behandelt.

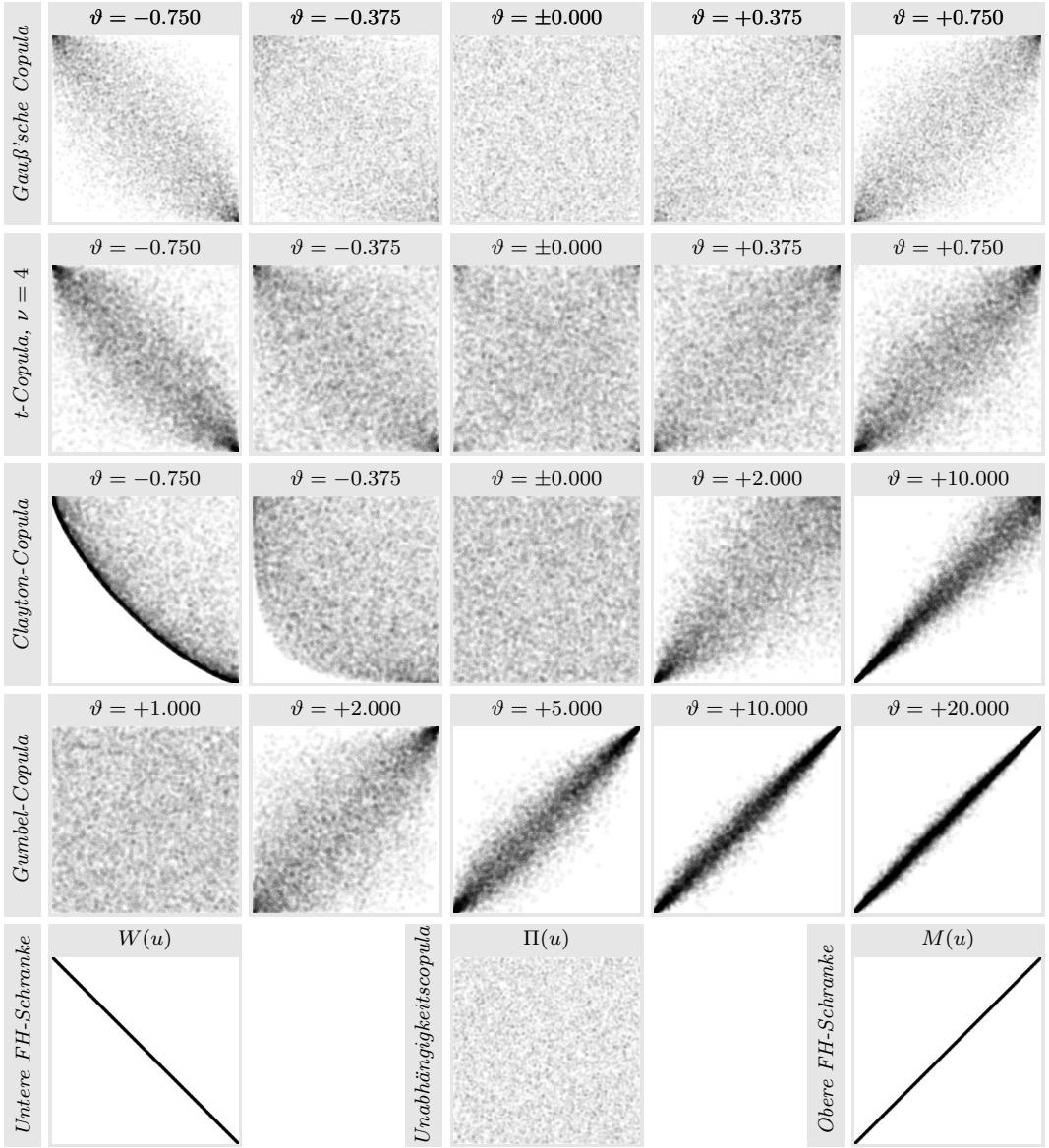


Abbildung 2.2: Streudiagramme der bivariaten Gauss-, t-, Clayton- und Gumbel-Copula für verschiedene Parameter ϑ sowie der Frechét-Hoeffding-Schranken und der Unabhängigkeitscopula. Es sind Stichproben der Größe $n = 10\,000$ auf $[0, 1]^2$ dargestellt. Quelle der Grafik: Eigene Erstellung.

Korrelationskoeffizienten

Neben dem Pearson'schen Korrelationskoeffizienten, welcher die Korrelation zwischen den Zufallsvariablen $X_1 \sim F_1$ und $X_2 \sim F_2$ berechnet als

$$\rho(X_1, X_2) := \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}, \quad (2-15)$$

kann ferner die Korrelation bzgl. der in Wahrscheinlichkeit transformierten Zufallsvariablen $U_j := F_j(X_j)$, $j \in \{1, 2\}$, betrachtet werden.

2.4.1 Definition (Spearman'sches Rho). *Sei $X = (X_1, X_2) \sim F$ ein Zufallsvektor bzgl. einer Verteilung F mit stetigen Marginalen F_1 und F_2 . Dann ist das Spearman'sche Rho als Korrelationskoeffizient definiert als*

$$\rho_S(X_1, X_2) := \rho(F_1(X_1), F_2(X_2)),$$

wobei $\rho(\cdot, \cdot)$ die Funktion für das Pearson'sche Rho aus Formel 2-15 ist (vgl. McNeil et al. (2015): S. 245, Proposition 7.34).

Bei der Definition der Pearson'schen und Spearman'schen Korrelationskoeffizienten werden die Differenzen $(X_j - \mathbb{E}[X_j])$ und $(U_j - 0.5)$ der Zufallsvariablen bzw. Transformierten zum jeweiligen Erwartungswert herangezogen. Das Kendall'sche Tau als ein weiteres Abhängigkeitsmaß nutzt einen Ansatz auf Basis von Konkordanz und Diskordanz: Sei $x_1, \dots, x_n \in \mathbb{R}^2$ eine Stichprobe von Elementen $x_i = (x_{i1}, x_{i2})$. Zwei Elemente x_i und x_j mit $i \neq j$ heißen genau dann *konkordant*, wenn die Äquivalenz

$$x_{i1} > x_{j1} \iff x_{i2} > x_{j2} \quad (2-16)$$

erfüllt ist. Gilt andererseits die Äquivalenz

$$x_{i1} > x_{j1} \iff x_{i2} < x_{j2}, \quad (2-17)$$

so heißen x_i und x_j diskordant. Diese Definitionen können alternativ über die Vorzeichenfunktion $\text{sign}(x) := \mathbb{1}_{\{x>0\}} - \mathbb{1}_{\{x<0\}}$ formuliert werden:

$$(2-16) \iff \text{sign}((x_{i1} - x_{j1})(x_{i2} - x_{j2})) = +1,$$

$$(2-17) \iff \text{sign}((x_{i1} - x_{j1})(x_{i2} - x_{j2})) = -1.$$

Wird angenommen, dass zwei Zufallsvariablen X_1 und X_2 in einer positiven Abhängigkeitsbeziehung stehen, sprich dass das Wachstum von X_1 und X_2 voneinander abhängen, so kann gefolgert werden, dass auch die Wahrscheinlichkeit der Konkordanz höher sein muss als die Wahrscheinlichkeit der Diskordanz (vgl. McNeil et al. (2015): S. 244). Dies motiviert die folgende Definition des Kendall'schen Taus.

2.4.2 Definition (Kendall'sches Tau). *Sei $X = (X_1, X_2) \sim F$ ein stetig verteilter Zufallsvektor. Dann ist das Kendall'sche Tau als Koeffizient für die Korrelation zwischen X_1 und X_2 definiert als*

$$\tau(X_1, X_2) := \mathbb{E}[\text{sign}((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2))], \quad (2-18)$$

wobei $\tilde{X} = (\tilde{X}_1, \tilde{X}_2)$ eine i.i.d. Kopie von X ist.

Im multivariaten Fall, d. h. für $X = (X_1, \dots, X_d)$ und $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_d)$ lässt sich das Kendall'sche Tau verallgemeinern als Kovarianzmatrix

$$\tau(X) = \text{Cov}(S), \quad S = (S_1, \dots, S_d), \quad S_j := \text{sign}(X_j - \tilde{X}_j)$$

(vgl. McNeil et al. (2015): S. 244, Definition 7.31).

Im Umgang mit bivariaten Copulas kann das Kendall'sche Tau explizit über die Copula $C : [0, 1]^2 \rightarrow [0, 1]$ berechnet werden; das wird im folgenden Lemma gezeigt.

2.4.3 Lemma (Kendall'sches Tau für bivariate Copulas). *Sei $F : \mathbb{R}^2 \rightarrow [0, 1]$ eine stetige Verteilungsfunktion mit Copula $C : [0, 1]^2 \rightarrow [0, 1]$. Sei weiter $X = (X_1, X_2) \sim F$ ein Zufallsvektor. Dann gilt*

$$\tau(X_1, X_2) = 4 \int_{[0,1]^2} C(u) dC(u) - 1. \quad (2-19)$$

Beweis. Der Beweis ist in McNeil et al. (2015) als Beweis zur Proposition 7.32 zu finden (vgl. McNeil et al. (2015): S. 244-245). \square

Die Formel (2-19) erlaubt eine direkte Berechnung des Kendall'schen Taus über die Copula selbst; es werden keine Informationen über die Marginalen benötigt. Weiter kann mit der Formel (2-19) der Parameter $\vartheta \in \Theta$ einer Copulafamilie $\{C_\vartheta : \vartheta \in \Theta\}$ über die Beziehung

$$\tau \in [-1, +1] \quad \xleftrightarrow{(2-19)} \quad C : [0, 1]^2 \rightarrow [0, 1] \quad \xrightarrow[\substack{\exists \vartheta \in \Theta: \\ C=C_\vartheta}]{} \quad \vartheta \in \Theta$$

für die Berechnung des Koeffizienten genutzt werden.

2.4.4 Beispiel (Berechnung der Kendall'schen Taus aus Copula-Parameter). *Für die vier parametrischen Copulafamilien aus den Beispielen 2.3.1 bis 2.3.4 kann das Kendall'sche Tau im bivariaten Fall über die Parameter $\vartheta \in \Theta$ berechnet werden.*

Für die Gauß'sche Copula und die t-Copula mit $\vartheta \in \Theta = [-1, 1]$ gilt die Abbildungsvorschrift

$$\tau : [-1, 1] \rightarrow [-1, 1], \quad \tau(\vartheta) := \frac{2}{\pi} \arcsin(\vartheta).$$

Das Kendall'sche Tau der bivariaten Clayton-Copula mit dem Parameterraum $\Theta = (0, \infty)$ wird berechnet über die Vorschrift

$$\tau : [0, \infty) \rightarrow [0, 1], \quad \tau(\vartheta) := \frac{\vartheta}{\vartheta + 2},$$

und bzgl. des erweiterten Parameterraums $\tilde{\Theta} = [-1, \infty) \setminus \{0\}$ gilt

$$\tau : [-1, \infty) \rightarrow [-1, 1], \quad \tau(\vartheta) := \frac{\vartheta}{\vartheta + 2},$$

wobei in beiden Fällen die Unabhängigkeitscopula ($\vartheta \rightarrow 0$) ebenso abgedeckt wird. Für die Gumbel-Copula mit $\vartheta \in \Theta = [1, \infty)$ gilt die Vorschrift

$$\tau : [1, \infty) \rightarrow [0, 1], \quad \tau(\vartheta) := 1 - \frac{1}{\vartheta}$$

zur Berechnung des Kendall'schen Taus (vgl. Czado (2019): S. 55, Table 3.2).

Copula-Konstruktionen

2.4.5 Definition (Survival-Copula). Sei $C : [0, 1]^d \rightarrow [0, 1]$ eine beliebige Copula und dazu $U = (U_1, \dots, U_d) \sim C$ ein Zufallsvektor. Die zu C zugehörige Survival-Copula $\bar{C} : [0, 1]^d \rightarrow [0, 1]$ beschreibt die Verteilung des Vektors $1 - U = (1 - U_1, \dots, 1 - U_d)$. Es gilt also die Äquivalenz

$$U \sim C \iff 1 - U \sim \bar{C}.$$

zwischen diesen beiden Copulas. Im bivariaten Fall wird die Survival-Copula konstruiert als Funktion $\bar{C} : [0, 1]^d \rightarrow [0, 1]$ mit

$$\bar{C}(u_1, u_2) := C(1 - u_1, 1 - u_2) + u_1 + u_2 - 1$$

(vgl. Kojadinovic (2017): S. 9; Nelsen (2006): S. 32, Formel (2.6.1)).

Neben der Survival-Copula sind zwei weitere Copula-Konstruktionen von Interesse. Die Definitionen werden entsprechend der Nutzung in den Abschnitten 3.4.2 und 3.4.3 in Hofert et al. (2018) definiert.

Eine gemischte Copula (Mixture of Copulas) bezeichnet eine gewichtete Konvexkombination von einzelnen Copulas.

2.4.6 Definition (Gemischte Copulas). Seien $C_1, \dots, C_m : [0, 1]^d \rightarrow [0, 1]$ Copulas und $w = (w_1, \dots, w_m) \in [0, 1]^m$ ein Gewichtsvektor mit $\sum_{i=1}^m w_i = 1$. Die (gewichtete) Mischung der Copulas C_1, \dots, C_m wird über die Copula

$$\text{mix}_w(C_1, \dots, C_n)(u) := \sum_{i=1}^m w_i C_i(u) \quad (u \in [0, 1]^d)$$

konstruiert. Für die gemischte Copula sind Zufallsvektoren V darstellbar als

$$V = \sum_{i=1}^m M_i U_i \sim \text{mix}_w(C_1, \dots, C_n),$$

wobei $U_i \sim C_i$, $i \in \{1, \dots, m\}$, unabhängige Zufallsvektoren seien und der Vektor $M = (M_1, \dots, M_m) \sim \mathcal{M}(1, w)$ multinomialverteilt und unabhängig von den U_1, \dots, U_m sei (vgl. Hofert et al. (2020): S. 127-128).

Gemischte Copulas finden im Abschnitt 5.2 Anwendung; beim Jittering, d. h. bei dem Hinzufügen eines Rauschens in die Stichprobe, wird in einer Methode das Rauschen bzgl. einer gemischten Copula unter Berücksichtigung des Kendall'schen Taus generiert. Eine andere Konstruktion nach Khoudraji (1995) wird zur Modellierung asymmetrischer bzw. nicht-austauschbaren Copulas (siehe Abschnitt 3.3) verwendet. Hierfür wird das Khoudraji's Device eingeführt und darüber sogenannte Khoudraji-Copulas definiert.

2.4.7 Definition (Khoudraji's Device). *Es seien $C_1, C_2 : [0, 1]^d \rightarrow [0, 1]$ zwei Copulas und $s \in [0, 1]^d$ ein Formparametervektor.*

Das Khoudraji's Device $\text{kho}_s(C_1, C_2)$ mit Parameter s konstruiert aus den beiden Copulas C_1 und C_2 eine neue Copula

$$K(u) := \text{kho}_s(C_1, C_2)(u) := C_1(u^{1-s})C_2(u^s) \quad (u \in [0, 1]^d),$$

wobei hier die Schreibweise $u^t = (u_1^{t_1}, \dots, u_d^{t_d})$ für $u, t \in [0, 1]^d$ gelte.

Des Weiteren lassen sich K -verteilte Zufallsvektoren aus C_1 - und C_2 -verteilten Vektoren konstruieren: Seien $U_1 \sim C_1$ und $U_2 \sim C_2$ unabhängige Zufallsvektoren der Form $U_j = (U_{j1}, \dots, U_{jd})$ und $s \in [0, 1]$ der Formparameter des Khoudraji's Device. Mit den Konventionen $u^{1/0} = 0$ für $u \in [0, 1)$ und $1^{1/0}$ gilt für den konstruierten Vektor

$$V := (V_1, \dots, V_d) := \left(U_{11}^{\frac{1}{1-s_1}} \vee U_{21}^{\frac{1}{s_1}}, \dots, U_{1n}^{\frac{1}{1-s_n}} \vee U_{2d}^{\frac{1}{s_d}} \right),$$

dass $V \sim K := \text{kho}_s(C_1, C_2)(u)$ (vgl. Hofert et al. (2018): S. 120-121).

Über das Khoudraji's Device lassen sich aus bekannten Copulafamilien neue Copulafamilien generieren. Unter der Wahl der Unabhängigkeitscopula $\Pi : [0, 1]^d \rightarrow [0, 1]$ als eine der beiden Copulas im Khoudraji's Device werden die sogenannten Khoudraji-Copulas definiert.

2.4.8 Beispiel (Khoudraji-Copulas). *Zur Copulafamilie $\{C_\vartheta : \vartheta \in \Theta\}$ ist die zugehörige Khoudraji-Copulafamilie $\{\text{kho}_s(C_\vartheta, \Pi) : \vartheta \in \Theta\}$ über das Khoudraji's Device bzgl. der Produktcopula Π definierbar. Die Khoudraji-Copulas zur Copula C_ϑ besitzen dabei die Darstellung*

$$\text{kho}_s C_\vartheta(u) := \text{kho}_s(C_\vartheta, \Pi)(u) := u^{1-s} C_\vartheta(u^s) \quad (u \in [0, 1]^d).$$

Für die Clayton-Copulas sowie die anderen vorgestellten Copulafamilien werden diese konstruierten Copulas als Khoudraji-Clayton-Copula, Khoudraji-Gauß-Copula etc. betitelt.

Partielle Ableitungen und Dichten von Copulas

Für Aussagen über die Konvergenz der Teststatistiken in Abschnitt 3.3 sind Regularitätsanforderungen an die partiellen Ableitungen $\dot{C}_j(u)$, $j \in \{1, \dots, d\}$, zu erfüllen. Diese werden für $u \in [0, 1]^d$ mit $0 \leq u \leq 1$ definiert als

$$\dot{C}_j(u) := \frac{\partial}{\partial u_j} C(u) := \begin{cases} \lim_{h \rightarrow 0} \frac{C(u + 2he_j) - 0}{2h}, & \text{für } u_j = 0, \\ \lim_{h \rightarrow 0} \frac{C(u + he_j) - C(u - he_j)}{2h}, & \text{für } u_j \in (0, 1), \\ \lim_{h \rightarrow 0} \frac{C(u) - C(u - 2he_j)}{2h}, & \text{für } u_j = 1. \end{cases}$$

Hierbei bezeichnet e_j den j -ten Einheitsvektor (vgl. Segers (2012): S. 767; in Anlehnung an die beidseitige Notation in Genest et al. (2012): S. 818).

Neben den partiellen Ableitungen \dot{C}_j wird in dieser Arbeit die Dichtefunktion der Copulas für die Definition von Maximum-Pseudo-Likelihood-Schätzern (MPLE) benötigt. Falls existent, d. h. falls die Copula die absolute Stetigkeit erfüllt, wird die Dichtefunktion auf folgende Weise definiert:

2.4.9 Definition (Dichtefunktion einer Copula C). *Sei $C : [0, 1]^d \rightarrow [0, 1]$ eine absolut stetige Copula. Die Dichtefunktion $c : [0, 1]^d \rightarrow \mathbb{R}_{\geq}$ zur Copula C ist definiert als*

$$c(u) := \frac{\partial}{\partial u} C(u) := \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \cdots \partial u_d} \quad (u \in [0, 1]^d) \quad (2-20)$$

(vgl. McNeil et al. (2015): S. 233).

Alternativ kann die Dichtefunktion über die Darstellung (2–5) aus dem Satz von Sklar bestimmt werden.

2.4.10 Lemma. *Sei $F : \mathbb{R}^d \rightarrow [0, 1]$ eine absolut stetige Verteilungsfunktion mit Copula $C : [0, 1]^d \rightarrow [0, 1]$ und strikt monoton wachsenden, stetigen Marginalen $F_1, \dots, F_d : \mathbb{R} \rightarrow [0, 1]$.*

Dann ist die Darstellung $C(u) = F(F_1^\leftarrow(u_1), \dots, F_d^\leftarrow(u_d))$ differenzierbar und die Dichtefunktion $c : [0, 1]^d \rightarrow \mathbb{R}_{\geq}$ zur Copula C ist darstellbar als

$$c(u_1, \dots, u_d) := \frac{f(F_1^\leftarrow(u_1), \dots, F_d^\leftarrow(u_d))}{\prod_{j=1}^d f_j(F_j^\leftarrow(u_j))} \quad (u \in [0, 1]^d) \quad (2-21)$$

Es bezeichnen f, f_1, \dots, f_d die Dichtefunktionen von F, F_1, \dots, F_d (vgl. McNeil et al. (2015): S. 233).

Diese alternative Darstellung erleichtert die Handhabe von Dichten, da diese über die (multivariaten) Verteilungsfunktionen bestimmt werden können.

2.5 Betrachtung bivariater Copula-Modelle

In dieser Arbeit liegt der Schwerpunkt auf bivariate Copula-Modelle und Methoden. Die ausschließliche Betrachtung des bivariaten Falls wird über die folgenden Punkte begründet:

Einerseits sind gewisse Copulas im multivariaten Fall nicht oder nur eingeschränkt definierbar. Als Beispiele sind die Clayton-Copulas und die untere Frechét-Hoeffding-Schranke zu nennen. Ebenso wurden in der Forschung einige der statistischen Tests aufgrund der Komplexität hauptsächlich im bivariaten Fall behandelt, hierzu ist der Test auf Austauschbarkeit in Abschnitt 3.3 vorwegzunehmen: Bei diesem Test für eine d -dimensionale Copula sind die Permutationen auf $\{1, \dots, p\}$ zu berücksichtigen, wodurch die Rechenkomplexität im multivariaten Fall steigt.

Andererseits können multivariate Modellierungen über mehrere bivariate Copula-Modelle durchgeführt werden. Mit sogenannten Vine-Copulas können multivariate Verteilungen über deren Marginale und bivariate Copulas konstruiert werden. Diese bilden eine Baumstruktur, genannt Regular Vine Tree Structure (vgl. Czado (2019): S. 77-79, 101-103; Krämer (2011): S. 7-9).

Die bereits bekannte Formel (2-21) erlaubt die Berechnung der Dichtefunktion f einer multivariaten Verteilungsfunktion F als

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \cdot \prod_{j=1}^d f_j(x_j) \quad (x \in \mathbb{R}^d),$$

wobei f_1, \dots, f_d und c die Dichten der Marginale F_1, \dots, F_d und der Copula C bzgl. F sind. Das folgende Lemma zeigt eine alternative Darstellung der Dichten f und c ohne die Nutzung einer multivariaten Copula.

2.5.1 Lemma (Kanonische Pair-Copula-Konstruktion). *Sei $F : \mathbb{R}^d \rightarrow [0, 1]$ eine absolut stetige multivariate Verteilungsfunktion. $F_1, \dots, F_d : \mathbb{R} \rightarrow [0, 1]$ und $C : [0, 1]^d \rightarrow [0, 1]$ sind die Marginale und die Copula bzgl. F . Es bezeichnen f, f_1, \dots, f_d und c die jeweiligen Dichtefunktionen.*

Dann besitzt die Dichte $f : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq}$ eine (kanonische) Pair-Copula-Konstruktion (C-PCC) über bivariate Copula-Dichten. Diese ist gegeben als:

$$f(x_1, \dots, x_d) = \left[\prod_{j=1}^{d-1} \prod_{k=1}^{j-1} c_{j,j+k|1,\dots,j-1} \right] \left[\prod_{j=1}^d f_j(x_j) \right] \quad (x \in \mathbb{R}^d), \quad (2-22)$$

wobei die Faktoren der Pair-Copulas $c_{j,j+k|1,\dots,j-1}$ definiert sind als

$$c_{j,j+k|1,\dots,j-1} := c_{j,j+k|1,\dots,j-1}(F(x_j|x_1, \dots, x_{j-1}), F(x_{j+k}|x_1, \dots, x_{j-1}))$$

(vgl. Czado (2019): S. 89-90, Theorem 4.8; Czado (2010): S. 2-3).

Der Beweis ist im Appendix E.2 zu finden. Unter den Voraussetzungen der beiden Lemmata 2.4.10 und 2.5.1 folgt aus den beiden Gleichungen (2–21) und (2–22) direkt die Darstellbarkeit multivariater Copulas über bivariate Copulas:

$$c(u_1, \dots, u_d) = \prod_{j=1}^{d-1} \prod_{k=1}^{j-1} c_{j,j+k|1,\dots,j-1}(u_j, u_{j+k}|u_1, \dots, u_{j-1}) \quad (u \in [0, 1]^d).$$

Diese Gleichung sagt gleichzeitig aus, dass über bivariate Copulas multivariate Modelle konstruiert werden können, sodass eine primäre Betrachtung von bivariaten Copulas in dieser Arbeit keine Einschränkung darstellt. Copula-Modelle in höheren Dimensionen werden aus diesem Grund in dieser Arbeit lediglich referenziert und nicht tiefergehend in der Simulationsstudie behandelt.

Kapitel 3

Bindungen und rangbasierte Copula-Methoden

Eine der zentralen Aussagen in Kapitel 2 wurde im Satz von Sklar 2.2.5 formuliert: Die Copula $C : [0, 1]^d \rightarrow [0, 1]$ zu der Verteilung $F : \mathbb{R}^d \rightarrow [0, 1]$ ist auf der Menge $\text{ran}(F_1) \times \dots \times \text{ran}(F_d) \subseteq [0, 1]^d$ bzgl. der Marginalen $F_1, \dots, F_d : \mathbb{R} \rightarrow [0, 1]$ eindeutig definiert. Weiter ist die Copula auf $[0, 1]^d$ eindeutig, falls alle Marginale F_1, \dots, F_d die Stetigkeit erfüllen.

Einhergehend mit einer Stetigkeitsannahme ist, dass die zugehörige Realisierung bzw. Stichprobe fast sicher einzigartige Werte innehaben sollte. Ein veranschaulichendes Beispiel: Für die Zufallsvariable $X \sim G$ zur stetigen Verteilungsfunktion $G : \mathbb{R}^d \rightarrow [0, 1]$ gilt $\mathbb{P}(X = x) = 0$ für jeden Punkt $x \in \mathbb{R}^d$. Dies lässt sich zu $\mathbb{P}(X_1 = X_2) = 0$ für i.i.d. $X_1, X_2 \sim G$ weiterführen. Die Zufallsvariablen X_1 und X_2 sind also fast sicher ungleich.

Der Fall, dass in einer Folge oder Stichprobe (x_1, \dots, x_n) ein Wertepaar (x_i, x_j) mit $x_i = x_j$ für $i \neq j$ vorliegt, wird als Bindung bezeichnet. In diesem Fall sind x_i und x_j nicht unterscheidbar.

Die Begriffe Bindung und Rang werden im Abschnitt 3.1 formal definiert und diskutiert. Im Abschnitt 3.2 werden die empirischen Verteilungsfunktionen und Copulas als rangbasierte Schätzer behandelt. Unter Berücksichtigung von Regularitätsanforderungen wird über den empirischen Copula-Prozess eine Konvergenzaussage formuliert.

Der dritte Abschnitt 3.3 behandelt rangbasierte Copula-Methoden bzw. statistische Tests in sowohl theoretischen als auch praxisorientierten Fassungen. Über den empirischen Copula-Prozess wird die Konvergenz der Teststatistiken unter angenommener Nullhypothesen festgestellt. Zur Approximation der p-Werte sind jeweils nachgehend Bootstrap-Algorithmen angegeben; diese sind nicht speziell zur Handhabe von gebundenen Datensätzen ausgezeichnet worden.

3.1 Bindungen und Ränge

Dieser Abschnitt zu Bindungen und Rangvergabe nutzt die Definitionen 2.3.1 und 3.1.2 in Hafner (2001) und Example 6.1.1 in Hofert et al. (2018). Die Definitionen werden für mehrdimensionale Stichproben präzisiert.

3.1.1 Definition (Bindungen). *Für eine beliebige Stichprobe $x_1, \dots, x_n \in \mathbb{R}^d$ mit $x_i = (x_{i1}, \dots, x_{id})$, $i \in \{1, \dots, n\}$, sind folgende Eigenschaften definiert:*

- (i.) Ungebundene / bindungsfreie Stichprobe: *Gilt $x_i \neq x_j$ für alle $i \neq j$, so nennt man die Stichprobe x_1, \dots, x_n ungebunden bzw. bindungsfrei.*
- (ii.) Bindung (der Größe $k \geq 2$): *Existieren zu einem Element x_{i_1} mit dem Wert $x_{i_1} = x$ genau $(k - 1)$ weitere Elemente x_{i_2}, \dots, x_{i_k} , sodass*

$$x = x_{i_1} = x_{i_2} = \dots = x_{i_k},$$

so nennt man $(x_{j_1}, \dots, x_{j_k})$ eine Bindung der Größe k vom Wert x . In der Literatur werden Bindungen der Größe k auch $(k - 1)$ -fache Bindungen genannt (vgl. Hafner (2001): S. 16, Definition 2.3.1).

Bei mehrdimensionalen Stichproben kann es sinnvoll sein einen weiteren Bindungsbegriff einzuführen, der Bindungen in den Koordinaten berücksichtigt:

- (iii.) Vollständige und teilweise Bindungen: *Existieren zu k -vielen Elementen x_{i_1}, \dots, x_{i_k} eine nichtleere Indexmenge $J \subseteq \{1, \dots, d\}$ und dazu eine Projektionsabbildung $\pi_J : \mathbb{R}^d \rightarrow \mathbb{R}^{|J|}$, $x_i = (x_{i1}, \dots, x_{id}) \mapsto (x_{ij})_{j \in J}$ derart, dass die Projektionen der x_{i_1}, \dots, x_{i_k} eine Bindung*

$$\pi_J(x_{i_1}) = \pi_J(x_{i_2}) = \dots = \pi_J(x_{i_k})$$

der Größe k bilden, so liegt bei ebendiesen Elementen eine teilweise Bindung in $|J|$ -vielen Koordinaten und der Größe k vor.

Eine Bindung im Sinne der Definition (ii.), d. h. falls die teilweise Bindung für alle Koordinaten $J = \{1, \dots, d\}$ gilt, wird in diesem Zusammenhang als vollständige Bindung bezeichnet.

3.1.2 Beispiel (Bindungen in mehrdimensionalen Stichproben). *Für die zweidimensionale Stichprobe x_1, \dots, x_{12} mit den Werten*

$$\begin{pmatrix} 1 \\ 6 \end{pmatrix}, \begin{pmatrix} 8 \\ 9 \end{pmatrix}, \begin{pmatrix} 3 \\ 8 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 6 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \end{pmatrix}, \begin{pmatrix} 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

können sowohl vollständige als auch teilweise Bindungen gesichtet werden.

Die beiden Vektoren x_4 und x_{11} bilden eine (vollständige) Bindung der Größe 2 vom Wert (3, 9). Es sind acht teilweise gebundene Elemente in drei Bindungen bzgl. der ersten Koordinate festzustellen:

x_1, x_7	teilw. Bindung der Größe 2 vom Wert 1,
x_3, x_4, x_8, x_{11}	teilw. Bindung der Größe 4 vom Wert 3,
x_5, x_{12}	teilw. Bindung der Größe 2 vom Wert 6.

Die teilweisen Bindungen zur zweiten Koordinate umfassen neun der zwölf Elemente der Stichprobe in vier Bindungen:

x_1, x_{10}	teilw. Bindung der Größe 2 vom Wert 6,
x_2, x_4, x_{11}	teilw. Bindung der Größe 3 vom Wert 9,
x_3, x_6	teilw. Bindung der Größe 2 vom Wert 8,
x_5, x_9	teilw. Bindung der Größe 2 vom Wert 1.

Es ist erkennbar, dass eine vollständige Bindung der Größe 2 vorliegt und jedes Element dieser Stichprobe teilweise gebunden ist.

3.1.3 Bemerkung. Das Beispiel 3.1.2 zeigt auf, dass der Umfang und die Struktur der teilweisen und vollständigen Bindungen bei mehrdimensionalen Stichproben stark voneinander abweichen können.

Für die beiden Elemente $x_9 = (2, 1)$ und $x_{12} = (6, 4)$ gilt $x_9 \neq x_{12}$, sodass diese Elemente untereinander ungebunden sind. Wird nun $x_5 = (6, 1)$ hinzugezogen, so sind x_5 und x_9 sowie x_5 und x_{12} teilweise gebunden. Die Elemente x_9 und x_{12} stehen somit wegen der teilweisen Bindungen zu x_5 in einer indirekten Beziehung zueinander.

Bei der Identifizierung von Bindungen innerhalb einer Stichprobe sind sowohl eine ausschließliche Betrachtung vollständiger Bindungen, als auch eine Hinzunahme von teilweisen Bindungen plausibel. Da bei Letzteren bereits einzelne Koordinaten ausschlaggebend sein können, ist im Sachzusammenhang zu entscheiden, wieviele Koordinaten mindestens gebunden sein sollen und ob möglicherweise einzelne Koordinaten ignoriert werden können. Ferner kann je nach Datensatz eine Betrachtung von nur vollständigen Bindungen nicht ausreichen, da beispielsweise eine einzelne Koordinate ungebunden sein kann, sodass dort keine vollständigen Bindungen zu finden wären.

3.1.4 Definition (Ränge). Sei $x_1, \dots, x_n \in \mathbb{R}^d$ eine beliebige Stichprobe mit Elementen der Form $x_i = (x_{i1}, \dots, x_{id})$, $i \in \{1, \dots, n\}$. Die Standardhandhabe für Ränge sind die sogenannten maximalen Ränge; für Stichproben mit Bindungen ist es sinnvoll, verschiedene Rang-Begriffe einzuführen:

- Maximaler Rang (Max-Rang): *Der maximale Rang $r_i = (r_{i1}, \dots, r_{id})$ von x_i innerhalb x_1, \dots, x_d ist komponentenweise definiert über*

$$r_{ij} = r_{\max}(x_{ij} | x_{1j}, \dots, x_{nj}) := \sum_{k=1}^n \mathbb{1}_{(x_{kj} \leq x_{ij})}. \quad (3-1)$$

- Minimaler Rang (Min-Rang): *Der minimale Rang $r_i = (r_{i1}, \dots, r_{id})$ von x_i innerhalb x_1, \dots, x_d ist komponentenweise definiert über*

$$r_{ij} = r_{\min}(x_{ij} | x_{1j}, \dots, x_{nj}) := \sum_{k=1}^n \mathbb{1}_{(x_{kj} < x_{ij})} + 1. \quad (3-2)$$

- Mittlerer Rang (Mid-Rang): *Der mittlere Rang $r_i = (r_{i1}, \dots, r_{id})$ von x_i innerhalb x_1, \dots, x_d ist komponentenweise definiert über*

$$r_{ij} = r_{\text{mid}}(x_{ij} | x_{1j}, \dots, x_{nj}) := \sum_{k=1}^n \left(\mathbb{1}_{(x_{kj} < x_{ij})} + \frac{1}{2} \mathbb{1}_{(x_{kj} = x_{ij})} \right) \quad (3-3)$$

$$= \frac{1}{2} (r_{\max}(x_{ij} | x_{1j}, \dots, x_{nj}) + r_{\min}(x_{ij} | x_{1j}, \dots, x_{nj})). \quad (3-4)$$

Unabhängig der Berechnungsart der Ränge r_1, \dots, r_n sind die zugehörigen auf das Intervall $(0, 1)$ -skalierten Ränge $u_1, \dots, u_n \in (0, 1)$ definiert als

$$u_i := \frac{r_i}{n+1} = \left(\frac{r_{i1}}{n+1}, \dots, \frac{r_{id}}{n+1} \right).$$

Die $(0, 1)$ -skalierten Ränge werden Pseudo-Beobachtungen oder Pseudostichprobe genannt.

Bei diesen Berechnungsmethoden ist zu betonen, dass alle gebundenen Einträge mit gleichem Wert auch den gleichen Rang erhalten. Ist gewünscht, dass die Ränge unterschiedlich sein sollen, kann alternativ die Rangvergabe bei Bindungen zufällig durchgeführt werden. Diese Ränge werden mit r_{rnd} bezeichnet. Die zufällige Rangvergabe wird als Randomisierungsmethode in Kapitel 5.2 besprochen und hier nicht nähergehend behandelt.

3.1.5 Bemerkung. (i.) *Bei Stichproben ohne Bindungen stimmen die verschiedenen Rangberechnungsmethoden (3-1), (3-2), (3-3) bzw. (3-4) aus Definition 3.1.4 sowie die zufällige Rangvergabe r_{rnd} in Algorithmus 5.2.1 überein. Allgemein gilt für die Berechnungsmethoden*

$$r_{\min} \leqq r_{\text{mid}}, r_{\text{rnd}} \leqq r_{\max}.$$

(ii.) Da die Rangvergabe bei mehrdimensionalen Stichproben je Koordinate durchgeführt wird, haben die teilweisen Bindungen einen direkten Einfluss darauf. Bei der Nutzung von rangbasierten Methoden und Tests auf gebundenen Datensätzen ist somit die Mitigation von Bindungen jeglicher Art von besonderem Interesse. Die in Bemerkung 3.1.3 genannten Aspekte bleiben ebenso bei der Rangvergabe relevant.

3.1.6 Beispiel (Rangvergabe bei Stichproben mit Bindungen). Für die im Beispiel 3.1.2 betrachtete Stichprobe x_1, \dots, x_{12} werden nachfolgend die unterschiedlichen Rangvergabemethoden angewendet:

1. Koord.	1	8	3	3	6	4	1	3	2	5	3	6
Min-Ränge	1.0	12.0	4.0	4.0	10.0	8.0	1.0	4.0	3.0	9.0	4.0	10.0
Mid-Ränge	1.5	12.0	5.5	5.5	10.5	8.0	1.5	5.5	3.0	9.0	5.5	10.5
Max-Ränge	2.0	12.0	7.0	7.0	11.0	8.0	2.0	7.0	3.0	9.0	7.0	11.0
2. Koord.	6	9	8	9	1	8	5	2	1	6	9	4
Min-Ränge	6.0	10.0	8.0	10.0	1.0	8.0	5.0	3.0	1.0	6.0	10.0	4.0
Mid-Ränge	6.5	11.0	8.5	11.0	1.5	8.5	5.0	3.0	1.5	6.5	11.0	4.0
Max-Ränge	7.0	12.0	9.0	12.0	2.0	9.0	5.0	3.0	2.0	7.0	12.0	4.0

Bei einer Stichprobe kann neben den Eigenschaften *ungebunden* und *gebunden* auch die Konkordanz und Diskordanz definiert werden. Diese Eigenschaften wurden im Abschnitt 2.4 für die Definition des Kendall'schen Tau's genutzt. Zu diesem Koeffizienten wird das empirische Pendant definiert:

3.1.7 Definition (Empirische Kendall'sche Tau's (Tau-A, Tau-B)). Es sei x_1, \dots, x_n eine bivariate Stichprobe mit Elementen der Form $x_i = (x_{i1}, x_{i2})$ gegeben. Für diese Stichprobe werden die Kennzahlen

$$\begin{aligned} n_0 &:= n(n-1)/2 \\ n_1 &:= |\{(i, j) : x_{i1} = x_{j1}\}|, \quad n_c := |\{(i, j) : x_i \text{ und } x_j \text{ sind konkordant}\}|, \\ n_2 &:= |\{(i, j) : x_{i2} = x_{j2}\}|, \quad n_d := |\{(i, j) : x_i \text{ und } x_j \text{ sind diskordant}\}| \end{aligned}$$

berechnet, wobei jeweils $i, j \in \{1, \dots, n\}$ mit $i < j$ betrachtet werden. n_1 und n_2 bezeichnen somit die Anzahlen der teilweise gebundenen Paare (x_i, x_j) . Die Zahlen n_c und n_d bezeichnen die Paare mit Konkordanz und Diskordanz. Das empirische Kendall'sche Tau, auch Kendall-Rangkorrelationskoeffizient oder Tau-A genannt, wird unter der Annahme $n_1 = n_2 = 0$ berechnet als

$$\hat{\tau}_{A,n} := \frac{n_c - n_d}{n_0} \in [-1, +1], \quad (3-5)$$

wobei (3–5) auch bzgl. (2–18) eine Darstellung besitzt der Form

$$\hat{\tau}_{A,n}(x_1, x_2) = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^{i-1} \text{sign}((x_{i1} - x_{j1})(x_{i2} - x_{j2}))$$

(vgl. McNeil et al. (2015): S. 267, Formel (7.52); Agresti (2010): S. 189). Für gebundene Stichproben, d. h. $(n_1, n_2) \geq (0, 0)$, wird ein angepasstes Kendall'sches Tau, das sogenannte Tau-B, definiert als

$$\hat{\tau}_{B,n} := \frac{n_c - n_d}{((n_0 - n_1)(n_0 - n_2))^{1/2}} \in [-1, +1].$$

Für ungebundene Stichproben gilt $\hat{\tau}_{A,n} = \hat{\tau}_{B,n}$ (vgl. Agresti (2010): S. 188).

3.1.8 Bemerkung. In dieser Arbeit wird für $\hat{\tau}_n$ das Tau-B $\hat{\tau}_{B,n}$ genutzt. Im Fall der Bindungsfreiheit gilt $\hat{\tau}_{A,n} = \hat{\tau}_{B,n}$, aber das Tau-A $\hat{\tau}_{A,n}$ benötigt weniger Rechenaufwand.

3.2 Die empirische Copula als rangbasierter nichtparametrischer Copula-Schätzer

3.2.1 Definition (Empirische Verteilungsfunktion und empirische Copula). Sei $F : \mathbb{R}^d \rightarrow [0, 1]$ eine beliebige multivariate Verteilungsfunktion mit zugehöriger Copula $C : [0, 1]^d \rightarrow [0, 1]$. Die betrachtete Stichprobe bestehe aus i.i.d. Zufallsvektoren $X_1, \dots, X_n \sim F$.

Die empirische Verteilungsfunktion $\hat{F}_n : \mathbb{R}^d \rightarrow [0, 1]$ als konsistenter nichtparametrischer Schätzer für F ist bzgl. der Stichprobe definiert als

$$\hat{F}_n(x) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}} = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1}_{\{X_{ij} \leq x_j\}} \quad (x \in \mathbb{R}^d).$$

Auf Basis der Stichprobe X_1, \dots, X_n lässt sich auch ein Schätzer für die Copula definieren: Die empirische Copula $\hat{C}_n : [0, 1]^d \rightarrow [0, 1]$ wird analog zur empirischen Verteilungsfunktion konstruiert als

$$\hat{C}_n(u) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{U_i \leq u\}} = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1}_{\{U_{ij} \leq u_j\}} \quad (u \in [0, 1]^d), \quad (3-6)$$

wobei die Auswertung über die Pseudostichprobe der $(0, 1)$ -skalierten Ränge U_1, \dots, U_n durchgeführt wird (vgl. Hofert et al. (2018): S. 158).

Durch die Konstruktion (3–6) der empirischen Copula über Ränge wird diese Funktion als rangbasierter Schätzer bezeichnet. Für Methoden und Tests, denen \hat{C}_n oder die Ränge U_1, \dots, U_n zugrunde liegen, wird entsprechend von rangbasierten Methoden und Tests gesprochen. Diese werden im Folgeabschnitt 3.3 behandelt.

Für eine empirische Verteilungsfunktion $\hat{F}_n : \mathbb{R} \rightarrow [0, 1]$ zu einer eindimensionalen Stichprobe lässt sich mit Definition 2.1.2 die empirische Quantilfunktion $\hat{F}_n^\leftarrow : [0, 1] \rightarrow \overline{\mathbb{R}}$ als

$$\begin{aligned}\hat{F}_n^\leftarrow(\alpha) &:= \inf\{x \in \mathbb{R} : \hat{F}_n(x) \geqq \alpha\} \\ &= \inf \left\{ x \in \mathbb{R} : \sum_{i=1}^n \mathbf{1}_{\{X_i \leqq x\}} \geqq \alpha n \right\}\end{aligned}$$

für $\alpha \in [0, 1]$ definieren. Diese Funktion stellt einen Schätzer für die Quantilfunktion F^\leftarrow zu F dar.

Eine alternative Definition der empirischen Copula als Widerspiegelung der Konstruktionsaussage im Satz von Sklar 2.2.5 kann über das Plug-In-Prinzip, d. h. dem Ineinandereinsetzen von bereits bekannten Schätzern, formuliert werden.

3.2.2 Definition (Empirische Copula über das Plug-In-Prinzip). *Sei eine Verteilungsfunktion $F : \mathbb{R}^d \rightarrow [0, 1]$ mit Copula $C : [0, 1]^d \rightarrow [0, 1]$ und Marginalen $F_1, \dots, F_d : \mathbb{R} \rightarrow [0, 1]$ gegeben. Es bezeichnen $\hat{F}_n, \hat{F}_{n,1}, \dots, \hat{F}_{n,d}$ die empirischen Schätzer zu F, F_1, \dots, F_d aus Definition 3.2.1.*

Im Satz von Sklar, genauer in Formel (2–5), wurde unter genügender Stetigkeit die Copula $C : [0, 1]^d \rightarrow [0, 1]$ konstruiert als $C = F(F_1^\leftarrow, \dots, F_d^\leftarrow)$. Durch das Ersetzen der Verteilungsfunktionen in ebendieser Formel durch die empirischen Verteilungsfunktionen lässt sich der Schätzer \tilde{C}_n als

$$\tilde{C}_n(u) := \hat{F}_n(\hat{F}_{n,1}^\leftarrow(u_1), \dots, \hat{F}_{n,d}^\leftarrow(u_d)) \quad (u \in [0, 1]^d) \quad (3-7)$$

konstruieren. $\hat{F}_{n,j}^\leftarrow$ bezeichnet jeweils die empirische Quantilfunktion zu $\hat{F}_{n,j}$ für alle $j \in \{1, \dots, d\}$ (vgl. Kojadinovic (2017): S. 4).

Die empirischen Copulas \hat{C}_n und \tilde{C}_n unterscheiden sich dabei in der Größenordnung n^{-1} ; dies wird insbesondere durch die Gleichung

$$(n+1)u_{ij} = r_{ij} = r_{\max}(x_{ij} | x_{1j}, \dots, x_{nj}) = n\hat{F}_{n,j}(x_{ij})$$

zwischen der Pseudostichprobe u_{1j}, \dots, u_{nj} , den Max-Rängen r_{1j}, \dots, r_{nj} und der empirischen Verteilung $\hat{F}_{n,j}$, welche bzgl. der x_{1j}, \dots, x_{nj} konstruiert sind, ersichtlich. Da die Berechnungskomplexität von \hat{C}_n geringer ist, wird die Plug-In-Version in dieser Arbeit nicht genutzt (vgl. Kojadinovic (2017): S. 4).

Konvergenz des empirischen Copula-Prozesses

Vorbereitend für den Abschnitt über die statistischen Tests sind der empirische Copula-Prozess \hat{C}_n und zugehörige Anforderungen zu definieren, da über diesen Prozess grundlegende Konvergenzaussagen für rangbasierte Teststatistiken formuliert werden können. Der empirische Copula-Prozess wird dafür über die Copula C und der empirischen Copula \hat{C}_n als nichtparametrischer Schätzer für ebendiese Copula definiert:

3.2.3 Definition (Empirischer Copula-Prozess). *Sei $C : [0, 1]^d \rightarrow [0, 1]$ eine Copula und $\hat{C}_n : [0, 1]^d \rightarrow [0, 1]$ die empirische Copula als Schätzer für C . Dann heißt*

$$\hat{\mathbb{C}}_n := \sqrt{n}(\hat{C}_n - C) \in \ell^\infty([0, 1]^d)$$

empirischer Copula-Prozess. $\ell^\infty([0, 1]^d)$ bezeichnet der Raum der beschränkten Funktionen auf $[0, 1]^d$ mit Supremumsnorm (vgl. Segers (2012): S. 765).

Bisher waren keine zusätzlichen Anforderungen an die Copula C notwendig. Um Aussagen zur schwachen Konvergenz von $\hat{\mathbb{C}}_n$ durchzuführen, sind nun weitere Anforderungen an die partiellen Ableitungen $\dot{C}_j(u)$, falls sie existieren, zu erfüllen.

3.2.4 Definition (Regularität). *Eine Copula $C : [0, 1]^d \rightarrow [0, 1]$ heißt regulär, wenn die Glattheitseigenschaft*

$$\dot{C}_j(u) \text{ existiert und ist stetig auf } \{u \in [0, 1]^d : u_j \in (0, 1)\} \quad (3-8)$$

für alle $j \in \{1, \dots, d\}$ erfüllt ist (vgl. Segers (2012): S. 767, Condition 2.1).

3.2.5 Satz (Schwache Konvergenz des empirischen Copula-Prozesses). *Ist die Copula $C : [0, 1]^d \rightarrow [0, 1]$ regulär, so gilt für den empirischen Copula-Prozess die schwache Konvergenz*

$$\hat{\mathbb{C}}_n := \sqrt{n}(\hat{C}_n - C) \xrightarrow{d} \mathbb{C}_C \quad (n \rightarrow \infty) \quad (3-9)$$

in $\ell^\infty([0, 1]^d)$ gegen

$$\mathbb{C}_C(u) = \mathbb{B}_C(u) - \sum_{j=1}^d \dot{C}_j(u) \mathbb{B}_C(1, \dots, 1, u_j, 1, \dots, 1) \quad (3-10)$$

für $u \in [0, 1]^d$, wobei \mathbb{B}_C eine C -Brown'sche Brücke ist.

Beweis. Der Beweis ist in Segers (2012) als Beweis zur Proposition 3.1 zu finden (vgl. Segers (2012): S. 768-771). \square

Die Anwendbarkeit dieser Konvergenzaussage für die vorgestellten Copulas wird über die folgende Bemerkung bestätigt.

3.2.6 Bemerkung. *Die Regularitätsanforderungen sind für die Copulas aus den Beispielen 2.3.1, 2.3.2, 2.3.3 und 2.3.4 unter gewissen Einschränkungen an die Parameter erfüllt:*

Besitzt die Korrelationsmatrix P vollen Rang, so erfüllt die Gauß'sche Copula und die t-Copula die Regularitätsanforderung (3–8) (vgl. Segers (2012): S. 776, Example 5.1; Hofert et al. (2018): S. 82, Remark 3.1.2).

Für archimedische Copulas sind die partiellen Ableitungen \dot{C}_j über den jeweiligen Generator ψ darstellbar als

$$\dot{C}_j(u) = \frac{\psi'(u_j)}{\psi'(C(u))} \quad (u \in [0, 1]^d : u_j \in (0, 1)),$$

falls ψ stetig differenzierbar ist. Diese Darstellung gilt für die Familien der Clayton- und Gumbel-Copulas (vgl. Segers (2012): S. 777, Example 5.2).

Die Unabhängigkeitscopula ist regulär. Die Frechét-Hoeffding-Schranken sind nicht regulär (vgl. Segers (2012): S. 779, Example 5.4).

3.3 Rangbasierte Copula-Methoden

In diesem Abschnitt werden insgesamt vier Tests und Methoden für Copulas, die für die Anwendung in der Finanzwirtschaft relevant sind, vorgestellt. Hierfür werden die grundlegenden Definitionen der zu prüfenden Sachverhalte, eine kurze Einordnung der Relevanz sowie theoretische und praxisnahe Tests über geeignete Approximationen behandelt.

Dazu gehören die *Radial-Symmetrie* und *Austauschbarkeit* der Copulas, die *Modellanpassung* unter Annahme der Zugehörigkeit zu einer Copulafamilie, und die *Güte der Anpassung*. Es ist hervorzuheben, dass die Tests und Methoden in diesem Abschnitt bzgl. Bindungen nicht modifiziert sind.

Radial-Symmetrie

3.3.1 Definition (Radial-Symmetrie). Eine Copula $C : [0, 1]^d \rightarrow [0, 1]$ wird radial-symmetrisch genannt, falls sie mit der zugehörigen Survival-Copula $\bar{C} : [0, 1]^d \rightarrow [0, 1]$ übereinstimmt, d. h. falls die Gleichung

$$C(u) = \bar{C}(u) \quad (u \in [0, 1]^d) \tag{3-11}$$

erfüllt ist (vgl. Kojadinovic (2017): S. 9).

Die Radial-Symmetrie-Eigenschaft von Copulas ist hilfreich bei der Modellierung und Interpretation der Abhängigkeitsstruktur zwischen Finanzinstrumenten: Für einen Zufallsvektor $U = (U_1, \dots, U_d) \sim C$ für die Abhängigkeitsstruktur von z. B. Renditen gilt unter (3–11) die Gleichung

$$\mathbb{P}(U \leq u) = \mathbb{P}(U \geq 1 - u)$$

für $u = (u_1, \dots, u_d), 1 - u = (1 - u_1, \dots, 1 - u_d) \in [0, 1]^d$. Dies impliziert, dass in diesem Modell ein gemeinsamer Wertgewinn der Finanzinstrumente gleichwahrscheinlich mit einem gemeinsamen Verlust derselben Größe ist (vgl. Smillie (2008): S. 8). Eine solche Struktur kann über die folgenden Tests auf Radial-Symmetrie bestätigt oder abgelehnt werden:

3.3.2 Definition (Test auf Radial-Symmetrie). *Seien $U_1, \dots, U_n \stackrel{i.i.d.}{\sim} C$ Zufallsvektoren bzgl. einer regulären Copula $C : [0, 1]^d \rightarrow [0, 1]$ oder eine Pseudostichprobe bzgl. einer multivariaten Verteilung $F : \mathbb{R}^d \rightarrow [0, 1]$ mit einer solchen regulären Copula C zugrundeliegend. $\hat{C}_n, \hat{\bar{C}}_n : [0, 1]^d \rightarrow [0, 1]$ bezeichnen die empirischen Versionen der Copula und Survival-Copula.*

Unter der Nullhypothese der Radial-Symmetrie

$$\mathcal{H}_0 : C = \bar{C}$$

ist ein konsistenter statistischer Test auf Radial-Symmetrie über den Abstand zwischen \hat{C}_n und $\hat{\bar{C}}_n$ definierbar: Dazu wird die Teststatistik

$$\begin{aligned} \mathcal{S}_n &:= n \int_{[0,1]^d} (\hat{C}_n(u) - \hat{\bar{C}}_n(u))^2 d\hat{C}_n(u) \\ &= \sum_{i=1}^n (\hat{C}_n(U_i) - \hat{\bar{C}}_n(U_i))^2 \end{aligned} \quad (3-12)$$

betrachtet (vgl. Genest et al. (2012): S. 1108). Für diese Teststatistik gilt unter \mathcal{H}_0 für $n \rightarrow \infty$ die schwache Konvergenz

$$\mathcal{S}_n \xrightarrow[\mathcal{H}_0]{d} \mathcal{S} := \int_{[0,1]^d} (\mathbb{C}_C(u) - \mathbb{C}_C(1-u))^2 dC(u),$$

wobei \mathbb{C}_C der Grenzwert (3–10) des empirischen Copula-Prozesses in (3–9) sei (vgl. Genest und Nešlehová (2013): S. 1111, Proposition 1).

Da die Konvergenz des empirischen Copula-Prozess nach Bemerkung 3.2.6 für die Copula-Familien ohne die Randfälle der Frechét-Hoeffding-Schranken

gilt, wird die Anwendbarkeit und Teststatistik-Konvergenz dieses Tests sichergestellt. Analoge Aussagen gelten ebenso für die weiteren Tests.

Da der Grenzwert \mathcal{S} wegen \mathbb{C}_C von der Copula C abhängt, sind in der Praxis alternative Methoden für die Berechnung der Verteilungen notwendig. Beispielsweise kann die Verteilung und somit der p-Wert zur Statistik \mathcal{S}_n über Bootstrap-Methoden approximiert werden. Der folgende Algorithmus zur Durchführung des Tests auf Radial-Symmetrie mit einer solchen Vorgehensweise ist in Genest und Nešlehová (2013) für den bivariaten Fall formuliert. Ein Ansatz zur multivariaten Verallgemeinerung wird in den an den Algorithmus anschließenden Bemerkungen angegeben.

3.3.3 Algorithmus (Rangbasierter Test auf Radial-Symmetrie mit Approximation nach Genest und Nešlehová (2013)). *Seien $X_1, \dots, X_n \sim F$ i.i.d. Zufallsvektoren mit Komponenten $X_i = (X_{i1}, X_{i2})$ und $F : \mathbb{R}^2 \rightarrow [0, 1]$ eine Verteilungsfunktion. Die Parameter $N \in \mathbb{N}$ und $b_n \in (0, 1/2)$ legen die Iterationsanzahl und die Bandbreite für die Approximation fest.*

Der Test auf Radial-Symmetrie inklusive der Approximation der p-Werte wird mit folgendem Algorithmus durchgeführt:

1. Pseudostichprobe: Berechne für die Stichprobe X_1, \dots, X_n die $(0, 1)$ -skalierten Ränge U_1, \dots, U_n mit Komponenten $U_i = (U_{i1}, U_{i2})$.
2. Berechnung der Teststatistik: Berechne \mathcal{S}_n aus Formel (3–12), d. h.

$$\mathcal{S}_n := \sum_{i=1}^n (\hat{C}_n(U_i) - \hat{\dot{C}}_n(U_i))^2,$$

wobei \hat{C}_n und $\hat{\dot{C}}_n$ bzgl. der U_1, \dots, U_n konstruiert sind.

3. Definition von Hilfsfunktionen: Die Funktion $P_n : [0, 1]^2 \rightarrow \{0, \pm 1\}^n$, $u = (u_1, u_2) \mapsto P(u)$, wird als Spaltenvektor der Form

$$P(u) := (P_1(u), \dots, P_n(u))^\top, \quad P_i(u) := \prod_{j=1}^2 \mathbb{1}_{\{U_{ij} \leq u_j\}} - \prod_{j=1}^2 \mathbb{1}_{\{U_{ij} > 1 - u_j\}}$$

definiert. Die Approximationen $\hat{C}_{1n}, \hat{C}_{2n} : [0, 1]^2 \rightarrow \mathbb{R}$ der partiellen Ableitungen \dot{C}_1 und \dot{C}_2 werden zur Bandbreite b_n konstruiert als

$$\hat{\dot{C}}_{jn}(u) := \frac{\hat{C}_n(u + b_n e_j) - \hat{C}_n(u - b_n e_j)}{2b_n},$$

wobei e_j der j -te kanonische Einheitsvektor ist.

4. Bootstrap-Iterationen: Für jede Iteration $k \in \{1, \dots, N\}$ werden die folgenden Berechnungen durchgeführt:

(i.) Zufallsaspekt je Iteration: Sei $\xi^{(k)} := (\xi_1^{(k)}, \dots, \xi_n^{(k)})$ ein Zufallsvektor mit $\xi_1^{(k)}, \dots, \xi_n^{(k)} \stackrel{i.i.d.}{\sim} \text{Exp}(1)$ und $\xi_j^{(k)} \perp\!\!\!\perp X_1, \dots, X_n$. Der standardisierte Zufallsvektor $\Xi_n^{(k)}$ zu $\xi_n^{(k)}$ ist:

$$\Xi_n^{(k)} := \frac{1}{\sqrt{n}} \left(\frac{\xi_1^{(k)}}{\bar{\xi}_n^{(k)}} - 1, \dots, \frac{\xi_n^{(k)}}{\bar{\xi}_n^{(k)}} - 1 \right), \quad \bar{\xi}_n^{(k)} := \frac{1}{n} \sum_{i=1}^n \xi_i^{(k)}.$$

(ii.) Bootstrap-Version $\hat{\mathbb{D}}_n^{(k)}(u)$ für $\mathbb{D}(u) := \mathbb{C}_C(u) - \mathbb{C}_C(1-u)$: Die Funktion $\hat{\mathbb{D}}_n^{(k)}(u) : [0, 1]^2 \rightarrow \mathbb{R}$ wird über das Skalarprodukt

$$\hat{\mathbb{D}}_n^{(k)}(u) := (\Xi_n^{(k)}) (P(u) - \hat{C}_{1n}(u)P(u_1, 1) - \hat{C}_{2n}(u)P(1, u_2)) \quad (3-13)$$

definiert.

(iii.) Bootstrap-Version $\mathcal{S}_n^{(k)}$ der Teststatistik \mathcal{S}_n : Die Teststatistik $\mathcal{S}_n^{(k)}$ wird über $\hat{\mathbb{D}}_n^{(k)}$ und \hat{C}_n berechnet als

$$\mathcal{S}_n^{(k)} := \int_{[0,1]^2} (\hat{\mathbb{D}}_n^{(k)}(u))^2 d\hat{C}_n(u) = \frac{1}{n} \sum_{i=1}^n (\hat{\mathbb{D}}_n^{(k)}(U_i))^2.$$

5. Berechnung des p-Wertes: Der für die Teststatistik \mathcal{S}_n approximierte p-Wert $p(\mathcal{S}_n) \in (0, 1)$ lautet

$$p(\mathcal{S}_n) = \frac{1}{N+1} \sum_{k=1}^N (\mathbb{1}_{\{\mathcal{S}_n^{(k)} > \mathcal{S}_n\}} + 0.5).$$

Die R-Implementierung ist im Appendix D.2.1 zu finden.

3.3.4 Bemerkung. Zum Algorithmus 3.3.3 ist Folgendes anzumerken:

- (i.) Wahl der Bandbreite: In Genest und Nešlehová (2013) und in der Implementierung von Quessy (2021) wird $b_n = b/\sqrt{n}$ für ein $b \in \mathbb{N}$ als Bandbreite gewählt.
- (ii.) Alternative Berechnung des p-Wertes: In Anlehnung an die Algorithmen in Kojadinovic (2017) wird die Berechnung des p-Wertes über die Multiplikation mit $N/(N+1)$ derart angepasst, dass die Werte in $(0, 1)$ statt in $[0, 1]$ liegen (vgl. Kojadinovic (2017): S. 7).

- (iii.) Verteilungswahl: In Schritt 4.(i.) können bei der Verteilung der $\xi_j^{(k)}$ statt der Standard-Exponentialverteilung $\text{Exp}(1)$ auch andere nicht-negative Verteilungen mit Erwartungswert und Varianz gleich Eins genutzt werden (vgl. Genest und Nešlehová (2013): S. 1113).
- (iv.) Code-Grundlage: Die R-Implementierung des Radial-Symmetrie-Tests basiert auf der MATLAB-Implementierung von Quessy (2021). Über geeignete Matrix-Notationen wurden dort die Berechnungen vereinfacht.
- (v.) Multivariate Verallgemeinerungen des Tests: Der bivariate Test kann auf d -viele Dimensionen verallgemeinert werden. Dafür sind die betrachteten Zufallsvektoren, Ränge und Funktionen d -dimensional aufzufassen. Die Berechnung von $\hat{\mathbb{D}}_n^{(k)}$ ist entsprechend der Darstellung (3–10) anzupassen (vgl. Segers (2012): S. 768–772). Quessy (2021) stellt den zugehörigen multivariaten Test auf Radial-Symmetrie als MATLAB-Implementierung zur Verfügung.

3.3.5 Beispiel (Radial-Symmetrie). Die Nullhypothese der Radial-Symmetrie (3–11), folgend mit \mathcal{H}_0 bezeichnet, ist für die Familien der Gauß'schen Copulas und t -Copulas erfüllt. Diese Familien werden in diesem Zusammenhang auch als $\mathcal{H}_{0,RS}$ -Copulas bezeichnet. Die Alternativhypothese $\mathcal{H}_1 : C \neq \bar{C}$ gilt für die Familien der Clayton- und Gumbel-Copulas. Für diese Familien gilt die Bezeichnung als $\mathcal{H}_{1,RS}$ -Copulas (vgl. McNeil et al. (2015): S. 233).

Austauschbarkeit

Eine weitere Eigenschaft, die über rangbasierte Methoden getestet werden kann, ist die sogenannte Austauschbarkeit der Copula:

3.3.6 Definition (Austauschbarkeit). Eine Copula $C : [0, 1]^d \rightarrow [0, 1]$ mit $u \mapsto C(u)$ heißt austauschbar, falls die Copula unter beliebiger Vertauschung der Koordinaten von u erhalten bleibt, d. h. falls die Gleichung

$$C(u_1, \dots, u_d) = C(u_{\pi(1)}, \dots, u_{\pi(d)}) \quad (u \in [0, 1]^d) \quad (3-14)$$

für jede Permutation π auf $\{1, \dots, d\}$ gilt (vgl. Kojadinovic (2017): S. 3).

Austauschbare Copulas werden in der Praxis bei der Modellierung von Abhängigkeiten innerhalb homogener Gruppen von Finanzinstrumenten, Unternehmen, Risikotreibern etc. genutzt (vgl. McNeil et al. (2015): S. 234). Im Fall der Austauschbarkeit sind verallgemeinernde Modellierungen möglich, in denen die betrachteten Komponenten aus der homogenen Gruppe frei wählbar und untereinander austauschbar sind.

Wie beim Test auf Radial-Symmetrie werden eine Teststatistik, deren Konvergenz und ein anwendungsorientierter Test zur Austauschbarkeit definiert. Im Gegensatz zur Radial-Symmetrie, bei der die empirische Copula und Survival-Copula benötigt werden, wären bei einem multivariaten Test auf Austauschbarkeit $d!$ -viele Permutationen zu berücksichtigen. Entsprechend werden die Tests wie in Genest et al. (2012) im bivariaten Fall definiert. Eine aufwändige multivariate Verallgemeinerung ist in Harder und Stadtmüller (2016) im Detail einsehbar.

3.3.7 Definition (Test auf Austauschbarkeit). *Seien $U_1, \dots, U_n \stackrel{i.i.d.}{\sim} C$ Zufallsvektoren bzgl. einer regulären bivariaten Copula $C : [0, 1]^2 \rightarrow [0, 1]$ oder eine Pseudostichprobe bzgl. einer multivariaten Verteilung $F : \mathbb{R}^d \rightarrow [0, 1]$ mit einer solchen regulären Copula C zugrundeliegend. $\hat{C}_n : [0, 1]^2 \rightarrow [0, 1]$ bezeichnet die empirische Copula.*

Unter der Nullhypothese der Austauschbarkeit

$$\mathcal{H}_0 : \forall (u_1, u_2) \in [0, 1]^2 : C(u_1, u_2) = C(u_2, u_1)$$

ist ein konsistenter statistischer Test auf Radial-Symmetrie definierbar. Die dazu betrachtete Teststatistik zur Austauschbarkeit lautet

$$\begin{aligned} \mathcal{A}_n &:= n \int_{[0,1]^d} (\hat{C}_n(u_1, u_2) - \hat{C}_n(u_2, u_1))^2 d\hat{C}_n(u_1, u_2) \\ &= \sum_{i=1}^n (\hat{C}_n(U_{i1}, U_{i2}) - \hat{C}_n(U_{i2}, U_{i1}))^2 \end{aligned} \quad (3-15)$$

(vgl. Genest et al. (2012): S. 814.). Für diese Teststatistik gelten für $n \rightarrow \infty$ die schwache Konvergenz unter \mathcal{H}_0 und die stochastische Konvergenz

$$\begin{aligned} \mathcal{A}_n &\xrightarrow[\mathcal{H}_0]{d} \mathcal{A} := \int_{[0,1]^d} (\mathbb{C}_C(u_1, u_2) - \mathbb{C}_C(u_2, u_1))^2 dC(u_1, u_2), \\ \mathcal{A}_n &\xrightarrow{\mathbb{P}} \tilde{\mathcal{A}} := n \int_{[0,1]^d} (C(u_1, u_2) - C(u_2, u_1))^2 dC(u_1, u_2), \end{aligned}$$

wobei \mathbb{C}_C der Grenzwert (3-10) des empirischen Copula-Prozesses in (3-9) sei (vgl. Genest et al. (2012): S. 814, 816-817, 830, Propositionen 2 bis 4).

3.3.8 Algorithmus (Rangbasierter Test auf Austauschbarkeit mit Approximation nach Genest et al. (2012)). *Seien $X_1, \dots, X_n \sim F$ i.i.d. Zufallsvektoren mit Komponenten $X_i = (X_{i1}, X_{i2})$ und $F : \mathbb{R}^2 \rightarrow [0, 1]$ eine Verteilungsfunktion. Die Parameter $N \in \mathbb{N}$ und $b_n \in (0, 1/2)$ legen die Iterationsanzahl und die Bandbreite für die Approximation fest.*

Der Test auf Austauschbarkeit inklusive der Approximation der p -Werte wird mit folgendem Algorithmus durchgeführt:

1. Pseudostichprobe: Berechne für die Stichprobe X_1, \dots, X_n die $(0, 1)$ -skalierten Ränge U_1, \dots, U_n mit Komponenten $U_i = (U_{i1}, U_{i2})$.
2. Berechnung der Teststatistik: Berechne \mathcal{A}_n aus Formel (3–15), d. h.

$$\mathcal{A}_n := \sum_{i=1}^n (\hat{C}_n(U_{i1}, U_{i2}) - \hat{C}_n(U_{i2}, U_{i1}))^2,$$

wobei die empirische Copula \hat{C}_n bzgl. U_1, \dots, U_n konstruiert ist.

3. Definition von Hilfsfunktionen: Die Funktion $P_n : [0, 1]^2 \rightarrow \{0, \pm 1\}^n$, $u = (u_1, u_2) \mapsto P(u)$, wird als Spaltenvektor der Form

$$P(u) := (P_1(u), \dots, P_n(u))^\top, \quad P_i(u) := \prod_{j=1}^2 \mathbb{1}_{\{U_{ij} \leq u_j\}} - \prod_{j=1}^2 \mathbb{1}_{\{U_{ij} > 1 - u_j\}}$$

definiert. Die Approximationen $\hat{\dot{C}}_{1n}, \hat{\dot{C}}_{2n} : [0, 1]^2 \rightarrow \mathbb{R}$ der partiellen Ableitungen \dot{C}_1 und \dot{C}_2 werden zur Bandbreite b_n konstruiert als

$$\hat{\dot{C}}_{1n}(u) := \begin{cases} \frac{\hat{C}_n(2b_n, u_2)}{2b_n}, & \text{für } u_1 \in [0, b_n], \\ \frac{\hat{C}_n(u_1 + b_n, u_2) - \hat{C}_n(u_1 - b_n, u_2)}{2b_n}, & \text{für } u_1 \in [b_n, 1 - b_n], \\ \frac{\hat{C}_n(1, u_2) - \hat{C}_n(1 - 2b_n, u_2)}{2b_n}, & \text{für } u_1 \in (1 - b_n, 1], \end{cases}$$

und analog

$$\hat{\dot{C}}_{2n}(u) := \begin{cases} \frac{\hat{C}_n(u_1, 2b_n)}{2b_n}, & \text{für } u_2 \in [0, b_n], \\ \frac{\hat{C}_n(u_1, u_2 + b_n) - \hat{C}_n(u_1, u_2 - b_n)}{2b_n}, & \text{für } u_2 \in [b_n, 1 - b_n], \\ \frac{\hat{C}_n(u_1, 1) - \hat{C}_n(u_1, 1 - 2b_n)}{2b_n}, & \text{für } u_2 \in (1 - b_n, 1]. \end{cases}$$

4. Bootstrap-Iterationen: Für jede Iteration $k \in \{1, \dots, N\}$ werden die folgenden Berechnungen durchgeführt:

- (i.) Zufallsaspekt je Iteration: Sei $\xi^{(k)} := (\xi_1^{(k)}, \dots, \xi_n^{(k)})$ ein Zufallsvektor mit $\xi_1^{(k)}, \dots, \xi_n^{(k)} \stackrel{i.i.d.}{\sim} \text{Exp}(1)$ und $\xi^{(j)} \perp\!\!\!\perp X_1, \dots, X_n$. Der standardisierte Zufallsvektor $\Xi_n^{(k)}$ zu $\xi_n^{(k)}$ ist:

$$\Xi_n^{(k)} := \frac{1}{\sqrt{n}} \left(\frac{\xi_1^{(k)}}{\bar{\xi}_n^{(k)}} - 1, \dots, \frac{\xi_n^{(k)}}{\bar{\xi}_n^{(k)}} - 1 \right), \quad \bar{\xi}_n^{(k)} := \frac{1}{n} \sum_{i=1}^n \xi_i^{(k)}.$$

- (ii.) Bootstrap-Version $\hat{\mathbb{D}}_n^{(k)}(u)$ für $\mathbb{D}(u) := \mathbb{C}_C(u_1, u_2) - \mathbb{C}_C(u_2, u_1)$: Die Funktion $\hat{\mathbb{D}}_n^{(k)}(u) : [0, 1]^d \rightarrow \mathbb{R}$ wird über das Skalarprodukt

$$\hat{\mathbb{D}}_n^{(k)}(u) := (\Xi_n^{(k)}) (P(u) - \hat{C}_{1n}(u)P(u_1, 1) - \hat{C}_{2n}(u)P(1, u_2)) \quad (3-16)$$

definiert.

- (iii.) Bootstrap-Version $\mathcal{A}_n^{(k)}$ der Teststatistik \mathcal{A}_n : Die Teststatistik $\mathcal{A}_n^{(k)}$ wird über $\hat{\mathbb{D}}_n^{(k)}$ und \hat{C}_n berechnet als

$$\mathcal{A}_n^{(k)} := \int_{[0,1]^d} (\hat{\mathbb{D}}_n^{(k)}(u))^2 d\hat{C}_n(u) = \frac{1}{n} \sum_{i=1}^n (\hat{\mathbb{D}}_n^{(k)}(U_i))^2.$$

5. Berechnung des p-Wertes: Der für die Teststatistik \mathcal{A}_n approximierte p-Wert $p(\mathcal{A}_n) \in (0, 1)$ lautet

$$p(\mathcal{A}_n) = \frac{1}{N+1} \sum_{k=1}^N (1_{\{\mathcal{A}_n^{(k)} > \mathcal{A}_n\}} + 0.5).$$

Die R-Implementierung ist im Appendix D.2.2 bereitgestellt.

3.3.9 Bemerkung. Zum Algorithmus 3.3.8 ist Folgendes anzumerken:

- (i.) Alternative Berechnung des p-Wertes: Wie im Algorithmus 3.3.8 wird die Berechnung des p-Wertes angepasst (vgl. Kojadinovic (2017): S. 7).
- (ii.) Code-Grundlage: Quessy (2021) stellt keine Implementierung bereit. Die R-Implementierung des Tests auf Austauschbarkeit wurde auf Basis der Implementierung für den Test auf Radial-Symmetrie erstellt.
- (iii.) Partielle Ableitungen: Die Berechnung der Approximationen $\hat{C}_{jn}(u)$ unterscheidet sich zwischen den Algorithmen 3.3.3 und 3.3.8 in der Konstruktion, da die Approximationen in dieser Form in den jeweiligen Veröffentlichungen formuliert wurden. Beide Definitionen sind in der Verwendung in den Algorithmen zulässig (vgl. Bücher und Kojadinovic (2016): S. 19-20).

3.3.10 Beispiel (Austauschbarkeit). *Die vier Familien der Gauß'schen, t-, Clayton- und Gumbel-Copulas erfüllen alle die Nullhypothese der Austauschbarkeit (3–14), und werden in diesem Zusammenhang $\mathcal{H}_{0,EX}$ -Copulas genannt (vgl. McNeil et al. (2015): S. 235). Die Khoudraji-Copulafamilien bzgl. der vorgestellten Familien erfüllen die Hypothese der Austauschbarkeit nicht und werden als $\mathcal{H}_{1,EX}$ -Copulas bezeichnet (Kojadinovic (2017): S. 3).*

In Kojadinovic (2017) werden für das Khoudraji's Device $kho_s(C_\vartheta, \Pi)(u)$, $u \in [0, 1]^2$, die Parameter $s = (s_1, s_2)$ mit $s_1 \in \{0.2, 0.4, 0.6, 0.8\}$ und $s_2 = 0.95$ gewählt (vgl. Kojadinovic (2017): S. 9, Table 2).

Modellanpassung und Anpassungsgüte

Um die Güte einer Anpassung $C_{\hat{\vartheta}_n}$ zur Copula C zu bestimmen bzw. zu testen, ist eine Annahme an die Zugehörigkeit der Copula C zu einer Copulafamilie \mathcal{C}_0 zu tätigen. Diese Zugehörigkeit wird dann als Nullhypothese

$$\mathcal{H}_0 : C \in \mathcal{C}_0 = \{C_\vartheta : \vartheta \in \Theta\}$$

formuliert. Um eine Einschränkung der Hypothese auf

$$\mathcal{H}'_0 : F \in \mathcal{F}_0 = \{F = C(F_1, \dots, F_d) : C \in \mathcal{C}_0, F_j \in \mathcal{F}_j\}$$

wegen einer umfassenden Modellierung einer Verteilungsfunktion F inklusive Copula C und Marginale F_1, \dots, F_d zu vermeiden, wird ausgenutzt, dass die Copula C unter isoton-stetigen Transformationen erhalten bleibt und dass die Invarianz auch für Rangstatistiken gilt. Somit werden bzgl. der ursprünglichen Hypothese \mathcal{H}_0 rangbasierte Schätzmethoden und Tests über die Anpassungsgüte definiert.

Sei angenommen, dass für die Copula $C : [0, 1]^d \rightarrow [0, 1]$ die Hypothese \mathcal{H}_0 mit $C = C_{\vartheta_0} \in \mathcal{C}_0$ für den (unbekannten) Parameter $\vartheta_0 \in \Theta$ erfüllt ist. Ein Schätzer $\hat{\vartheta}_n$ für den Parameter $\vartheta \in \Theta$ kann über verschiedene Ansätze berechnet werden: Zwei Verfahren sind die Maximum-Pseudo-Likelihood-Schätzung (MPLE) über die Dichtefunktion zur Copulafamilie \mathcal{C}_0 und die Momenten-Methode (MME), die das inverse Kendall'sche Tau zur Parameterschätzung nutzt. In Genest et al. (2009) wurden für die Modellanpassung und den Test der Anpassungsgüte MPLE- und ITS-Methoden verwendet. Die ITS-Methoden werden in dieser Arbeit nicht genutzt. Die Resultate dieser Studie nennen mit Hinblick auf das Testen der Anpassungsgüte eine Präferenz von MPLE-Methoden (vgl. Genest et al. (2009): S. 209–211). Kojadinovic (2017) nutzt für die Parameterschätzung sowohl MPLE als auch MME; für den bzgl. Bindungen modifizierten Test der Anpassungsgüte ist der MPL-Schätzer zu bevorzugen, da dessen Verhalten am konsistentesten

war (vgl. Kojadinovic (2017): S. 21). Auch Li et al. (2020) verwendet für die Modellanpassung und den Test der Anpassungsgüte einen MPLE-Ansatz mit Intervall-Zensur. Den Resultaten und der Verwendung in den genannten Studien entsprechend, wird in dieser Arbeit die MPLE als zugrundeliegende Schätzmethode betrachtet. Die folgende Definition konstruiert den rangbasierten MPL-Schätzer $\hat{\vartheta}_n$ für den Copula-Parameter $\vartheta_0 \in \Theta$:

3.3.11 Definition (Rangbasierte Copula-Modellanpassung über MPLE).
Sei $C := C_{\vartheta_0} : [0, 1]^d \rightarrow [0, 1]$ eine Copula aus der Familie $\mathcal{C}_0 := \{C_\vartheta : \vartheta \in \Theta\}$ mit unbekanntem Parameter $\vartheta_0 \in \Theta$, die die Abhängigkeitsstruktur einer Verteilungsfunktion $F : \mathbb{R}^d \rightarrow [0, 1]$, $F = C(F_1, \dots, F_d)$, modelliert. Es bezeichnet $c_\vartheta(u) := c(u; \vartheta)$ die Dichtefunktion zur Copula $C_\vartheta \in \mathcal{C}_0$. Weiter sei für eine i.i.d. Stichprobe $X_1, \dots, X_n \sim F$ die Pseudostichprobe U_1, \dots, U_n mit $U_i = (U_{i1}, \dots, U_{id})$ gegeben.

Dann ist die Log-Pseudo-Likelihood-Funktion $\mathcal{L} : \Theta \rightarrow \mathbb{R}$ definiert als

$$\mathcal{L}(\vartheta; u_1, \dots, u_n) := \sum_{i=1}^n \ln c(u_i; \vartheta).$$

Der Maximum-Pseudo-Likelihood-Schätzer $\hat{\vartheta}_n$ für den unbekannten Parameter $\vartheta_0 \in \Theta$ wird über die Maximierung von \mathcal{L} auf der Pseudostichprobe bestimmt:

$$\hat{\vartheta}_n = \hat{\vartheta}_n(U_1, \dots, U_n) = \arg \max_{\vartheta \in \Theta} \mathcal{L}(\vartheta; U_1, \dots, U_n)$$

(vgl. McNeil et al. (2015): S. 270, Formel (7.55)).

Zur technischen Umsetzung der Modellanpassung in R wird die Funktion `fitCopula()` aus dem R-Paket Copula von Hofert et al. (2020) genutzt. Da neben dem Parameter ϑ_0 auch die zugrundeliegende Copulafamilie \mathcal{C}_0 unbekannt sein kann, ist bei einer Modellanpassung zu prüfen, ob diese für die zugrundeliegenden Daten plausibel ist. Für die Definition eines Tests der Anpassungsgüte ist ein weiterer empirischer Prozess notwendig.

3.3.12 Definition (Empirischer Prozess für den Test der Anpassungsgüte).
Sei $C : [0, 1]^d \rightarrow [0, 1]$, $C = C_{\vartheta_0} \in \mathcal{C}_0$, eine Copula der Familie $\mathcal{C}_0 := \{C_\vartheta : \vartheta \in \Theta\}$. Für eine i.i.d. Pseudostichprobe U_1, \dots, U_n bzgl. der Copula C sind die empirische Copula \hat{C}_n als nichtparametrischer Schätzer und der MPL-Schätzer $C_{\hat{\vartheta}_n} \in \mathcal{C}_0$ mit $\hat{\vartheta}_n = \hat{\vartheta}_n(U_1, \dots, U_n)$ für ϑ_0 betrachtet.

Der empirische Prozess, der dem Test der Anpassungsgüte zur Nullhypothese $\mathcal{H}_0 : C \in \mathcal{C}_0 = \{C_\vartheta : \vartheta \in \Theta\}$ zugrunde liegt, lautet

$$\hat{\mathbb{C}}_n^{\text{GoF}} := \sqrt{n} (\hat{C}_n - C_{\hat{\vartheta}_n}) \in \ell^\infty([0, 1]^d)$$

(vgl. Genest et al. (2009): S. 200-201; Kojadinovic (2017): S. 18).

3.3.13 Definition (Test der Anpassungsgüte). Seien $U_1, \dots, U_n \stackrel{i.i.d.}{\sim} C$ Zufallsvektoren bzgl. einer regulären bivariaten Copula $C : [0, 1]^2 \rightarrow [0, 1]$ oder eine Pseudostichprobe bzgl. einer multivariaten Verteilung $F : \mathbb{R}^d \rightarrow [0, 1]$ mit einer solchen regulären Copula C zugrundeliegend. \hat{C}_n ist die dazugehörige empirische Copula. Sei weiter $\mathcal{C}_0 = \{C_\vartheta : \vartheta \in \Theta\}$ eine Copulafamilie, über welche die Nullhypothese

$$\mathcal{H}_0 : C \in \mathcal{C}_0$$

formuliert wird. Aus dieser Familie soll ferner der parametrische Schätzer $C_{\hat{\vartheta}_n} \in \mathcal{C}_0$ mit $\hat{\vartheta}_n$ über MPLE bestimmt werden. Unter \mathcal{H}_0 sei $C = C_{\vartheta_0}$ für den (unbekannten) Parameter $\vartheta_0 \in \Theta$.

Hierzu ist ein konsistenter statistischer Test der Anpassungsgüte definierbar. Die betrachtete Teststatistik lautet

$$\begin{aligned} \mathcal{G}_n &:= n \int_{[0,1]^d} (\hat{C}_n^{\text{GoF}}(u))^2 d\hat{C}_n(u) \\ &= \sum_{i=1}^n (\hat{C}_n(U_i) - C_{\hat{\vartheta}_n}(U_i))^2 \end{aligned} \quad (3-17)$$

(vgl. Genest et al. (2009): S. 201; Kojadinovic (2017): S. 19). Für diese Teststatistik werden unter \mathcal{H}_0 und $n \rightarrow \infty$ die schwachen Konvergenzen

$$\hat{\theta}_n := \sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \xrightarrow{d} \theta \sim \mathcal{N}(0, \cdot)$$

und

$$\begin{aligned} \hat{C}_n^{\text{GoF}} &= \sqrt{n}(\hat{C}_n - C_{\vartheta_0}) - \sqrt{n}(C_{\hat{\vartheta}_n} - C_{\vartheta_0}) \\ &= \hat{C}_n - (\partial_\vartheta C_{\vartheta_0})^\top \hat{\theta}_n \xrightarrow{d} \mathbb{C}_C - (\partial_\vartheta C_{\vartheta_0})^\top \hat{\theta}_n \end{aligned}$$

genutzt, um die Konvergenz der Teststatistik sicherzustellen:

$$\mathcal{G}_n \xrightarrow[\mathcal{H}_0]{d} \mathcal{G} := \int_{[0,1]^d} (\mathbb{C}_C(u) - (\partial_\vartheta C_{\vartheta_0})^\top \hat{\theta}_n)^2 dC(u).$$

\mathbb{C}_C sei hier der Grenzwert (3-10) des empirischen Copula-Prozesses in (3-9) (vgl. Genest und Rémillard (2008): S. 1107-1108; Rémillard (2017): S. 8, 21: Proposition 1 und zugehöriger Beweis).

3.3.14 Bemerkung. Neben dem MPLE $\hat{\vartheta}_n$ sind auch andere Schätzer nutzbar, sodass die Konvergenz von \mathcal{G}_n erhalten bleibt. Diese ist für MPLE und MME erfüllt (vgl. Genest und Rémillard (2008): S. 1108-1110). Kojadinovic (2017) sagt aus, dass die MPLE unter mittleren Rängen, die Statistik \mathcal{G}_n unter maximalen Rängen berechnet werden solle, um bestmögliche Resultate zu erhalten (vgl. Kojadinovic (2017): S. 19).

Für diesen Test ist der Grenzwert \mathcal{G} abhängig von der Copula C definiert. Wie bei den vorigen beiden Tests auf Radial-Symmetrie und Austauschbarkeit ist eine Approximation der Verteilung zur Anwendung notwendig.

3.3.15 Algorithmus (Rangbasierter Test der Anpassungsgüte mit Approximation nach Genest et al. (2009)). *Seien $X_1, \dots, X_n \sim F$ i.i.d. Zufallsvektoren mit Komponenten $X_i = (X_{i1}, \dots, X_{id})$ und $F : \mathbb{R}^d \rightarrow [0, 1]$ eine Verteilungsfunktion. Sei $C = C_{\vartheta_0} \in \mathcal{C}_0$ die zu F gehörige Copula aus der Familie $\mathcal{C}_0 = \{C_\vartheta : \vartheta \in \Theta\}$.*

Der Parameter $N \in \mathbb{N}$ legt die Anzahl der Bootstrap-Iterationen fest, $M \in \mathbb{N}$, $M \geq n$, definiert – falls notwendig – die Iterationsanzahl der Monte-Carlo-Approximation $\hat{C}_{M, \hat{\vartheta}_n}^*$ für $C_{\hat{\vartheta}_n}$.

Der Test der Anpassungsgüte inklusive der Approximation der p -Werte wird mit folgendem Algorithmus durchgeführt:

1. Pseudostichproben: Berechne für die Stichprobe X_1, \dots, X_n die $(0, 1)$ -skalierten Ränge U_1, \dots, U_n mit Komponenten $U_i = (U_{i1}, \dots, U_{id})$.
2. Rangbasierte Copula-Schätzer: Schätze den Parameter $\vartheta_0 \in \Theta$ über den MPL-Schätzer $\hat{\vartheta}_n := \hat{\vartheta}_n(U_1, \dots, U_n)$. Konstruiere $C_{\hat{\vartheta}_n}$ und die empirische Copula \hat{C}_n als rangbasierte Schätzer bzgl. U_1, \dots, U_n .
3. Berechnung der Teststatistik: Besitzt die Copula $C_{\hat{\vartheta}_n}$ eine explizite Darstellung, so kann die Teststatistik \mathcal{G}_n über die Formel (3–17) direkt berechnet werden, d. h.

$$\mathcal{G}_n := \sum_{i=1}^n (\hat{C}_n(U_i) - C_{\hat{\vartheta}_n}(U_i))^2.$$

Ist keine solche Darstellung vorliegend, so kann $C_{\hat{\vartheta}_n}$ über einen Monte-Carlo-Ansatz approximiert werden:

(i.) Monte-Carlo-Stichprobe: Generiere für $C_{\hat{\vartheta}_n}$ die Stichprobe

$$V_1, \dots, V_M \stackrel{i.i.d.}{\sim} C_{\hat{\vartheta}_n}.$$

(ii.) Approximation der Funktion $C_{\hat{\vartheta}_n}$: Konstruiere bzgl. V_1, \dots, V_M die Funktion $\hat{C}_{M, \hat{\vartheta}_n}^* : [0, 1]^d \rightarrow [0, 1]$ als empirische Copula, d. h.

$$\hat{C}_{M, \hat{\vartheta}_n}^*(v) := \frac{1}{M} \sum_{k=1}^M \mathbf{1}_{\{V_k \leq v\}} \quad (v \in [0, 1]^d).$$

- (iii.) Berechnung der Teststatistik: Berechne die Teststatistik aus der Formel (3–17) bzgl. der Approximation $\hat{C}_{M,\hat{\vartheta}_n}^*$, d. h.

$$\mathcal{G}_n := \sum_{i=1}^n (\hat{C}_n(U_i) - \hat{C}_{M,\hat{\vartheta}_n}^*(U_i))^2.$$

4. Bootstrap-Iterationen: Für jede Iteration $k \in \{1, \dots, N\}$ werden die folgenden Berechnungen durchgeführt:

- (i.) Zufallsaspekt je Iteration: Generiere bzgl. der Copula $C_{\hat{\vartheta}_n}$ die Stichprobe $Y_1^{(k)}, \dots, Y_n^{(k)}$ $i.i.d.$ $\sim C_{\hat{\vartheta}_n}$. Berechne dazugehörig die Pseudostichprobe $U_1^{(k)}, \dots, U_n^{(k)}$ der $(0, 1)$ -skalierten Ränge.
- (ii.) Rangbasierte Copula-Schätzer: Schätze den Parameter $\vartheta_0 \in \Theta$ über den MPL-Schätzer $\hat{\vartheta}_n^{(k)} := \hat{\vartheta}_n^{(k)}(U_1^{(k)}, \dots, U_n^{(k)})$. Konstruiere $C_{\hat{\vartheta}_n^{(k)}}$ und die empirische Copula $\hat{C}_n^{(k)}$ bzgl. $U_1^{(k)}, \dots, U_n^{(k)}$.
- (iii.) Berechnung der Teststatistik: Besitzt die Copula $C_{\hat{\vartheta}_n^{(k)}}$ eine explizite Darstellung, so kann die Bootstrap-Teststatistik $\mathcal{G}_n^{(k)}$ über die Formel (3–17) direkt berechnet werden, d. h.

$$\mathcal{G}_n^{(k)} := \sum_{i=1}^n (\hat{C}_n^{(k)}(U_i^{(k)}) - C_{\hat{\vartheta}_n^{(k)}}(U_i^{(k)}))^2.$$

Ist keine solche Darstellung vorliegend, so kann $C_{\hat{\vartheta}_n^{(k)}}$ analog zu der Approximation 3.(i.) bis 3.(iii.) bestimmt werden.

5. Berechnung des p-Wertes: Der für die Teststatistik \mathcal{G}_n approximierte p-Wert $p(\mathcal{G}_n) \in (0, 1)$ lautet

$$p(\mathcal{G}_n) = \frac{1}{N+1} \sum_{k=1}^N (\mathbb{1}_{\{\mathcal{G}_n^{(k)} > \mathcal{G}_n\}} + 0.5).$$

Im R-Paket Copula von Hofert et al. (2020) können bei der Funktion `gofCopula()`, die den Test der Anpassungsgüte aus Kojadinovic (2017) durchführt, die Bindungshandhabe mit `ties=FALSE` deaktiviert werden. Dies entspricht dem Test der Anpassungsgüte aus Genest et al. (2009) (vgl. Kojadinovic (2017): S. 20).

Kapitel 4

Empirische und synthetische Datensätze

Dieses Kapitel betrachtet die Rahmenbedingungen auf Finanzmärkten mit dem Ziel der Entwicklung eines Generators für synthetische Datensätze zur Anwendung in Simulationsstudien.

Im Abschnitt 4.1 werden regulatorische Anforderungen sowie die Rahmenbedingungen der Börsen und Datenanbieter, die zu Bindungen führen können, präsentiert. Im zweiten Abschnitt 4.2 wird der datengenerierende Algorithmus entwickelt. Dafür wird ein existierender Algorithmus nach Kojadinovic (2017) als Grundlage herangezogen und entsprechend der Informationen aus den vorigen Abschnitten ergänzt. In den beiden Abschnitten 4.3 und 4.4 werden eine Auswahl empirischer Datensätze bzgl. Bindungen, Datenpräzision und fehlerhafter Einträge analysiert; Ziel ist die Modellierung der Rahmenbedingungen für die Simulationsstudie.

4.1 Rahmenbedingungen auf Finanzmärkten

Innerhalb empirischer Finanzmarktdaten können im Allgemeinen – und trotz Aufbereitung und Verarbeitung seitens der bereitstellenden Unternehmen – Bindungen und Datenfehler vorkommen. Da synthetische Datensätze von Algorithmen unter gewissen Annahmen und unter Idealbedingungen generiert werden, sind zur Wahrung der Praxisrelevanz der Resultate zusätzlich die Praxisbedingungen in den Algorithmen zu berücksichtigen.

Hierzu werden die Rahmenbedingungen für Finanzmärkte aufbereitet und bzgl. Relevanz und Plausibilität validiert. Eine begründete Auswahl der Einschränkungen und Annahmen wird in den datengenerierenden Algorithmen eingebunden.

Preisband	LB ₁	LB ₂	LB ₃	LB ₄	LB ₅	LB ₆
$0.00 \leq P < 0.10$	$5 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	$1 \cdot 10^{-4}$
$0.10 \leq P < 0.20$	$1 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	$1 \cdot 10^{-4}$
$0.20 \leq P < 0.50$	$2 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	$1 \cdot 10^{-4}$
$0.50 \leq P < 1.00$	$5 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$1 \cdot 10^{-4}$
$1.00 \leq P < 2.00$	$1 \cdot 10^{-2}$	$5 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	$2 \cdot 10^{-4}$
$2.00 \leq P < 5.00$	$2 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	$5 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$5 \cdot 10^{-4}$
$5.00 \leq P < 10.00$	$5 \cdot 10^{-2}$	$2 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	$5 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$1 \cdot 10^{-3}$
$10.00 \leq P < 20.00$	$1 \cdot 10^{-1}$	$5 \cdot 10^{-2}$	$2 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	$5 \cdot 10^{-3}$	$2 \cdot 10^{-3}$
$20.00 \leq P < 50.00$	$2 \cdot 10^{-1}$	$1 \cdot 10^{-1}$	$5 \cdot 10^{-2}$	$2 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	$5 \cdot 10^{-3}$
$50.00 \leq P < 100.00$	$5 \cdot 10^{-1}$	$2 \cdot 10^{-1}$	$1 \cdot 10^{-1}$	$5 \cdot 10^{-2}$	$2 \cdot 10^{-2}$	$1 \cdot 10^{-2}$
$100.00 \leq P < 200.00$	$1 \cdot 10^0$	$4 \cdot 10^{-1}$	$2 \cdot 10^{-1}$	$1 \cdot 10^{-1}$	$5 \cdot 10^{-2}$	$2 \cdot 10^{-2}$
$200.00 \leq P < 500.00$	$2 \cdot 10^0$	$1 \cdot 10^0$	$4 \cdot 10^{-1}$	$2 \cdot 10^{-1}$	$1 \cdot 10^{-1}$	$5 \cdot 10^{-2}$
$500.00 \leq P < 1000.00$	$5 \cdot 10^0$	$2 \cdot 10^0$	$1 \cdot 10^0$	$4 \cdot 10^{-1}$	$2 \cdot 10^{-1}$	$1 \cdot 10^{-1}$
$1000.00 \leq P < 2000.00$	$1 \cdot 10^1$	$5 \cdot 10^0$	$2 \cdot 10^0$	$1 \cdot 10^0$	$4 \cdot 10^{-1}$	$2 \cdot 10^{-1}$
$2000.00 \leq P < 5000.00$	$2 \cdot 10^1$	$1 \cdot 10^1$	$5 \cdot 10^0$	$2 \cdot 10^0$	$1 \cdot 10^0$	$4 \cdot 10^{-1}$
$5000.00 \leq P < 10000.00$	$5 \cdot 10^1$	$2 \cdot 10^1$	$1 \cdot 10^1$	$5 \cdot 10^0$	$2 \cdot 10^0$	$1 \cdot 10^0$
$10000.00 \leq P < 20000.00$	$1 \cdot 10^2$	$5 \cdot 10^1$	$2 \cdot 10^1$	$1 \cdot 10^1$	$5 \cdot 10^0$	$2 \cdot 10^0$
$20000.00 \leq P < 50000.00$	$2 \cdot 10^2$	$1 \cdot 10^2$	$5 \cdot 10^1$	$2 \cdot 10^1$	$1 \cdot 10^1$	$5 \cdot 10^0$
$50000.00 \leq P < \infty$	$5 \cdot 10^2$	$2 \cdot 10^2$	$1 \cdot 10^2$	$5 \cdot 10^1$	$2 \cdot 10^1$	$1 \cdot 10^1$

Liquiditätsbänder: LB₁: $0 \leq \text{ADNT} < 10$, LB₂: $10 \leq \text{ADNT} < 80$, LB₃: $80 \leq \text{ADNT} < 600$, LB₄: $600 \leq \text{ADNT} < 2000$, LB₅: $2000 \leq \text{ADNT} < 9000$, LB₆: $9000 \leq \text{ADNT} < \infty$.

Tabelle 4.1: Tickgrößen-Tabelle für Aktien, Zertifikate und börsengehandelte Fonds (ETFs). Die Tickgrößen bzw. Präzisionsanforderungen sind strukturiert nach Preis- und Liquiditätsbewertungen. Quelle: EU (2017): S. 417.

Regulatorische Rahmenbedingungen:

Der Handel mit Finanzinstrumenten unterliegt diversen nationalen und internationalen Gesetzen und Richtlinien, die den Wertpapierhandel per se, die Veröffentlichung von Handelsaktivitäten und die Preisstrukturen regulieren. So sind auch Tickgrößen als grundlegende kleinste Einheit für die Bemessung von Kursen und Kursänderungen reguliert.

Die Europäische Union (EU) reguliert die Größe von Ticks durch Rechtsvorschriften. Als Ziele der Tickgrößen-Regulierung sind der Erhalt der Marktstabilität und die Auslastungssteuerung der Systemkapazitäten genannt (vgl. EU (2014), Art. 48 (6)). Das umzusetzende Tickgrößen-System soll hierbei an das Liquiditätsprofil und den Typ des Finanzinstrumentes adaptiert sein (vgl. EU (2014), Art. 49). In Tabelle 4.1 sind die Vorgaben an die Tickgröße für Aktien, Zertifikate und börsengehandelte Fonds (ETFs), veröffentlicht in der Ergänzung EU (2017) für EU (2014), dargestellt. Es sind die Tickgrößen

nach 19 Preisniveaus und sechs Liquiditätsbändern aufgeschlüsselt. Die Liquidität ist über die durchschnittlichen Transaktionsanzahl (ADNT, Average Daily Number of Transactions) bemessen. Die möglichen Tickgrößen liegen dabei zwischen 0.0001 EUR und 500.00 EUR. Ähnliche weniger komplexe Rahmenbedingungen sind zum Beispiel in den Vereinigten Staaten von Amerika gegeben. Wertpapiere mit einem Kurs unter 1.00 USD haben eine Tickgröße von 0.0001 USD und Wertpapiere mit einem Kurs von oder über 1.00 USD haben eine Tickgröße von 0.01 USD (vgl. SEC (2020): S. 857, 17 CFR § 242.612).

Die Begrenzung der minimalen Tickgröße hat somit einen direkten Einfluss auf die Präzision der bereitgestellten Daten derart, dass die Preisbewertungen bzgl. der minimalen Tickgröße gerastet bzw. gerundet werden und somit Bindungen auftreten können. Entsprechend ist es sinnvoll, diese Einschränkung im datengenerierenden Algorithmus zu berücksichtigen. Hierfür wäre die Rasterung mit einem Rasterabstand $r \in \mathbb{Q}$ wie z. B. die Tickgrößen-Zuordnungen in Tabelle 4.1, oder als vereinfachende Annahme die Rundung der Daten zur Präzision 10^k , $k \in \mathbb{Z}$, zielführend.

Minimalpreise als Anforderung der Börsen:

Zusätzlich zu den regulatorischen Anforderungen sind Einschränkungen durch die jeweiligen Börsen, an denen die Finanzinstrumente gehandelt werden, möglich. Für eine Notierung an einer Börse können gewisse Mindestanforderungen, wie die *Anzahl der ausgegebenen Wertpapiere*, die initiale *Marktkapitalisierung* oder auch ein *minimaler Durchschnittspreis* vorgegeben sein. Zusätzlich zur natürlichen Preisuntergrenze von 0.00 USD sind zum Beispiel an den US-amerikanischen Börsen NYSE (New York Stock Exchange) und NASDAQ (National Association of Securities Dealers Automated Quotations) ein Mindestpreis von 1.00 USD, durchschnittlich über 30 Tage, gefordert. Weiter ist ein initialer Mindestpreis von 4.00 USD pro Wertpapier zur Notierung auf der Börse NASDAQ festgelegt (vgl. NASDAQ (2021b): S. 6; NASDAQ (2021a): S. 3-4; NYSE (o.J.): § 802.01 (C)).

Da also Einschränkungen des Messbereichs seitens der Börsen vorliegen können, ist eine Berücksichtigung im datengenerierenden Algorithmus sinnvoll. Es ist zu beachten, dass die Preisuntergrenze der zwei genannten Börsen im Durchschnitt über 30 Tage gilt; an einzelnen Tagen sind Preise zwischen 0.00 USD und 1.00 USD zulässig, so dass eine Einschränkung des Messbereichs abseits der natürlichen Preisuntergrenze nur in begründeten Fällen durchzuführen ist.

Komplexität und fehlerhafte Datenübertragungen:

Auf Finanzmärkten interagieren eine Vielzahl von Akteuren und Institutionen, die den unterschiedlichen regionalen und globalen Rahmenbedingungen folgen müssen. Weiter findet die Handelsaktivität auf mehreren Handelsplätzen und -plattformen statt. Somit stellen Finanzmärkte hochkomplexe und stark vernetzte Systeme dar.

Als veranschaulichende Beispiele für diese Komplexität werden zwei der drei im Abschnitt 4.3 genutzten Informationsdienste und den zugrundeliegenden Datenquellen genannt:

- Der Anbieter *Yahoo! Finance* benötigt zur Marktdeckung des Services Daten von 79 Börsen aus 50 Ländern. Hierfür werden die Daten über Drittanbieter wie *Morningstar*, *Commodity Systems Inc.* und weiteren Unternehmen bezogen (vgl. *Yahoo Inc.* (2021b)).
- Die Plattform *Investing.com* des Unternehmens *Fusion Media Ltd.* bietet Finanzmarktdaten von offiziellen Quellen wie Börsen und Unternehmen sowie durch die jeweiligen Marktteilnehmer bereitgestellte Daten an (vgl. *Fusion Media Ltd.* (2021b)). Es ist also für diese Plattform anzunehmen, dass sowohl voll- und teilautomatisierte, als auch manuelle Datenbereitstellungen vorliegen.

Zusätzlich dazu gelten an Börsen unterschiedliche Handelstage; diese sind von regionalen Gegebenheiten abhängig. An sogenannten Börsenfeiertagen, die in den Handelskalendern der jeweiligen Börsen veröffentlicht werden, findet keine oder eine nur verringerte Handelsaktivität statt (vgl. *Buske* (2021)). Ein direktes, außerbörsliches Handeln (OTC, Over The Counter) von Wertpapieren ist an diesen Tagen weiterhin möglich: Der OTC-Handel ist dabei an rechtliche Regulierungen gebunden. Die Rahmenbedingungen einzelner Börsen sind beim OTC-Handel nicht zu berücksichtigen (vgl. *Heldt* (2018)).

Die an diesen Beispielen aufgezeigte Komplexität und Heterogenität ist bei der Beschaffung und Verarbeitung der Finanzmarktdaten seitens der Datenanbieter zu berücksichtigen. Da unter diesen Bedingungen technische Fehler in der Verarbeitung und Übertragung sowie menschliche Eingabefehler auftreten können, wird darüber eine fehlerfreie Datenbereitstellung erschwert (vgl. *Dhevi* (2014): S. 255). Somit sind fehlende Daten in Form von NA- oder Null-Einträgen, das wiederholte Veröffentlichen voriger Werte oder der Datenbezug von anderen Quellen möglich. Eine umfassende Modellierung der Ursachen für unvollständige oder fehlerhafte Daten ist nicht möglich und zielführend, da Informationen über die Prozesse der einzelnen Akteure nicht

öffentlich vorliegen. Alternativ dazu ist die zufällige Modellierung der Wertauslassung bzw. Ersetzung mit Wahrscheinlichkeit $p \in [0, 1)$ ohne weitere Informationen möglich. Dies wird im Algorithmus berücksichtigt. NA-Werte oder Wertwiederholungen können zu Bindungen in der Stichprobe führen.

Betrachtung von Renditen:

Die oben genannten Einschränkungen und Rahmenbedingungen wurden speziell für Preisdaten formuliert. Für viele Anwendungen und Modelle für Finanzmärkte sind Renditen bzw. log-Renditen von Interesse, da diese die relativen Veränderungen von Komponenten eines Portfolios, Risikofaktoren etc. beschreiben. Die log-Renditen $(x_t)_{t=1,\dots,n}$ für eine Preiszeitreihe $(y_t)_{t=0,\dots,n}$ sind definiert als

$$x_t := \ln \left(\frac{y_t}{y_{t-1}} \right) \quad (t \in \{1, \dots, n\}) \quad (4-1)$$

(vgl. McNeil et al. (2015): S. 49, 79-80). Die log-Renditen unterliegen somit nach Konstruktion (4-1) nicht denselben Einschränkungen wie die Preisdaten. Dennoch ist es plausibel anzunehmen, dass die durch die Informationsdienste bereitgestellten Daten unvollständig oder gerundet vorliegen.

Entsprechend ist die Annahme, dass in empirischen Renditedaten Bindungen vorliegen können, gerechtfertigt. Unter Idealbedingungen, d. h. sofern nachträglich keine weiteren Effekte einwirken, stehen die Bindungen bei Preis- und Renditedaten in Beziehung zueinander. Da die Rendite x_t auf Basis der aufeinanderfolgenden Preise $y_{t-1}, y_t > 0$ berechnet wird, gilt

$$x_t = x_s \iff \frac{y_t}{y_{t-1}} = \frac{y_s}{y_{s-1}}$$

für Bindungen im Nicht-NA-Fall. Bei NA- bzw. Null-Werten für den Preis y_t sind die beiden zugehörigen Renditen nicht bestimmbar. Es gilt

$$y_t = 0 \text{ oder } y_t = \text{NA} \implies x_t = x_{t+1} = \text{NA},$$

da die Logarithmen aus Formel (4-1) nicht ausgewertet werden können.

4.2 Generierung synthetischer Datensätze

Aufbauend auf die im vorigen Abschnitt 4.1 behandelten Rahmenbedingungen wird in diesem Abschnitt ein Algorithmus zur Generierung synthetischer Datensätze für die Anwendung in Simulationsstudien entwickelt. Dieser soll unter Modellierung der Rahmenbedingungen kontrolliert Bindungen hervorrufen, sodass die Copula-Methoden darüber analysiert werden können.

Zu Beginn wird ein bereits existierender Algorithmus als Entwicklungsgrundlage gesichtet. Der Algorithmus ist aus Kojadinovic (2017), Procedure 2.1, unter Anlehnung an die hier vorliegenden Notationen entnommen.

4.2.1 Algorithmus (Generierung von Daten nach Kojadinovic (2017)). *Sei $C : [0, 1]^d \rightarrow [0, 1]$ eine Copula und bezeichne $k \in \mathbb{N}$ die Stichprobengröße. Zur Rasterung des Hyperwürfels $[0, 1]^d$ werden die Parameter $k \in \mathbb{N} \cup \{\infty\}$ und $t \in \{0.5, 1, 2\}$ genutzt. Im Fall $k = \infty$ wird keine Rasterung durchgeführt. Der Algorithmus zur Datengenerierung ist definiert als:*

1. Generierung einer Stichprobe bzgl. der Copula C : Realisiere bzgl. der Zufallsvektoren $U_1, \dots, U_n \stackrel{i.i.d.}{\sim} C$ eine Stichprobe u_1, \dots, u_n mit den Koordinaten notiert als $u_i = (u_{i1}, \dots, u_{id})$ für alle $i \in \{1, \dots, n\}$.
2. Rasterung – Reduktion der Messpräzision: Falls $k < \infty$ gilt, führe die folgenden weiteren Schritte durch:
 - (i.) Konstruktion des Rasters: Die (eindimensionale) Rasterung des Intervalls $[0, 1]$ ist konstruiert als disjunkte Vereinigung

$$[0, 1] = A_0 \dot{\cup} A_1 \dot{\cup} \dots \dot{\cup} A_{k-1}$$

von Intervallen

$$A_0 := [a_0, a_1], \quad A_l := (a_l, a_{l+1}] \quad (l \in \{0, \dots, k-1\})$$

zwischen den Punkten $a_l := (\frac{l}{k})^t$ für alle $l \in \{0, \dots, k\}$.

- (ii.) Rasterung der Datenpunkte: Überschreibe die u_{ij} mit den Mittelpunkten des jeweils umschließenden Intervalls A_l . Setze dazu

$$u_{ij} := \sum_{l=0}^{k-1} \frac{a_{l+1} - a_l}{2} \cdot \mathbb{1}_{A_l}(u_{ij})$$

für alle $(i, j) \in \{1, \dots, n\} \times \{1, \dots, d\}$.

Nach der Durchführung der obigen Schritte wird die modifizierte Stichprobe u_1, \dots, u_n mit $u_i := (u_{i1}, \dots, u_{id})$ ausgegeben.

Es ist anzumerken, dass die Rasterung im Algorithmus 4.2.1 je nach Wahl des Parameters $t \in \{0.5, 1, 2\}$ eine abnehmende, konstante oder wachsende Rasterbreite besitzt (vgl. Kojadinovic (2017), S. 3). Da der Algorithmus 4.2.1 nicht die Erkenntnisse und Rahmenbedingungen aus dem Abschnitt 4.1 abbilden kann, sind Anpassungen und Ergänzungen notwendig.

4.2.2 Definition (Rundungsfunktionen). *Bezeichne $\lfloor \cdot \rfloor, \lceil \cdot \rceil : \mathbb{R} \rightarrow \mathbb{Z}$ die untere und die obere Gaußklammer, auch Abrundungs- und Aufrundungsfunktion genannt. Diese sind formal definiert als*

$$\lfloor x \rfloor := \max \{z \in \mathbb{Z} \mid z \leqq x\} \quad \text{und} \quad \lceil x \rceil := \min \{z \in \mathbb{Z} \mid z \geqq x\}.$$

Die symmetrische Rundungsfunktion $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$ ist darüber definiert als

$$\lfloor x \rfloor := \operatorname{sgn}(x) \cdot \left\lceil \lfloor 2|x| \rfloor / 2 \right\rceil,$$

wobei bzgl. der zu rundenden Dezimalstelle bei den Ziffern $\{0, \dots, 4\}$ abgerundet und bei $\{5, \dots, 9\}$ aufgerundet wird.

4.2.3 Algorithmus (Generierung synthetischer Datensätze mit Bindungen). *Seien $C : [0, 1]^d \rightarrow [0, 1]$ eine Copula und $F_1, \dots, F_d : \mathbb{R} \rightarrow [0, 1]$ Marginale mit zugehörigen Quantilfunktionen $F_1^\leftarrow, \dots, F_d^\leftarrow$. Die Marginale können auch empirischer Natur oder $\text{Unif}_{[0,1]}$ -Marginale sein, falls ausschließlich Pseudobechtungen notwendig sind. Die Zahl $n \in \mathbb{N}$ beschreibe die Stichprobengröße.*

Um gängige Datenfehler und Messeinschränkungen zu modellieren, werden weitere Parameter für den Algorithmus benötigt:

- Messpräzision: *Der Vektor $r = (r_1, \dots, r_d)$ mit $r_j \in [0, \infty)$ beschreibt die Rasterung bzw. Rundung zum Wert r_j der jeweiligen j -ten Koordinate. Im Fall $r_j = 0$ wird keine Modifikation der Werte durchgeführt.*
- Zulässiger Messbereich: *Die Vektoren $x_{\text{lower}}, x_{\text{upper}} \in \overline{\mathbb{R}}^d$ parametrisieren den Messbereich $\bigtimes_{j=1}^d [x_{j;\text{lower}}, x_{j;\text{upper}}]$. Werte außerhalb dieser Menge werden auf den nächstgelegenen zulässigen Wert abgeändert.*
- Fehlerhafte Datenübertragung: *Der Vektor $p = (p_1, \dots, p_d) \in [0, 1]^d$ modelliert die Wahrscheinlichkeit fehlerhafter Datenübertragungen, bei welchen die x_{ij} durch Standardwerte $x_{j;\text{default}}$ ersetzt werden. Diese sind fixierte Werte $x_{ij;\text{default}} \equiv x_{j;\text{default}} \in \mathbb{R}$ oder datenadaptive Werte, wie ein Moving Average $x_{ij;\text{default}} = k^{-1} \cdot (x_{(i-1)j} + \dots + x_{(i-1-k)j})$ über k Zeitpunkte. Auch die Wahl $x_{j;\text{default}} := \text{NA}$ ist möglich.*

Der Algorithmus ist definiert über die folgenden fünf Schritte, wobei zu Beginn “perfekte” Daten generiert und diese dann entsprechend der Parameter angepasst und überschrieben werden. Die folgenden Anweisungen verstehen sich für alle $(i, j) \in \{1, \dots, n\} \times \{1, \dots, d\}$.

1. Generierung einer Stichprobe bezüglich der Copula C : Realisiere bezüglich der Zufallsvektoren $U_1, \dots, U_n \stackrel{i.i.d.}{\sim} C$ eine Stichprobe u_1, \dots, u_n mit den Koordinaten notiert als $u_i = (u_{i1}, \dots, u_{id})$ für alle $i \in \{1, \dots, n\}$.
2. Quantiltransformation über die Marginale – Zielstrukturmodellierung: Setze die Vektoren $x_i := (x_{i1}, \dots, x_{id})$ über die Quantiltransformationen

$$x_{ij} := F_j^{-1}(u_{ij}).$$

3. Rundung – Reduktion der Messpräzision: Überschreibe die x_{ij} mit den gerundeten Werten, falls notwendig ($r_j < \infty$). Es gilt die Vorschrift

$$x_{ij} := \begin{cases} \lfloor x_{ij} \cdot r_j^{-1} \rfloor \cdot r_j, & r_j > 0, \\ x_{ij}, & r_j = 0 \end{cases}.$$

4. Auslassung von Werten – Unvollständige Datenübertragung: Modelliere zu Beginn die zufällige Auswahl der zu überschreibenden Elemente und überschreibe diese entsprechend.

- (i.) Zufallsauswahl: Generiere eine Stichprobe $b_1, \dots, b_n \in \{0, 1\}^d$ mit $b_i = (b_{i1}, \dots, b_{id})$ bzgl. des Zufallsvektors

$$B_i := (B_{i1}, \dots, B_{id}) : B_{1j}, \dots, B_{dj} \stackrel{i.i.d.}{\sim} \text{Bin}_{1, p_j},$$

wobei $B_i \perp\!\!\!\perp B_k$ für $i \neq k$ gelte.

- (ii.) Überschreiben der Werte: Ersetze die x_{ij} , falls $b_{ij} = 1$, mit den jeweiligen Standardwerten. Die Ersetzung folgt der Vorschrift

$$x_{ij} := (1 - b_{ij}) \cdot x_{ij} + b_{ij} \cdot x_{ij; \text{default}}.$$

Im datenadaptiven Fall sind die Ersetzungen im aufsteigenden Index $i \in \{1, \dots, n\}$ durchzuführen.

5. Begrenzung von Werten – Einschränkung auf den Messbereich: Schränke die Werte x_{ij} auf das jeweilige Intervall $[x_{j;\text{lower}}, x_{j;\text{upper}}]$ ein. Setze

$$x_{ij} := ((x_{ij} \vee x_{j;\text{lower}}) \wedge x_{j;\text{upper}}).$$

Nach der Durchführung der obigen Schritte wird die modifizierte Stichprobe x_1, \dots, x_n mit $x_i := (x_{i1}, \dots, x_{id})$ ausgegeben.

Auswahl und Herleitung der Parameterwerte:

Zur Anwendung des Algorithmus sind die Parameter zu setzen. Einige Parameter können bereits aus dem Sachzusammenhang festgelegt werden. Zu diesen gehören der zulässige Messbereich, parametrisiert über die Vektoren $x_{\text{lower}}, x_{\text{upper}} \in \overline{\mathbb{R}}^d$, und die Wahl des Standardwert-Vektors $x_{\text{default}} \in \mathbb{R}^d$.

Für den Umgang mit Preisdaten ist es naheliegend, die natürliche Preisuntergrenze von 0.00 USD zu wählen, da börsenspezifische Mindestpreise durchschnittlicher Natur sind; eine Preisobergrenze ist nicht definiert. Für die Modellierung von Preisdaten ist die Wahl

$$x_{\text{lower}} = (0, \dots, 0) \quad \text{und} \quad x_{\text{upper}} = (+\infty, \dots, +\infty)$$

zweckgemäß. Im Umgang mit Renditen sind Einschränkungen dieser Art nicht notwendig, da Renditen nicht beschränkt werden, entsprechend ist

$$x_{\text{lower}} = (-\infty, \dots, -\infty) \quad \text{und} \quad x_{\text{upper}} = (+\infty, \dots, +\infty)$$

für die Modellierung von Renditedaten zu setzen.

Für die Wahl des Standardwert-Vektors für Preisdaten wird in dieser Arbeit aus Komplexitätsgründen

$$x_{\text{default}} = (0, \dots, 0)$$

gesetzt, da im Sachzusammenhang Preise strikt größer als 0.00 USD betrachtet werden und somit die Null als NA interpretierbar bleibt. Alternativ sind datenadaptive Ansätze möglich.

Unabhängig der Betrachtung von Preisen und Renditen sind bzgl. der anderen genannten Rahmenbedingungen die Parameter über eine entsprechende Stichprobe herzuleiten, da diese nicht direkt aus dem Sachzusammenhang bestimmbar sind. Für die Schätzung der Messpräzision $r = (r_1, \dots, r_d) \in \mathbb{R}_{\geq}^d$ ist beispielsweise die maximale Anzahl der Dezimalstellen derart, dass die letzte Ziffer ungleich Null ist, eine mögliche Grundlage zur Schätzung von r_j . Der Schätzer \hat{r} zur Messpräzision, welcher in dieser Arbeit zur Auswertung der empirischen Daten genutzt wird, ist mit Bezug auf eine Stichprobe x_1, \dots, x_n definiert als

$$\hat{r} := 10^{-k^*}$$

$$k^* := \min \{ k \in \mathbb{N}_0 : \lfloor x_i \cdot 10^k \rfloor = x_i \cdot 10^k \text{ für } i \in \{1, \dots, n\} \}.$$

Hierbei gelte $\lfloor \text{NA} \cdot 10^k \rfloor = \text{NA} = \text{NA} \cdot 10^k$ für alle k . Dies bleibt insbesondere bei einer Interpretation von NA als Null erhalten.

Bei den fehlerhaft übertragenen Werten ist zu beachten, dass eine eindeutige und vollständige Identifizierung solcher Werte in einer Stichprobe nicht garantiert werden kann. Einerseits kann eine Kennzeichnung als NA durch den Datenanbieter erfolgen, was eine direkte Identifizierung ermöglicht. Andererseits können fehlende Werte x_t ebenso durch den letzten bereitgestellten Wert x_{t-1} ersetzt oder aus anderen Quellen bezogen worden sein. Da eine Bindung aufeinanderfolgender gleicher Werte durch die Rasterung bzw. Rundung der Daten auftreten kann, kann eine solche Bindung nicht eindeutig zugeordnet werden.

Somit werden in der Analyse der empirischen Daten ausschließlich NA- bzw. Null-Werte für die Analyse von fehlerhaften Übertragungen herangezogen. Der hier verwendete Schätzer \hat{p} für die Wahrscheinlichkeit p einer fehlerhaften Datenübertragung wird bzgl. einer Stichprobe x_1, \dots, x_n als

$$\hat{p} := \frac{|\{i \in \{1, \dots, n\} : x_i = \text{NA}\}|}{n}$$

definiert, was dem NA-Anteil an der gesamten Stichprobe entspricht.

4.3 Sichtung empirischer Finanzmarktdaten

Für die realitätsnahe Modellierung der Rahmenbedingungen aus Abschnitt 4.1 im Algorithmus 4.2.3 ist eine Sichtung von empirischen Daten notwendig. Da Finanzinstrumente in viele unterschiedliche Gruppen (z. B. Regionen, Sektoren, Art) gebündelt werden können und zusätzlich jeweils mehrere Bezugsquellen genutzt werden können, ist die Betrachtung einer umfassenden Auswahl von Finanzmarktdaten sinnvoll.

In diesem Abschnitt werden 179 Finanzinstrumente aus drei populären Quellen bezogen; hierzu gehören 70 Aktienindizes aus vier Regionen, 70 Aktien aus sieben Sektoren bzw. Sektor-Zusammenfassungen sowie 19 Rohstoffe, zehn Währungen und zehn Kryptowährungen. Die hier verwendeten Gruppenbezeichnungen sind in Tabelle 4.2 vorgestellt. Die empirischen Daten wurden bei Verfügbarkeit aus öffentlichen und nicht-öffentlichen Quellen bezogen. Die drei gewählten Quellen lauten:

- Yahoo! Finance: *Yahoo! Finance ist ein öffentlicher und kostenfreier Informationsdienst des Unternehmens Yahoo!, auf dem neben Börsen- und Finanznachrichten auch Finanzmarktdaten und Analysetools bereitgestellt werden. Die Finanzmarktdaten werden von 79 Börsen in 50 Ländern über Anbieter wie Morningstar oder Refinitiv bezogen und ohne weitere Validierung und ohne Gewähr veröffentlicht (vgl. Yahoo Inc. (2021b)). Die Bezugsreferenz lautet: Yahoo Inc. (2021a).*

- Investing.com: *Das kostenfreie Finanzportal Investing.com des Anbieters Fusion Media Ltd. stellt Daten für 300 000 Finanzinstrumente, welche auf 250 Börsen gehandelt werden, bereit (vgl. Fusion Media Ltd. (2021c)). Neben Daten, die durch Börsen oder ähnliche Anbieter bereitgestellt werden, wird die Datenbank durch Informationen von Marktakteuren ergänzt, so dass hier insbesondere Daten von weniger populären Märkten veröffentlicht werden können (vgl. Fusion Media Ltd. (2021b)). Der Bezug erfolgte aus Fusion Media Ltd. (2021a).*
- Thomson Reuters Refinitiv Eikon: *Refinitiv Eikon ist eine nicht öffentlich zugängliche Datenbank, d. h. die Daten sind in einem kostenpflichtigen Abonnement verfügbar. Laut Werbeunterlagen seitens des Unternehmens Refinitiv bündelt Eikon Daten über 99 % der Wertpapiere aus 150 Ländern; diese werden aus 2 000 individuellen Quellen bezogen und verifiziert (vgl. Refinitiv (2019): S. 3). Die Bezugsreferenz für diese Datenbank lautet: Thomson Reuters (2021).*

Code	Art	Gruppenbezeichnung	Anz.
EU	Region	Wirtschaftsraum Europa inkl. UK, Russland	20
AMER	Region	Wirtschaftsraum Nord-, Mittel- und Südamerika	15
APAC	Region	Wirtschaftsraum Asien und (West-)Pazifik	15
MEA	Region	Wirtschaftsräume Naher / Mittlerer Osten, Afrika	20
EG/UT	Sektor	Energiewirtschaft und Versorgungsunternehmen	10
MT	Sektor	Materialien / Grundstoffe	10
FN	Sektor	Finanzdienstleistungen	10
ID	Sektor	Industrieunternehmen	10
CN	Sektor	Konsumgüter- und Handelsunternehmen	10
LS	Sektor	Gesundheitswesen	10
TC	Sektor	Technologieunternehmen	10
COMM	Andere	Rohstoffhandel (Futures)	19
FX	Andere	Wechselkurse und Währungen	10
CX	Andere	Kryptowährungen	10

Tabelle 4.2: Übersicht der Gruppenbezeichnungen, sortiert in Reihenfolge des Auftretens.

Direkte Veröffentlichungen seitens der Börsen oder Dienste, die sich nur auf eine Art der genannten Finanzinstrumente spezialisieren, wurden zwecks der direkten Vergleichbarkeit nicht genutzt. Die Daten wurden von den drei genannten Quellen als .csv- oder .xlsx-Datei heruntergeladen, falls verfügbar, und ansonsten die Daten über den Quellcode der Webseite extrahiert. Es wurden die größtmöglichen Zeiträume für Tagesdaten ausgewählt, wobei aus

technischen Gründen bei Eikon lediglich Daten ab dem Jahr 2000 verfügbar waren.

In den folgenden Tabellen 4.3 bis 4.6 werden die Finanzinstrumente mit Namen und einem Kürzel gekennzeichnet. Es ist anzumerken, dass die Namen und Kürzel nicht einheitlich in den Quellen vorliegen, so dass Abweichungen auftreten können. Bei der Auswahl der Finanzinstrumente wurde angestrebt, jeweils das ursprüngliche Instrument auszuwählen, damit die Betrachtung von Derivaten wie Optionen, Futures etc. vermieden wird. Ausnahmen sind die Finanzinstrumente für den Rohstoffhandel (COMM), bei dem Futures mit „_F“ gekennzeichnet sind; die Kürzel für Wechselkurse enden mit „_X“.

Die bei Verfügbarkeit betrachtete Preiszeitreihe entspricht den täglichen Schlusskursen (Closing Price) je Finanzinstrument. Die zugehörigen Renditen werden ad hoc berechnet, da diese nicht in allen Quellen bereitgestellt werden und eine niedrigere Präzision innehaben können, als eine direkte Berechnung aus den Preisanalysen. Hierfür wurden unabhängig der Art der bereitgestellten Renditen die log-Renditen berechnet und analysiert.

	Allgemeine Informationen			Yahoo! Finance			Investing.com			Refinitiv Eikon			
Nr.	Code	Name	Gruppe	n	Präz. (emp.)	Preisdaten Bind.	n	Präz. (emp.)	Preisdaten Bind.	n	Präz. (emp.)	Preisdaten Bind.	
1	SXSE	EURO STOXX 50	EU	3.6	10 ⁻²	2.7%	0.4%	0.9%	0.8%	5.2	10 ⁻²	0.5%	0.0%
2	DAX30	Deutscher Aktienindex	EU	8.4	10 ⁻²	2.8%	1.9%	3.2%	3.2%	5.2	10 ⁻²	0.5%	0.0%
3	FTSE	Financial Times Stock Exch.	EU	—	—	—	—	—	—	5.0	10 ⁻²	1.5%	0.0%
4	CAC40	Coatation Assistée en Continu	EU	3.7	10 ⁻²	1.5%	0.1%	0.2%	0.2%	8.6	10 ⁻²	2.4%	0.0%
5	RTSI	Russian Trading System Index	EU	2.1	10 ⁻²	3.5%	1.8%	2.8%	2.7%	6.5	10 ⁻²	3.9%	0.0%
6	MEX	Moscow Exchange	EU	7.2	10 ⁻¹	12.0%	1.9%	3.6%	3.4%	4.2	10 ⁻²	4.2%	0.0%
7	IBEX35	Indice Bursátil Español	EU	7.4	10 ⁻⁶	13.9%	1.3%	2.3%	2.2%	—	—	0.3%	0.0%
8	AEX25	Amsterdam Exch. Index	EU	6.2	10 ⁻¹	4.6%	3.6%	5.7%	5.6%	8.5	10 ⁻²	29.9%	0.0%
9	BIST	Borsa Istanbul 100	EU	7.9	10 ⁻²	6.6%	2.0%	3.6%	3.5%	8.4	10 ⁻²	5.8%	0.0%
10	SMI	Swiss Market Index	EU	3.2	10 ⁻²	4.0%	0.2%	0.6%	0.4%	8.7	10 ⁻²	7.2%	0.0%
11	OMX SE30	OMX Stockholm 30	EU	2.1	10 ⁻²	4.3%	3.9%	7.2%	7.1%	2.1	10 ⁻²	2.5%	0.0%
12	WTG30	Warszawski Indeks Gieldowy 30	EU	2.1	10 ⁻²	7.9%	3.4%	6.3%	6.1%	0.7	10 ⁻²	1.6%	0.0%
13	OSEBX	Oslø Børs Benchmark Index	EU	4.2	10 ⁻²	10.0%	2.2%	4.8%	4.0%	5.5	10 ⁻²	8.4%	0.0%
14	OBX	Oslo Børs Index	EU	7.3	10 ⁻²	6.1%	2.8%	5.1%	5.0%	0.1	10 ⁻²	0.1%	0.0%
15	ATX	Austrian Traded Index	EU	7.8	10 ⁻²	8.1%	1.9%	3.5%	3.3%	7.5	10 ⁻²	2.3%	0.0%
16	BEL20	Belgian 20 Index	EU	—	—	—	—	5.1	10 ⁻²	8.0%	0.0%	0.1%	0.0%
17	OMX C20	OMX Copenhagen 20	EU	2.1	10 ⁻²	2.1%	0.6%	1.4%	1.3%	5.1	10 ⁻²	2.0%	0.0%
18	OMX H25	OMX Helsinki 25	EU	—	—	—	—	5.1	10 ⁻²	2.0%	0.0%	0.1%	0.0%
19	ASE	Athens Stock Exch. Comp.	EU	6.2	10 ⁻²	2.9%	1.7%	2.9%	2.8%	5.1	10 ⁻²	93.0%	0.0%
20	ISEQ	Irish Stock Exch. Overall	EU	7.4	10 ⁻²	10.6%	0.0%	0.1%	0.0%	—	—	—	—
21	DJI	Dow Jones Industrial Avg.	AMER	12.7	10 ⁻²	12.6%	0.0%	0.1%	0.0%	3.7	10 ⁻²	0.8%	0.0%
22	COMPX	NASDAQ Comp.	AMER	9.0	10 ⁻³	3.3%	0.0%	0.1%	0.0%	10.4	10 ⁻¹	27.9%	0.0%
23	NDX	NASDAQ 100	AMER	11.7	10 ⁻²	12.1%	0.0%	0.3%	0.0%	9.0	10 ⁻²	4.7%	0.0%
24	SP500	Standard & Poor's 500	AMER	14.0	10 ⁻²	27.2%	0.0%	0.9%	0.0%	3.6	10 ⁻²	2.7%	0.4%
25	NYA	New York Stock Exch. Comp.	AMER	7.2	10 ⁻¹	10.7%	2.5%	4.9%	4.6%	10.5	10 ⁻¹	13.6%	0.0%
26	IBOVESPA	Indice Bovespa	AMER	2.1	10 ⁻²	14.1%	0.0%	1.6%	1.4%	4.6	10 ⁻²	0.5%	0.0%
27	IBX50	Brazil 50 Index	AMER	10.8	10 ⁻²	11.0%	2.3%	4.5%	4.4%	10.5	10 ⁻²	0.8%	0.0%
28	GSPTSE	Toronto Stock Exch. Comp.	AMER	—	—	—	—	6.9	10 ⁻²	9.1%	0.0%	0.1%	0.0%
29	IPSA	Bolsa de Comercio de Santiago	AMER	7.6	10 ⁻²	4.2%	2.3%	4.4%	4.3%	8.7	10 ⁻²	1.0%	0.0%
30	MXX	Índice de Precios y Cotizaciones AMER	AMER	6.2	10 ⁻²	4.2%	2.7%	4.8%	4.7%	3.0	10 ⁻²	4.7%	0.0%
31	MERV	Mercado d. Val. d. Buenos Aires AMER	AMER	—	—	—	—	—	—	—	—	—	—
32	COLCAP	Colombia Stock Exch. AMER	AMER	—	—	—	—	—	—	—	—	—	—
33	IBC	Índice Bursátil d. Capitalización AMER	AMER	—	—	—	—	—	—	—	—	—	—
34	SPBLPGPT	Lima Stock Exch.	AMER	6.1	10 ⁻²	4.0%	3.4%	6.3%	6.3%	2.5	10 ⁻²	0.4%	0.0%
35	BVG	Bolsa d. Val. Guayaquil	AMER	—	—	—	—	—	—	—	—	—	—
36	SSEC	Shanghai Stock Exch. Comp.	APAC	6.0	10 ⁻³	3.3%	2.9%	3.6%	3.6%	2.4	10 ⁻²	81.0%	0.0%
37	SZI	Shenzhen Stock Exch. Comp.	APAC	6.0	10 ⁻²	3.6%	2.8%	3.5%	3.5%	6.4	10 ⁻²	0.9%	0.0%
38	FTXIN	FTSE China 50	APAC	—	—	—	—	—	—	—	—	—	—
39	HSI	Hang Seng Index	APAC	8.8	10 ⁻²	6.3%	2.9%	5.3%	5.1%	5.1	10 ⁻²	1.2%	0.0%
40	N225	Nikkei 225	APAC	14.4	10 ⁻²	7.4%	4.3%	9.0%	7.6%	5.1	10 ⁻²	0.2%	0.0%
41	NSEI	Nifty 50	APAC	3.4	10 ⁻²	3.7%	1.0%	2.1%	1.9%	6.4	10 ⁻²	5.8%	0.0%
42	BSESN	BSE Sensex	APAC	6.0	10 ⁻²	2.5%	2.2%	4.2%	4.2%	2.6	10 ⁻²	0.0%	0.0%
43	KS50	KOSPI 50	APAC	—	—	—	—	—	—	—	—	—	—
44	AXJO	Australian Stock Exch. 200	APAC	—	—	—	—	—	—	—	—	—	—
45	NZ50	New Zealand Stock Exch. 50	APAC	—	—	—	—	—	—	5.3	10 ⁻²	2.5%	0.0%

Tabelle 4.3: Auflistung der Finanzmarkt-Datensätze 1 bis 45 von 179 über die Regionen EU, AMER und APAC. Es bezeichnen n die Stichprobengröße, Prädz. die empirische Preisdaten-Präzision, Bind. und NA die Bindungsanteile. Die Renditen wurden bzgl. der Preisdaten berechnet. Quelle: Eigene Erstellung auf Basis der Daten in Yahoo Inc. (2021a), Fusion Media Ltd. (2021a) und Thomson Reuters (2021).

Nr. Code	Name	Allgemeine Informationen			Yahoo! Finance			Investing.com			Refinitiv Eikon						
		Gruppe (Stdt.)	n Präz. (emp.)	Bind. NA	Renditen Bind. NA	n Preisdaten (emp.)	Bind. NA	Renditen Bind. NA	n Preisdaten (emp.)	Bind. NA	Renditen Bind. NA	n Preisdaten (emp.)	Bind. NA				
46 GTSM50	Taipei Exch. 50	APAC	-	-	-	-	-	1.5 10^{-2}	18.6%	0.0%	0.5%	0.0%	3.1 10^{-2}	32.5%	0.0%	0.5%	0.0%
47 SETI	Stock Exch. of Thailand	APAC	7.9 10^{-3}	5.1%	4.0%	6.7% 6.6%	9.7 10^{-2}	9.9%	0.0%	0.3%	0.0%	4.9 10^{-3}	4.7%	0.0%	0.1%	0.0%	
48 JKSE	Jakarta Comp. Index	APAC	-	-	-	-	-	7.6 10^{-2}	4.8%	0.0%	0.3%	0.0%	4.9 10^{-2}	0.2%	0.0%	0.0%	0.0%
49 KNSE	FTSE Bursa Malaysia KLCI	APAC	6.0 10^{-2}	7.2%	5.4%	10.0% 9.0%	6.6 10^{-2}	1.6%	0.0%	0.7%	0.0%	4.9 10^{-2}	1.1%	0.0%	0.7%	0.0%	
50 KSE	Karachi Stock Exch. 100	APAC	-	-	-	-	-	2.8 10^{-2}	7.8%	0.0%	1.4%	0.0%	4.6 10^{-3}	2.8%	0.0%	0.4%	0.0%
51 BAX	Bahrein All Share	MEA	-	-	-	-	-	5.8 10^{-2}	1.6%	0.0%	0.3%	0.0%	4.9 10^{-2}	1.0%	0.1%	0.2%	0.1%
52 EGX30	Egyptian Exch. 30	MEA	-	-	-	-	-	1.6 10^{-2}	7.7%	0.0%	1.1%	0.0%	1.5 10^{-2}	12.2%	6.3%	9.8%	9.2%
53 ISX60	Iraq Stock Exch.	MEA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
54 TA35	Tel Aviv 35	MEA	7.3 10^{-2}	10.4%	4.7%	7.9% 7.7%	-	-	-	-	-	4.9 10^{-2}	4.1%	0.0%	0.2%	0.0%	
55 AMGNRLX	Amman Stock Exch.	MEA	-	-	-	-	-	5.1 10^{-2}	4.2%	0.0%	0.3%	0.0%	4.9 10^{-5}	3.1%	0.0%	0.1%	0.0%
56 BLOM	BLOM Stock Index	MEA	-	-	-	-	-	5.4 10^{-2}	20.1%	0.0%	6.6%	0.0%	-	-	-	-	-
57 MSM30	Muscat Stock Exch. 30	MEA	-	-	-	-	-	5.1 10^{-2}	2.0%	0.0%	0.6%	0.0%	-	-	-	-	-
58 PLE	Al-Quds Index	MEA	-	-	-	-	-	4.9 10^{-2}	20.5%	0.0%	1.1%	0.0%	-	-	-	-	-
59 TASI	Tadawul All Share Index	MEA	-	-	-	-	-	6.0 10^{-2}	2.0%	0.0%	0.7%	0.0%	5.2 10^{-2}	1.0%	0.0%	0.3%	0.0%
60 MT30	MSCI Tadawul 30	MEA	-	-	-	-	-	0.6 10^{-2}	2.5%	0.0%	0.7%	0.0%	0.7 10^{-2}	6.1%	0.0%	2.6%	0.0%
61 ADI	Abu Dhabi Index	MEA	-	-	-	-	-	5.2 10^{-2}	1.5%	0.0%	0.1%	0.0%	5.1 10^{-3}	1.2%	0.0%	0.1%	0.0%
62 DFM	Dubai Financial Market	MEA	-	-	-	-	-	4.5 10^{-2}	1.8%	0.0%	0.1%	0.0%	3.6 10^{-3}	93.8%	0.0%	55.6%	0.0%
63 NGSE30	Nigerian Stock Exch. 30	MEA	-	-	-	-	-	-	-	-	-	-	2.9 10^{-2}	3.1%	0.0%	0.1%	0.0%
64 JTOP1	South Africa Top 40	MEA	-	-	-	-	-	6.5 10^{-2}	5.6%	0.0%	0.6%	0.0%	5.0 10^{-4}	0.5%	0.0%	0.2%	0.0%
65 MASL	Moroccan All Shares Index	MEA	-	-	-	-	-	4.9 10^{-2}	1.0%	0.0%	0.1%	0.0%	-	-	-	-	-
66 NSE20	Kenya NSE 20	MEA	-	-	-	-	-	2.3 10^{-2}	2.2%	0.0%	0.1%	0.0%	5.0 10^{-3}	2.1%	0.0%	0.4%	0.0%
67 DSE1	Tanzania All Share	MEA	-	-	-	-	-	2.4 10^{-2}	17.0%	0.0%	8.8%	0.0%	3.5 10^{-2}	21.3%	0.0%	3.5%	0.0%
68 TUNINDEX	Tunisia Stock Exch.	MEA	-	-	-	-	-	5.8 10^{-2}	3.7%	0.0%	0.5%	0.0%	4.9 10^{-2}	3.5%	0.0%	0.5%	0.0%
69 LASILZ	LSE All Share	MEA	-	-	-	-	-	2.2 10^{-2}	53.8%	0.0%	33.1%	0.0%	4.8 10^{-2}	43.5%	0.0%	25.8%	0.0%
70 ALSIUG	Uganda All Share	MEA	-	-	-	-	-	2.4 10^{-2}	54.9%	0.0%	6.7%	0.0%	3.3 10^{-2}	44.2%	0.0%	6.3%	0.0%
71 EOAN.DE	E.ON SE	EG/UT	5.3 10^{-6}	45.1%	0.0%	2.6%	0.0%	2.9 10^{-3}	49.4%	0.0%	2.2%	0.0%	5.1 10^{-6}	21.3%	0.0%	0.4%	0.0%
72 RWE.DE	RWE	EG/UT	6.3 10^{-5}	47.3%	0.0%	3.2%	0.0%	3.0 10^{-3}	44.8%	0.0%	1.7%	0.0%	4.9 10^{-2}	3.5%	0.0%	0.5%	0.0%
73 NRG	NRG Energy Inc.	EG/UT	4.4 10^{-3}	65.9%	0.0%	3.0%	0.0%	8.9 10^{-2}	100.0%	0.0%	100.0%	0.0%	4.4 10^{-4}	66.0%	0.0%	4.1%	0.0%
74 AEP	American Electric Power	EG/UT	15.0 10^{-5}	85.7%	0.0%	57.8%	0.0%	10.4 10^{-2}	80.8%	0.0%	39.7%	0.0%	5.0 10^{-3}	57.7%	0.0%	1.9%	0.0%
75 EXC	Exelon Corp.	EG/UT	12.2 10^{-5}	83.4%	0.0%	50.8%	0.0%	10.4 10^{-2}	83.8%	0.0%	42.0%	0.0%	5.0 10^{-3}	63.2%	0.0%	2.5%	0.0%
76 SO	Southern Co.	EG/UT	10.0 10^{-6}	86.5%	0.0%	36.8%	0.0%	10.4 10^{-2}	88.2%	0.0%	42.0%	0.0%	5.0 10^{-3}	75.4%	0.0%	4.0%	0.0%
77 ETR	Entergy Corp.	EG/UT	12.4 10^{-5}	81.4%	0.0%	51.4%	0.0%	10.4 10^{-2}	78.1%	0.0%	40.3%	0.0%	5.0 10^{-2}	56.1%	0.0%	1.3%	0.0%
78 FTS.TO	Fortis	EG/UT	6.8 10^{-4}	74.8%	1.3%	17.6%	2.5%	5.1 10^{-2}	71.9%	0.0%	8.3%	0.0%	5.0 10^{-4}	66.8%	0.0%	6.2%	0.0%
79 TERP	TerraForm Power	EG/UT	-	-	-	-	-	-	-	-	-	-	1.5 10^{-6}	49.4%	0.1%	3.4%	0.1%
80 ET	Energy Transfer LP	EG/UT	3.9 10^{-4}	56.0%	0.0%	3.4%	0.0%	3.9 10^{-2}	80.7%	0.0%	13.3%	0.0%	3.9 10^{-6}	78.7%	0.0%	4.1%	0.0%
81 HUN	Huntsman Corp.	MT	4.1 10^{-2}	78.7%	0.0%	3.9%	0.0%	4.1 10^{-2}	78.8%	0.0%	5.6%	0.0%	4.1 10^{-3}	78.7%	0.0%	5.6%	0.0%
82 LYB	LyondellBasell Industries	MT	2.8 10^{-2}	31.5%	0.0%	0.6%	0.0%	2.8 10^{-2}	31.5%	0.0%	0.8%	0.0%	2.8 10^{-4}	31.5%	0.0%	0.8%	0.0%
83 STLD	Steel Dynamics Inc.	MT	6.2 10^{-6}	64.0%	0.0%	11.1%	0.0%	6.2 10^{-2}	77.3%	0.0%	16.0%	0.0%	5.0 10^{-6}	54.3%	0.0%	3.6%	0.0%
84 EXP	Eagle Materials Inc.	MT	6.9 10^{-6}	51.1%	0.0%	12.3%	0.0%	6.9 10^{-2}	58.9%	0.0%	14.6%	0.0%	5.0 10^{-6}	33.7%	0.0%	1.4%	0.0%
85 RGLD	Royal Gold Inc.	MT	10.1 10^{-6}	69.8%	0.0%	47.7%	0.0%	9.9 10^{-2}	69.4%	0.0%	44.9%	0.0%	5.0 10^{-4}	38.7%	0.0%	1.4%	0.0%
86 SMG	The Scotts Miracle-Gro Co.	MT	7.4 10^{-6}	58.7%	0.0%	19.2%	0.0%	7.4 10^{-2}	63.7%	0.0%	18.3%	0.0%	5.0 10^{-5}	39.9%	0.0%	0.9%	0.0%
87 ECL	Ecobab Inc.	MT	12.2 10^{-6}	72.3%	0.0%	42.6%	0.0%	10.4 10^{-2}	69.5%	0.0%	35.4%	0.0%	5.0 10^{-3}	37.5%	0.0%	1.6%	0.0%
88 RIO	Rio Tinto Group	MT	7.8 10^{-6}	58.8%	0.0%	16.3%	0.0%	7.8 10^{-2}	72.4%	0.0%	14.7%	0.0%	5.1 10^{-5}	44.7%	0.2%	1.8%	0.4%
89 LIN.DE	LINDE	MT	4.3 10^{-3}	83.4%	4.0%	76.5%	4.8%	0.9 10^{-3}	34.5%	0.0%	2.1%	0.0%	5.0 10^{-3}	27.1%	0.0%	1.2%	0.0%
90 KVR	Quaker Chemical Corp.	MT	10.4 10^{-6}	73.3%	0.0%	44.1%	0.0%	10.2 10^{-2}	73.2%	0.0%	43.5%	0.0%	5.0 10^{-4}	45.2%	0.0%	3.8%	0.0%

Tabelle 4.4: Auflistung der Finanzmarkt-Datensätze 46 bis 90 von 179 über die Regionen **APAC** und **MEA**, und den Sektoren **EG/UT** und **MT**. Es bezeichnen *n* die Stichprobengröße, *Präz.* die empirische Preisdaten-Präzision, *Bind.* und *NA* die Bindungsanteile. Die Renditen wurden bzgl. der Preisdaten berechnet. Quelle: Eigene Erstellung auf Basis der Daten in Yahoo Inc. (2021a), Fusion Media Ltd. (2021a) und Thomson Reuters (2021).

Nr. Code	Name	Gruppe	Yahoo! Finance			Investing.com			Refinitiv Eikon		
			n (Tsd.)	Präz. (emp.)	Preisdaten Bind. NA	n (Tsd.)	Präz. (emp.)	Preisdaten Bind. NA	n (Tsd.)	Präz. (emp.)	Preisdaten Bind. NA
91 BAC	Bank of America Corp.	FN	12.2 10 ⁻⁶	82.8% 0.0%	43.2% 0.0%	10.4 10 ⁻²	86.5% 0.0%	29.1% 0.0%	5.0 10 ⁻⁴	63.5% 0.0%	5.2% 0.0%
92 V	Visa	FN	3.4 10 ⁻⁴	14.5% 0.0%	0.4% 0.0%	3.4 10 ⁻²	31.0% 0.0%	1.6% 0.0%	3.4 10 ⁻⁴	14.5% 0.0%	0.5% 0.0%
93 MA	Mastercard Inc.	FN	3.8 10 ⁻³	9.2% 0.0%	0.4% 0.0%	3.8 10 ⁻²	26.0% 0.0%	2.1% 0.0%	3.8 10 ⁻⁴	9.2% 0.0%	0.5% 0.0%
94 DB	Deutsche Bank AG	FN	6.2 10 ⁻⁵	52.2% 0.0%	3.6% 0.0%	3.7 10 ⁻³	23.7% 0.0%	2.2% 0.0%	5.1 10 ⁻⁶	25.7% 0.0%	0.7% 0.0%
95 CS	Credit Suisse Group AG	FN	6.6 10 ⁻⁵	65.7% 0.0%	12.2% 0.0%	5.0 10 ⁻²	64.8% 0.0%	3.5% 0.0%	5.0 10 ⁻⁴	52.8% 0.2%	3.5% 0.4%
96 PYPL	PayPal Holdings Inc.	FN	1.5 10 ⁻³	16.3% 0.0%	0.5% 0.0%	1.5 10 ⁻²	16.2% 0.0%	0.9% 0.0%	1.5 10 ⁻³	16.3% 0.0%	0.9% 0.0%
97 JPM	JPMorgan Chase & Co.	FN	10.4 10 ⁻⁶	72.7% 0.0%	18.9% 0.0%	10.4 10 ⁻²	76.4% 0.0%	21.5% 0.0%	5.0 10 ⁻⁴	52.7% 0.0%	1.8% 0.0%
98 GS	The Goldman Sachs Group Inc.	FN	5.6 10 ⁻⁵	29.0% 0.0%	0.4% 0.0%	5.6 10 ⁻²	29.9% 0.0%	0.7% 0.0%	5.0 10 ⁻⁴	25.7% 0.0%	0.5% 0.0%
99 AXP	American Express Co.	FN	12.4 10 ⁻⁶	67.9% 0.0%	24.1% 0.0%	10.4 10 ⁻²	70.0% 0.0%	22.5% 0.0%	5.0 10 ⁻⁵	34.5% 0.0%	0.8% 0.0%
100 ALV.DE	Allianz SE NA.O.N	FN	6.3 10 ⁻³	44.0% 0.0%	3.6% 0.0%	3.1 10 ⁻²	42.8% 0.0%	2.1% 0.0%	—	—	—
101 BA	Boeing Co.	ID	15.0 10 ⁻⁶	72.0% 0.0%	22.9% 0.0%	10.4 10 ⁻²	63.7% 0.0%	17.8% 0.0%	5.0 10 ⁻⁴	29.3% 0.0%	0.7% 0.0%
102 LMT	Lockheed Martin Co.	ID	11.2 10 ⁻⁶	62.2% 0.0%	17.7% 0.0%	1.7 10 ⁻²	31.8% 0.0%	1.9% 0.0%	5.0 10 ⁻²	26.2% 0.0%	0.7% 0.0%
103 CAT	Caterpillar Inc.	ID	15.0 10 ⁻⁶	74.1% 0.0%	35.3% 0.0%	2.3 10 ⁻³	33.7% 0.0%	2.9% 0.0%	5.0 10 ⁻⁴	30.3% 0.0%	0.6% 0.0%
104 GE	General Electric Co.	ID	15.0 10 ⁻⁶	74.9% 0.0%	26.4% 0.0%	1.5 10 ⁻²	79.2% 0.0%	15.5% 0.0%	5.0 10 ⁻⁵	74.3% 0.0%	3.6% 0.0%
105 MMM	3M Co.	ID	13.0 10 ⁻⁶	70.8% 0.0%	32.7% 0.0%	10.4 10 ⁻²	70.0% 0.0%	15.1% 0.0%	5.0 10 ⁻³	35.5% 0.0%	1.1% 0.0%
106 SHE.DE	Siemens AG	ID	6.3 10 ⁻⁵	47.5% 0.0%	3.2% 0.0%	3.1 10 ⁻²	43.6% 0.0%	1.1% 0.0%	5.1 10 ⁻⁵	46.5% 0.0%	0.8% 0.0%
107 TKA.DE	Thyssenkrupp AG	ID	5.7 10 ⁻⁵	68.0% 0.0%	4.3% 0.0%	3.1 10 ⁻³	55.0% 0.0%	1.9% 0.0%	5.1 10 ⁻³	63.3% 0.0%	3.1% 0.0%
108 CNI	Canadian National Railway Co.	ID	6.2 10 ⁻⁶	39.7% 0.0%	3.2% 0.0%	1.5 10 ⁻²	20.3% 0.0%	0.6% 0.0%	5.0 10 ⁻⁶	29.0% 0.0%	1.2% 0.0%
109 UNP	Union Pacific	ID	10.5 10 ⁻⁶	60.9% 0.0%	15.6% 0.0%	10.4 10 ⁻²	71.4% 0.0%	22.8% 0.0%	5.0 10 ⁻⁵	22.9% 0.0%	0.7% 0.0%
110 EAF	GraflTech international Ltd.	ID	0.8 10 ⁻²	54.4% 0.0%	2.5% 0.0%	0.8 10 ⁻²	54.4% 0.0%	2.9% 0.0%	0.8 10 ⁻⁵	53.6% 0.0%	2.9% 0.0%
111 HEN3.DE	Henkel AG & Co. KGaA	CN	5.9 10 ⁻⁶	57.0% 0.0%	3.8% 0.0%	3.2 10 ⁻²	38.4% 0.0%	2.5% 0.0%	5.1 10 ⁻⁵	42.9% 0.0%	2.0% 0.0%
112 BEL.DE	Berndorf AG	CN	5.9 10 ⁻⁵	54.0% 0.0%	4.0% 0.0%	3.1 10 ⁻³	46.4% 0.0%	2.1% 0.0%	5.1 10 ⁻⁵	46.8% 0.0%	2.0% 0.0%
113 AMZN	Amazon.com Inc.	CN	6.1 10 ⁻⁶	18.4% 0.0%	0.9% 0.0%	6.1 10 ⁻²	21.5% 0.0%	1.4% 0.0%	5.0 10 ⁻⁴	13.7% 0.0%	0.6% 0.0%
114 BABA	Alibaba Group Holding	CN	1.7 10 ⁻²	10.1% 0.0%	0.2% 0.0%	1.7 10 ⁻²	10.0% 0.0%	0.3% 0.0%	1.7 10 ⁻²	10.1% 0.0%	0.3% 0.0%
115 NKE	Nike Inc.	CN	10.2 10 ⁻⁶	58.2% 0.0%	18.4% 0.0%	10.2 10 ⁻²	78.5% 0.0%	42.3% 0.0%	5.0 10 ⁻⁵	24.2% 0.0%	1.0% 0.0%
116 SBUX	Starbucks	CN	7.3 10 ⁻⁶	52.8% 0.0%	7.3% 0.0%	7.3 10 ⁻²	72.7% 0.0%	24.1% 0.0%	5.0 10 ⁻⁵	37.6% 0.0%	1.9% 0.0%
117 CON.DE	Continental AG	CN	6.3 10 ⁻⁶	41.0% 0.0%	4.0% 0.0%	3.1 10 ⁻²	33.5% 0.0%	1.4% 0.0%	5.1 10 ⁻⁶	33.0% 0.0%	1.4% 0.0%
118 ASOS	ASOS	CN	5.0 10 ⁻³	60.5% 0.3%	28.2% 0.6%	5.0 10 ⁻¹	60.2% 0.0%	25.5% 0.0%	5.0 10 ⁻³	60.2% 0.0%	27.5% 0.1%
119 MAT	Mattel Inc.	CN	11.4 10 ⁻⁶	89.9% 0.0%	41.2% 0.0%	10.4 10 ⁻²	91.0% 0.0%	36.8% 0.0%	5.0 10 ⁻⁴	76.9% 0.0%	5.8% 0.0%
120 KO	Coca-Cola Company	CN	15.0 10 ⁻⁶	88.7% 0.0%	27.7% 0.0%	10.4 10 ⁻²	88.2% 0.0%	31.5% 0.0%	5.0 10 ⁻⁵	67.1% 0.0%	3.0% 0.0%
121 MRK.DE	Merck & Co. KGaA	LS	5.9 10 ⁻⁶	47.7% 0.0%	5.0% 0.0%	3.1 10 ⁻²	47.8% 0.0%	2.5% 0.0%	5.0 10 ⁻⁵	52.5% 0.0%	1.2% 0.0%
122 FMB.DE	Fresenius AG & Co. KGaA	LS	5.9 10 ⁻⁶	48.4% 0.0%	3.3% 0.0%	3.0 10 ⁻³	38.7% 0.0%	1.0% 0.0%	5.1 10 ⁻⁶	34.4% 0.0%	2.1% 0.1%
123 BAYN.DE	Bayer AG	LS	6.3 10 ⁻⁵	49.3% 0.0%	4.3% 0.0%	3.1 10 ⁻²	42.2% 0.0%	1.3% 0.0%	5.1 10 ⁻⁶	40.9% 0.0%	1.2% 0.0%
124 BSX	Boston Scientific Corp.	LS	7.3 10 ⁻⁶	77.4% 0.0%	19.6% 0.0%	7.3 10 ⁻²	84.6% 0.0%	18.2% 0.0%	5.0 10 ⁻⁴	71.6% 0.0%	7.4% 0.0%
125 JNJ	Johnson & Johnson	LS	15.0 10 ⁻⁶	77.1% 0.0%	20.6% 0.0%	10.4 10 ⁻²	72.0% 0.0%	20.4% 0.0%	5.0 10 ⁻⁴	53.0% 0.0%	2.0% 0.0%
126 PFE	Pfizer Inc.	LS	12.4 10 ⁻⁶	86.2% 0.0%	26.9% 0.0%	10.4 10 ⁻²	90.6% 0.0%	30.7% 0.0%	5.0 10 ⁻⁵	77.9% 0.0%	3.1% 0.0%
127 NVS	Novartis AG	LS	6.2 10 ⁻⁵	61.8% 0.0%	4.3% 0.0%	5.3 10 ⁻²	63.8% 0.0%	2.6% 0.0%	5.0 10 ⁻⁵	81.9% 0.1%	13.0% 0.3%
128 BIB	Biogen Inc.	LS	7.5 10 ⁻⁶	44.7% 0.0%	12.1% 0.0%	7.5 10 ⁻²	46.9% 0.0%	14.1% 0.0%	5.0 10 ⁻⁵	25.7% 0.0%	0.6% 0.0%
129 ADPT	Adaptive Biotechnologies Corp.	LS	0.5 10 ⁻³	14.0% 0.0%	0.6% 0.0%	0.5 10 ⁻²	14.1% 0.0%	0.6% 0.0%	0.5 10 ⁻³	13.9% 0.0%	0.6% 0.0%
130 EXEL	Exelixis Inc.	LS	5.3 10 ⁻⁶	83.9% 0.0%	13.7% 0.0%	5.4 10 ⁻²	85.3% 0.0%	14.1% 0.0%	5.0 10 ⁻⁴	83.1% 0.0%	13.2% 0.0%
131 AAPL	Apple Inc.	TC	10.2 10 ⁻⁶	54.1% 0.0%	18.7% 0.0%	10.2 10 ⁻²	81.7% 0.0%	63.3% 0.0%	5.0 10 ⁻⁶	12.7% 0.0%	0.6% 0.0%
132 XRX	Xerox Holdings Corp.	TC	15.0 10 ⁻⁶	87.1% 0.0%	21.6% 0.0%	10.4 10 ⁻²	87.4% 0.0%	20.8% 0.0%	5.0 10 ⁻⁵	81.0% 0.0%	7.3% 0.0%
133 CSCO	Cisco Systems	TC	7.9 10 ⁻⁶	71.4% 0.0%	5.9% 0.0%	7.9 10 ⁻²	80.5% 0.0%	19.7% 0.0%	5.0 10 ⁻⁴	72.9% 0.0%	4.3% 0.0%
134 INTC	Intel Corp.	TC	10.4 10 ⁻⁶	76.8% 0.0%	17.7% 0.0%	10.4 10 ⁻²	84.0% 0.0%	34.1% 0.0%	—	—	—
135 ORCL	Oracle	TC	8.9 10 ⁻⁶	70.1% 0.0%	13.6% 0.0%	8.9 10 ⁻²	77.4% 0.0%	29.9% 0.0%	5.0 10 ⁻⁴	60.7% 0.0%	4.2% 0.0%

Tabelle 4.5: Auflistung der Finanzmarkt-Datensätze 91 bis 135 von 179 über die Sektoren FN, ID, CN, LS und TC. Es bezeichnen n die Stichprobengröße, Präz. die empirische Preisdaten-Präzision, Bind. und NA die Bindungsanteile. Die Renditen wurden bzgl. der Preisdaten berechnet. Quelle: Eigene Erstellung auf Basis der Daten in Yahoo Inc. (2021a), Fusion Media Ltd. (2021a) und Thomson Reuters (2021).

Nr.	Code	Name	Gruppe	Yahoo! Finance			Investing.com			Refinitiv Eikon		
				n (Tsd.)	Präz. (emp.)	Bind. NA	n (Tsd.)	Preisdaten (emp.)	Renditen Bind. NA	n (Tsd.)	Präz. (emp.)	Renditen Bind. NA
136	SAP.DE	SAP SE	TC	6.0 10 ⁻⁵	42.8%	0.0%	3.2%	0.0%	5.8 10 ⁻²	48.6%	0.0%	1.6% 0.0%
137	DTE.DE	Deutsche Telekom AG	TC	6.3 10 ⁻⁵	71.4%	0.0%	7.4%	0.0%	3.1 10 ⁻³	65.1%	0.0%	5.1 10 ⁻³
138	VZ	Verizon Communications	TC	9.5 10 ⁻⁶	73.3%	0.0%	13.7%	0.0%	9.5 10 ⁻²	87.1%	0.0%	3.8% 0.0%
139	FB	Facebook Inc.	TC	2.3 10 ⁻²	10.9%	0.0%	0.0%	0.0%	2.3 10 ⁻²	10.8%	0.0%	16.3% 0.0%
140	GOOG	Alphabet Inc.	COMM	4.3 10 ⁻⁵	3.6%	0.0%	0.0%	0.0%	1.8 10 ⁻¹	18.8%	0.0%	0.4% 0.0%
141	T.F	Crude Oil	COMM	—	—	—	—	—	9.7 10 ⁻²	68.8%	0.0%	4.0% 0.0%
142	B.F	Brent Oil	COMM	—	—	—	—	—	8.4 10 ⁻²	63.2%	0.0%	3.5% 0.0%
143	NG.F	Natural Gas	COMM	5.3 10 ⁻¹	39.5%	2.1%	5.3%	3.9%	7.9 10 ⁻³	70.9%	0.0%	2.9% 0.0%
144	GCF.F	Gold	COMM	2.8 10 ⁻²	59.9%	0.0%	3.5%	0.0%	10.6 10 ⁻²	68.3%	0.0%	4.9% 0.0%
145	ZC.F	US Corn	COMM	5.5 10 ⁻²	58.9%	2.2%	4.9%	4.0%	10.7 10 ⁻²	93.6%	0.0%	29.2% 0.0%
146	CTLF	US Cotton	COMM	1.8 10 ⁻²	52.8%	2.8%	29.3%	3.9%	3.0 10 ⁻²	49.5%	0.0%	2.0% 0.0%
147	AL.F	Aluminum	COMM	5.3 10 ⁻⁴	59.3%	2.0%	6.4%	3.8%	1.2 10 ⁻²	28.6%	0.0%	1.9% 0.0%
148	HG.F	Copper	COMM	—	—	—	—	—	8.3 10 ⁻⁴	82.3%	0.0%	15.3% 0.0%
149	ZS.F	US Soybeans	COMM	5.3 10 ⁻⁴	21.0%	2.0%	4.0%	3.7%	8.2 10 ⁻²	72.5%	0.0%	5.9% 0.0%
150	RB.F	RBOB Gasoline	COMM	—	—	—	—	—	4.3 10 ⁻⁴	19.2%	0.0%	0.2% 0.0%
151	ZW.F	US Wheat	COMM	—	—	—	—	—	8.3 10 ⁻²	83.7%	0.0%	10.7% 0.0%
152	KC.F	US Coffee	COMM	—	—	—	—	—	—	—	—	—
153	BL2c1	Milling Wheat	COMM	—	—	—	—	—	6.0 10 ⁻²	98.5%	0.0%	53.4% 0.0%
154	MZNc1	Zinc	COMM	—	—	—	—	—	3.3 10 ⁻²	45.9%	0.0%	—
155	LEAD.F	Lead	COMM	—	—	—	—	—	3.2 10 ⁻²	55.6%	0.0%	1.4% 0.0%
156	LE.F	Live Cattle	COMM	4.9 10 ⁻³	76.0%	2.2%	8.7%	3.8%	10.5 10 ⁻³	88.5%	0.0%	9.0% 0.0%
157	HO.F	Heating Oil	COMM	5.3 10 ⁻⁴	22.8%	2.0%	4.1%	3.8%	10.5 10 ⁻⁴	46.1%	0.0%	1.3% 0.0%
158	PLF	Platinum	COMM	5.5 10 ⁻¹	44.4%	14.8%	17.4%	16.9%	3.7 10 ⁻²	20.1%	0.0%	0.4% 0.0%
159	LGO	Gas Oil	COMM	—	—	—	—	—	3.7 10 ⁻²	71.8%	0.0%	8.4% 0.0%
160	EURUSD.X	EUR / USD	FX	4.6 10 ⁻⁶	32.7%	1.1%	2.0%	1.4%	10.8 10 ⁻⁴	78.7%	0.0%	2.6% 0.0%
161	CADUSD.X	CAD / USD	FX	4.7 10 ⁻⁶	38.5%	0.7%	1.6%	1.1%	3.2 10 ⁻³	55.6%	0.0%	1.0% 0.0%
162	JPYUSD.X	JPY / USD	FX	6.4 10 ⁻⁶	76.8%	1.0%	4.2%	1.5%	10.5 10 ⁻⁴	86.2%	0.0%	5.4% 0.0%
163	RUBUSD.X	RUB / USD	FX	6.6 10 ⁻⁶	34.9%	3.4%	9.0%	4.0%	6.7 10 ⁻⁵	60.3%	0.0%	2.7% 0.0%
164	CNYUSD.X	CNY / USD	FX	5.2 10 ⁻⁶	52.6%	5.4%	25.5%	6.8%	7.3 10 ⁻⁴	93.6%	0.0%	5.2 10 ⁻²
165	GBPEUR.X	GBP / EUR	FX	4.7 10 ⁻⁵	52.6%	1.1%	2.5%	1.4%	4.8 10 ⁻⁴	98.4%	0.0%	76.9% 0.0%
166	SEKEUR.X	SEK / EUR	FX	3.3 10 ⁻⁶	53.2%	0.8%	2.7%	1.0%	9.2 10 ⁻⁴	79.9%	0.0%	2.1% 0.0%
167	CHFEUR.X	CHF / EUR	FX	3.3 10 ⁻⁶	34.6%	1.1%	3.1%	1.4%	8.3 10 ⁻⁴	99.5%	0.0%	75.6% 0.0%
168	PLN/EUR	PLN / EUR	FX	4.4 10 ⁻⁶	59.2%	1.1%	2.5%	1.4%	9.0 10 ⁻⁵	73.6%	0.0%	26.9% 0.0%
169	EURRUB.X	EUR / RUB	FX	4.6 10 ⁻⁵	16.2%	11.2%	12.5%	11.5%	7.0 10 ⁻⁴	22.3%	0.0%	0.6% 0.0%
170	BTCUSD.X	Bitcoin	CX	2.5 10 ⁻⁴	0.4%	0.2%	0.4%	0.3%	3.4 10 ⁻¹	21.0%	0.0%	7.7% 0.0%
171	ETHUSD.X	Ethereum	CX	1.5 10 ⁻²	57.2%	0.0%	4.7%	0.0%	2.0 10 ⁻²	10.3%	0.0%	0.9% 0.0%
172	USDTUSD.X	Tether	CX	—	—	—	—	—	—	—	—	1.0 10 ⁻²
173	XRPUSD.X	Ripple	CX	—	—	—	—	—	2.4 10 ⁻⁵	30.5%	0.0%	5.5% 0.0%
174	LINKUSD.X	Chainlink	CX	1.4 10 ⁻⁶	0.3%	0.3%	0.5%	0.5%	1.3 10 ⁻²	63.8%	0.0%	32.7% 0.0%
175	BCHUSD.X	Bitcoin Cash	CX	—	—	—	—	—	0.9 10 ⁻¹	43.1%	0.0%	2.3% 0.0%
176	LTCTUSD.X	Litecoin	CX	2.5 10 ⁻⁵	0.4%	0.2%	0.3%	0.3%	1.8 10 ⁻³	12.2%	0.0%	3.0% 0.0%
177	BNBUSD.X	Binance Coin	CX	—	—	—	—	—	—	—	—	—
178	PDOTUSD.X	Polkadot	CX	1.4 10 ⁻⁶	0.9%	0.3%	0.5%	0.5%	0.2 10 ⁻³	0.0%	0.0%	0.2% 0.0%
179	ADAUSD.X	Cardano	CX	—	—	—	—	—	1.3 10 ⁻⁶	1.1%	0.0%	—

Tabelle 4.6: Auflistung der Finanzmarkt-Datensätze 136 bis 179 von 179 über den Sektor *TC* sowie über die sonstigen Gruppen *COMM*, *FX* und *CX*. Es bezeichnen *n* die Stichprobengröße, *Präz.* die empirische Preisdaten-Präzision, *Bind.* und *NA* die Bindungsanteile. Die Renditen wurden bzgl. der Preisdaten berechnet. Quelle: Eigene Erstellung auf Basis der Daten in Yahoo Inc. (2021a), Fusion Media Ltd. (2021a) und Thomson Reuters (2021).

4.4 Auswertung der empirischen Daten

Die in den Tabellen 4.3 bis 4.6 präsentierten Finanzinstrumente und deren Datensätze wurden je Instrument und Quelle einzeln auf Bindungen untersucht. In diesem Abschnitt werden diese Informationen unter verschiedenen Gesichtspunkten analysiert. Insbesondere werden dabei die Daten mit Bezug auf die Quellen und Zeiträume gefiltert. Das Ziel ist der Aufbau des Verständnisses über das Auftreten und Eigenschaften von Bindungen in Finanzmarktdaten.

Allgemeine Auswertung der empirischen Daten

Zu Anfang stellt Tabelle 4.7 einen Überblick über alle Datensätze dar. Dort werden Minima und Maxima, Quartile, Mittelwerte und Standardabweichungen zu folgenden Informationen aufbereitet: Stichprobenumfänge, Präzision über die gemessenen Dezimalstellen, und Bindungsanteile für Preise und Renditen. Zusätzlich wurden die Bindungsanteile unter der Bedingung von auftretenden Bindungen bemessen.

Kennzahl	Allgemein						Gebundene Stichproben			
	n	Dez.	Preisdaten		Renditen		Preisdaten		Renditen	
			Bind.	NA	Bind.	NA	Bind.	NA	Bind.	NA
Minimum	158	1.0	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%
1. Quartil	4 134	2.0	5.5%	0.0%	0.5%	0.0%	5.8%	0.2%	0.6%	0.3%
Median	5 029	2.0	34.5%	0.0%	2.3%	0.0%	34.9%	1.1%	2.5%	1.4%
3. Quartil	7 301	4.0	64.8%	0.0%	8.3%	0.0%	65.1%	2.5%	8.9%	4.0%
Maximum	14 986	6.0	100%	14.8%	100%	16.9%	100%	14.8%	100%	16.9%
Mittelwert	5 710	3.2	37.5%	0.3%	8.5%	0.5%	37.8%	1.8%	8.8%	2.7%
Std.-Abw.	2 984	1.6	30.7%	1.2%	14.9%	1.7%	30.6%	2.4%	15.0%	3.1%

Tabelle 4.7: Auswertung der Datensätze bzgl. der gesamten Datengrundlage (433 Stichproben) und eingeschränkt auf Stichproben mit Bindungen in Preis- oder Renditedaten (429 Stichproben). *n* ist die Stichprobengröße, *Dez.* die Anzahl an Dezimalstellen, und *Bind.* und *NA* bezeichnen die Bindungsanteile. Quelle: Eigene Erstellung.

Bei angenommenen 250 Handelstagen pro Jahr beinhalten die Stichproben im Mittel Daten über 22.8 Jahre. Der Interquartilsabstand (IQR) umfasst 16.5 bis 29.2 Jahre. Die größte Stichprobe beinhaltet Daten über 59.9 Jahre.

Die insgesamt bemessene Präzision liegt zwischen 10^{-1} und 10^{-6} je Datensatz, wobei im Mittel 3.2 Dezimalstellen vorliegen.

Bei den Bindungsanteilen ist auffällig, dass diese zwischen den Preis- und Renditestichproben stark abweichen. Bei den über die Preise berechneten

Renditen ist der Bindungsanteil im Mittel das 0.22-Fache des Bindungsanteils der ursprünglichen Preisdaten. Der NA-Anteil steigt im Mittel auf das 1.73-Fache und Maximal auf das Doppelte des NA-Anteils der Preisdaten – dies bestätigt die Folgerungen aus Formel (4–1).

Sowohl Preis- als auch Renditedaten können dabei zu 100 % gebunden sein, wobei Bindungen durchschnittlich 37.5 % und 8.5 % der Stichproben ausmachen. Die berechneten Mediane lauten 34.5 % und 2.3 %, und der IQR ist das Intervall von 5.5 % bis 64.8 % respektiv von 0.5 % bis 8.3 %.

Zum Ende vom Abschnitt 4.2 wurden bereits für einige der Parameter im Algorithmus 4.2.3 Werte angenommen, da diese aus dem Sachzusammenhang gefolgert werden können. Für die noch fehlenden Werte zur Messpräzision und zur Modellierung fehlerhaft übertragener Werte können aus der empirischen Messpräzision und der NA-Anteile in Tabelle 4.7 erste Kandidaten für die Werte bzw. Wertebereiche herangezogen werden.

Die empirische Präzision der Preisdaten liegt zwischen 10^{-1} und 10^{-6} , wobei der IQR Präzisionen von 10^{-2} bis 10^{-4} umfasst. Die Median- und Most-Likely-Präzision beträgt 10^{-2} ; diese liegt 48 % aller Datensätze zugrunde. Der zweithöchste Anteil von 15 % besitzt die Präzision 10^{-6} . Aufgrund des hohen Anteils ist 10^{-2} zu bevorzugen, wobei unter geeigneten Annahmen ebenfalls die anderen Werte plausibel sein können. Hierzu sind die detaillierteren Auswertungen zu betrachten. Für Renditedaten ist aufgrund der Möglichkeit der direkten Berechnung eine beliebig hohe Präzision zulässig. Dies wird in folgender Annahme zusammengefasst:

4.4.1 Annahme (Messpräzision für den datengenerierenden Algorithmus).
Für den Parameter $r = (r_1, \dots, r_d)$ im Algorithmus 4.2.3 sind

$$r_j \in \{10^{-2}, 10^{-6}, 0\}$$

geeignete Kandidaten zur Modellierung von Finanzmarktrahmenbedingungen. Die Wahl $r_j = 0$ ist als eine beliebig hohe Messpräzision $10^{-\infty} = 0$ zu interpretieren; diese wird hinzugenommen, um den Idealfall abzudecken.

In 81 % der Stichproben sind keine NA-Bindungen bzw. NA-Werte vorliegend. Somit ist ein erster Kandidat für die Wahrscheinlichkeit einer fehlerhaften Datenübertragung der Wert Null. Die NA-Anteile in gebundenen Stichproben liegen im Mittel bei 1.8 % und 2.7 % und maximal bei 14.8 % und 16.9 % für die Preis- bzw. Renditedaten. Die IQR umspannen die Intervalle [0.2 %, 2.5 %] und [0.3 %, 4.0 %].

Um eine möglichst konservative Annahme mit Hinblick auf die Simulationsstudie durchzuführen, wird hier die Obergrenze des Preisdaten-IQR von 2.5 % gewählt. Da der NA-Anteil der Renditen das Doppelte des NA-Anteils der Preisdaten betragen kann, entspricht dies einem Wert von 5.0 %.

4.4.2 Annahme (Wahrscheinlichkeit fehlerhafter Datenübertragungen und der zugehörige Standardwert für den datengenerierenden Algorithmus). Für den Parameter $p = (p_1, \dots, p_d)$ im Algorithmus 4.2.3 sind

$$p_j \in \{0.0\%, 2.5\%, 5.0\%\}$$

geeignete Kandidaten zur Modellierung von Finanzmarktrahmenbedingungen.

In Bezug auf die Modellierungskomplexität wird bei der Modellierung der Auslassung ein Standardwert $x_{\text{default}} = 0$ betrachtet.

Quellenbezogene Auswertung der empirischen Daten

In Abschnitt 4.1 wurden die Unterschiede zwischen den Datenanbietern und deren Bezugsquellen verdeutlicht. Dementsprechend ist neben einer quellennahmigen Betrachtung aus Tabelle 4.7 auch eine Analyse je Quelle notwendig. Die Ergebnisse bzgl. Präzision und Bindungen je Quelle sind in Tabelle 4.8 dargestellt. Es wurden die beiden häufigsten Präzisionen der einzelnen Quellen (Top-2-Dec.) ergänzt, um die Annahme 4.4.1 zu validieren.

Kennzahl	Yahoo! Finance			Investing.com			Refinitiv Eikon		
	Dez.	Preis Bind.	Rend. Bind.	Dez.	Preis Bind.	Rend. Bind.	Dez.	Preis Bind.	Rend. Bind.
Minimum	1.0	0.3%	0.0%	1.0	0.0%	0.0%	1.0	0.0%	0.0%
1. Quartil	2.0	9.2%	2.7%	2.0	5.8%	0.4%	2.0	3.1%	0.2%
Median	5.0	45.1%	4.5%	2.0	38.4%	2.1%	4.0	27.1%	0.8%
2. Quartil	6.0	65.9%	15.6%	2.0	72.4%	15.3%	5.0	53.3%	3.0%
Maximum	6.0	89.9%	76.5%	6.0	100.0%	100.0%	6.0	99.7%	95.7%
Mittelwert	4.1	40.4%	10.9%	2.2	41.1%	10.8%	3.6	30.5%	3.6%
Std.-Abw.	1.8	29.4%	14.0%	0.8	33.4%	17.6%	1.4	27.4%	10.2%
Top-2-Dec. und Anteil	6.0	(39%-Anteil)		2.0	(79%-Anteil)		2.0	(30%-Anteil)	
	2.0	(29%-Anteil)		3.0	(8%-Anteil)		4.0	(27%-Anteil)	

Tabelle 4.8: Quellenbasierte Auswertung der Datensätze auf *Yahoo! Finance* (129 Stichproben), *Investing.com* (165 Stichproben) und *Refinitiv Eikon* (139 Stichproben). n ist die Stichprobengröße, Dez. die Anzahl an Dezimalstellen, und Bind. bezeichnen die Bindungsanteile. Quelle: Eigene Erstellung.

In den Daten aus Tabelle 4.8 werden die Auswirkungen der Unterschiede in Datenbeschaffung und Überprüfung durch die jeweiligen Quellen deutlich: Dass sowohl *Yahoo! Finance* als auch *Refinitiv Eikon* die Daten von ausschließlich geprüften Anbietern beziehen, spiegelt sich in einer im Mittel und IQR höheren Präzision wieder – der Vergleich wird mit der Quelle

Investing.com, die auch Daten der Marktteilnehmer nutzt, gezogen. Erstere genannte Quellen haben eine mittlere Präzision von 4.1 und 3.6 Dezimalstellen. Diese ist damit weitaus höher als die durchschnittlich 2.2 Dezimalstellen bei *Investing.com*. Ähnliche Resultate sind beim Median und IQR zu sehen. Da bei jeder der Quellen die Präzision 10^{-2} bzw. eine Präzision über 2.0 Dezimalstellen zu den Top-2 bzgl. des Anteils gehört, aber eine Präzision von 10^{-6} nur bei *Yahoo! Finance* einen größeren Anteil über 39 % innehaltet (*Investing.com*: 1 %, *Refinitiv Eikon*: 11 %), ist die Wahl von $r_j = 10^{-2}$ als Wert zur Annahme 4.4.1 angebracht.

Ferner sind über die Kennzahlen unterschiedliche Bindungsanteile bei den Quellen erkennbar. Bei den bereitgestellten Preisdaten sind sowohl im Mittel und Median als auch bzgl. des IQR erkennbar, dass *Refinitiv Eikon* geringere Bindungsanteile als *Yahoo! Finance* und diese beide geringere Anteile als *Investing.com* besitzen. Lediglich Minima und Maxima sind in derselben Größenordnung. Bei den berechneten Renditen ist erneut *Refinitiv Eikon* die Quelle mit dem niedrigsten Bindungsanteil in den vorig genannten Statistiken; *Yahoo! Finance* besitzt hier den höchsten Bindungsanteil.

Zeitraumbezogene Auswertung der empirischen Daten

Neben einer umfassenden und einer quellspezifischen Analyse wird folgend der zeitliche Verlauf analysiert. Da sich Finanzmärkte und die Rahmenbedingungen stetig verändern, können über Visualisierungen mögliche Tendenzen für zukünftige Entwicklungen abgeleitet sowie vergangene und weniger relevante Werte abgegrenzt werden.

In Abbildung 4.1 wurden die Kennzahlen aus den vorigen Tabellen über sechs Betrachtungsräume von zehn bzw. elf Jahren aufbereitet. Zur Veranschaulichung der Veränderungen wurden jeweils der Mittelwert plusminus der Standardabweichung, und der Median mit dem IQR eingezzeichnet.

Bei der empirischen Präzision ist festzustellen, dass in beiden betrachteten Statistiken eine Abnahmetendenz vorliegt. Die Median-Präzision fällt zwischen 1960 und 1979 von 10^{-6} auf 10^{-2} und bleibt ab 1980 bis 2021 konstant bei 10^{-2} . Der Mittelwert fällt über den gesamten Betrachtungszeitraum von 10^{-5} in 1960 bis 1969 auf den Wert $10^{2.78}$ in 2010 bis 2021. Dies spricht für die Wahl des Messpräzisionsparameters $r_j = 10^{-2}$ in der Annahme 4.4.1. Da im Zeitraum bis 1989 der IQR die Präzision 10^{-6} einschließt, ist ebenso die Wahl von $r_j = 10^{-6}$ begründbar. Eine Präzision von 10^{-4} gehört ab dem Zeitraum 1970 bis 1979 zum IQR und zum Standardabweichungskoeffizienten um den Mittelwert, weswegen dieser Wert plausibel ist. Die Menge $\{10^{-2}, 10^{-4}, 10^{-6}\}$ deckt damit die unterschiedlichen Präzisionen ab.

Der Bindungsanteil und der NA-Anteil sind jeweils für Preise und Rendi-

ten getrennt dargestellt: Beim Bindungs- bzw. NA-Anteil sind von 1960 bis 1989 stabile Plateaus im oberen bzw. unteren Drittel vorzufinden. Zwischen 1990 und 2021 fallen die Anteile der Preis- und Renditedaten im Mittel (Median) von 74 % (93 %) und 39 % (31 %) auf 29 % (24 %) und 3 % (1 %).

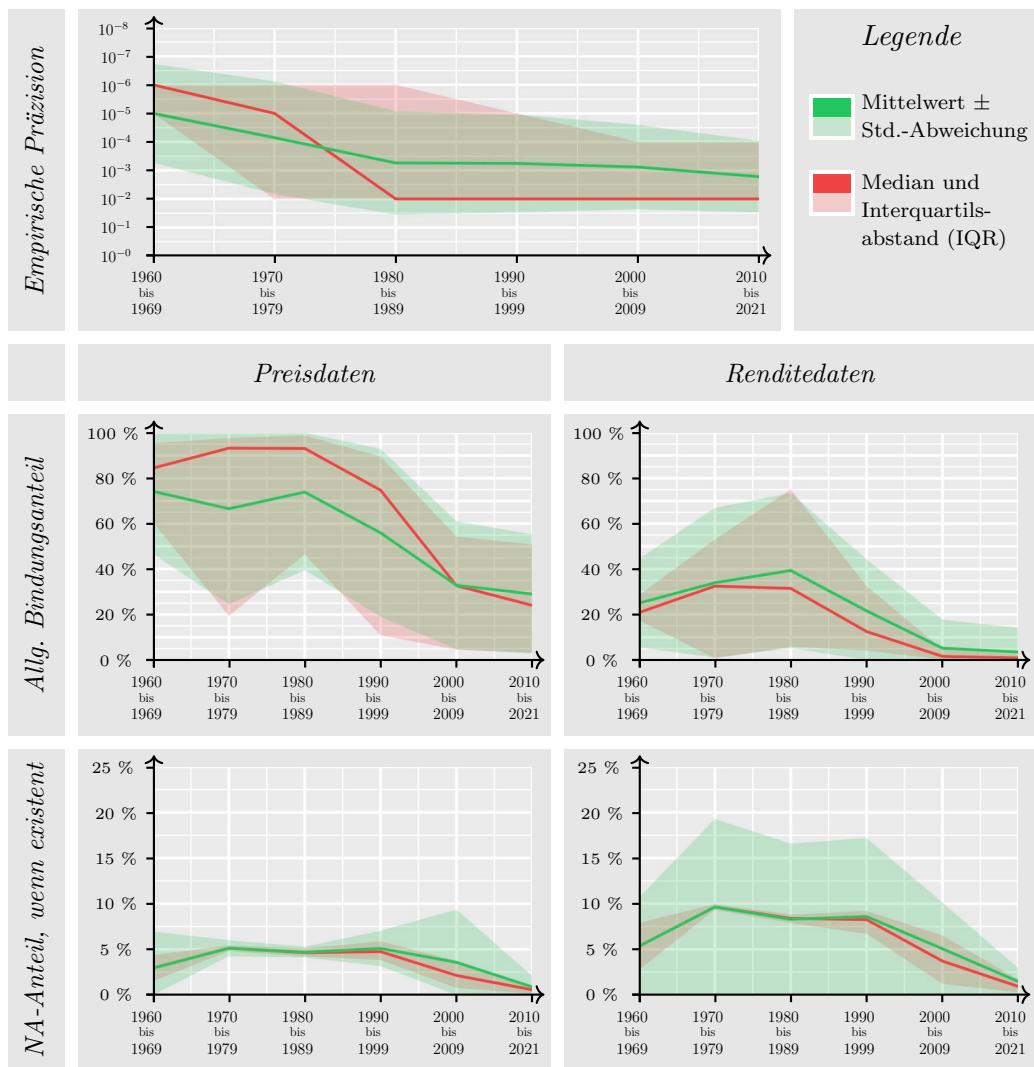


Abbildung 4.1: Visuelle Aufbereitung der Präzision und Bindungsanteile in Preis- und Renditedaten im zeitlichen Verlauf. Für die Zehnjahreszeiträume zwischen 1960 bis 2010 und dem Zeitraum 2010 bis 2021 wurden jeweils Präzisionen und Bindungsanteile als Mittelwert (grün) und Median (rot) zusammengefasst. Zum Mittelwert wurde die Standardabweichung (hellgrün) angegeben; der Median wird vom Interquartilsabstand (hellrot) begleitet. Quelle: Eigene Erstellung.

Bzgl. der NA-Bindungen gelten folgende weitere Beobachtungen: Der Mittelwert und Median der Preisdaten verzeichnen zu Beginn einen leichten An-

stieg von 3 % zu 5 % zwischen 1960 bis 1969 und 1970 bis 1979. Ab 1980 verbleibt der NA-Anteil auf diesem Niveau und fällt zwischen 1990 und 2021 auf unter 1 % ab. Die Kurven zu den NA-Bindungen in den berechneten Renditedaten spiegeln die Kurven zu den Preisdaten wider. Hierbei sind die Werte doppelt so hoch und besitzen eine höhere Standardabweichung. Diese Werte unterscheiden sich von den Erkenntnissen zu Tabelle 4.7, wobei aus einer zeitlichen Betrachtung eher die Werte

$$p_j \in \{0.0\%, 5.0\%, 10.0\%\}$$

als Kandidaten für die Modellierung der Wahrscheinlichkeit fehlerhafter Datenübertragungen in Frage kämen. Da bei beiden Herleitungen die Werte 0.0 % und 5.0 % enthalten sind, spricht dies nach dem Ausschlussprinzip für die Wahl dieser beiden Werte. Die Kurvenverläufe zu den Bindungen legen nahe, dass der Bindungsanteil von Finanzmarktdaten in der Tendenz fallend ist und gleichzeitig NA-Werte als Folge von z. B. fehlerhafter Datenübertragungen seltener auftreten. Die Annahmen 4.4.1 mit $r_j \in \{10^{-2}, 10^{-4}, 10^{-6}, 0\}$, und 4.4.2 mit $p_j \in \{0.0\%, 2.5\%, 5.0\%, 10.0\%\}$ unter Ergänzung der Werte sind zu bestätigen. Dies wird mit folgender Zusammenfassung festgehalten.

4.4.3 Annahme (Parameterannahmen für die Datengenerierung). *Für die Parameter $r = (r_1, \dots, r_d)$ und $p = (p_1, \dots, p_d)$ im Algorithmus 4.2.3 sind*

$$r_j \in \{10^{-2}, 10^{-4}, 10^{-6}, 0\}$$

und

$$p_j \in \{0.0\%, 2.5\%, 5.0\%, 10.0\%\}$$

geeignete Kandidaten zur Modellierung von Finanzmarktrahmenbedingungen. In Bezug auf die Modellierungskomplexität wird bei der Modellierung der Auslassung ein Standardwert $x_{\text{default}} = 0$ betrachtet.

Auswahl von repräsentativen Datensätzen

Im datengenerierenden Algorithmus werden Marginale zur Transformation der Daten berücksichtigt; bzgl. der empirischen Verteilungsfunktionen von ausgewählten Finanzmarktdaten werden die Bindungsstrukturen in die synthetischen Datensätze übertragen. Es werden drei Arten von Datensätzen bzgl. der Renditen berücksichtigt; es wird zwischen *Worst Case*, *Average Case* und *Best Case* Datensätzen unterschieden. Die Auswahl wurde dabei auf Basis der Bindungs- und NA-Anteile durchgeführt, wobei Worst Case bzw. Best Case Datensätze gleichzeitig möglichst hohe bzw. möglichst niedrige Bindungs- und NA-Anteile besitzen sollen; der Average Case betrachtet

Datensätze, deren Anteile nahe an den Medianen sind. Folgende Datensätze wurden ausgewählt:

<i>Bezeichnung</i>	<i>Nr.</i>	<i>Name</i>	<i>n</i>	<i>Dez.</i>	<i>Bind.</i>	<i>NA</i>
Worst Case 1	35	Bolsa d. Val. Guayaquil (EI)	787	2	95.68%	1.02%
Worst Case 2	89	LINDE (YF)	4 308	3	76.49%	4.76%
Worst Case 3	164	CNY / USD (YF)	5 234	6	25.54%	6.82%
Average Case 1	71	E.ON SE (YF)	5 505	6	2.56%	0.00%
Average Case 2	102	LMT (IN)	1 720	2	1.92%	0.00%
Average Case 3	160	EUR / USD (IN)	10 776	4	2.63%	0.00%
Best Case 1	2	Deutscher Aktienindex (EI)	5 067	2	0.00%	0.00%
Best Case 2	22	Nasdaq Composite (EI)	5 023	4	0.00%	0.00%
Best Case 3	48	Jakarta Composite (EI)	4 854	3	0.00%	0.00%

Tabelle 4.9: Auswahl repräsentativer Datensätze. *n* ist die Stichprobengröße und *Dez.* die Anzahl an Dezimalstellen. *Bind.* und *NA* sind die Bindungsanteile in den Renditen. YF: Yahoo! Finance, IN: Investing.com, und EI: Refinitiv Eikon. Quelle: Eigene Erstellung.

Aus den neun Datensätzen in Tabelle 4.9 werden in der Simulationsstudie sechs Paarungen *Worst* \times *Worst*, *Average* \times *Average*, *Best* \times *Best*, *Worst* \times *Average*, *Worst* \times *Best*, und *Average* \times *Best* berücksichtigt, wobei die Datensätze in jeder Iteration zufällig ausgewählt werden sollen. Auf Basis der Datensätze werden empirische Verteilungs- respektive Quantilfunktionen konstruiert, die dann als Marginale im datengenerierenden Algorithmus 4.2.3 berücksichtigt werden.

Kapitel 5

Methoden zur Handhabe von Bindungen

Die Tests und Methoden für Copulas im Kapitel 3 wurden unter der Annahme der Bindungsfreiheit formuliert. Wie in den Abschnitten 4.1 und 4.3 aufgezeigt, ist das Vorliegen von Bindungen in empirischen (Finanzmarkt-)Daten nicht auszuschließen, da in der Mehrheit der Finanzmarktdatensätze Bindungen vorliegen, sodass die Annahmen verletzt werden. Um die Copula-Methoden trotz dessen anwenden zu können, sind Ansätze zur Handhabung von Bindungen oder zur Mitigation der Auswirkungen notwendig.

Mögliche Ansätze werden in den folgenden Abschnitten vorgestellt. Hierzu gehören die *Naiven Methoden* in Abschnitt 5.1, die keine zusätzlichen Informationen neben der Bindungsstruktur benötigen, und die *Randomisierungsmethoden* in Abschnitt 5.2, die die Bindungen über zufällige Anpassungen der (Pseudo-)Stichproben lösen. Im Abschnitt 5.3 werden bei der *Intervall-Zensur* die möglichen Werte der unbekannten Parameter als Intervall berücksichtigt und die *Bootstrap-Methoden* im Abschnitt 5.4 sind alternative robuste Formulierungen der statistischen Tests.

5.1 Naive Methoden

Die naheliegendsten Ansätze der Handhabe von Bindungen sind die naiven Methoden *Keine Anpassung der Daten oder Methoden* oder die *Stichprobenreduktion*. Bei diesen Methoden werden keine zusätzlichen Annahmen an Rangvergaben getroffen und die Bindungen in den (Pseudo-)Stichproben nicht über Randomisierungen gelöst. Es werden somit bei der Stichprobenreduktion oder einer Betrachtung der ursprünglichen Daten keine zusätzlichen und möglicherweise falsche Informationen in die Daten eingebunden.

Keine Anpassung der Daten oder Methoden

Eine Nichtanpassung der Methode beim Auftreten von Bindungen kann unter Umständen sinnvoll sein: Eine Präferenz der unmodifizierten Methode gegenüber einer modifizierten Methode zur Anwendung auf gebundenen Daten kann beispielsweise ausgesprochen werden, wenn diese gleiche oder höhere Güte hat. Auch im Fall, dass die Performance der anderen Modifikationen abhängig von den vorliegenden Strukturen oder Prozesse sind, kann eine Nichtanpassung motiviert sein. Dies gilt, falls die Performance gleich oder besser als die mittlere Performance der anderen Methoden ist. Als Beispiel hierzu ist die Studie Genest et al. (2011) zu nennen: Dort wurde in einer Simulationsstudie beobachtet, dass eine zufällige Rangvergabe auf gebundenen Datensätzen bei zwei rangbasierten Tests (Extremale Abhängigkeit, Anpassungsgüte) den Effekt von Bindungen nicht genügend mitigieren kann (vgl. Genest et al. (2011): S. 144, Tabellen 2 und 4). Ferner ist zu nennen, dass das Auftreten von Bindungen in der Praxis hingenommen wird und die Methoden auf gebundenen Stichproben trotz möglicher Annahmeverletzungen angewendet werden (vgl. Hájek (1999): S. 136).

Stichprobenreduktion

Alternativ kann eine Löschung der gebundenen Elemente exklusiv je einer Repräsentanten durchgeführt werden. Hierbei gehen Informationen verloren, so dass genügend große Datensätze und möglichst niedrige Bindungsanteile notwendig sind, um eine Stichprobenreduktion bei Bindungen zu begründen (vgl. Hájek (1999): S. 135-136). Stichprobenreduktionen wurden in Genest et al. (2011) oder Bücher und Kojadinovic (2016) als Ansätze zur Handhaben von Bindungen genannt; dort wird diese Methode aufgrund des verzerrenden Effekts auf Tests der Extreme Value Dependence und Tests der Anpassungsgüte nicht empfohlen (vgl. Genest et al. (2011): S. 142-144).

Bei Stichproben über zwei oder mehr Dimensionen sind die in Definition 3.1.1 eingeführten vollständigen und teilweisen Bindungen unterschiedlich zu handhaben. Je nach Struktur der Bindungen kann die Stichprobe zu stark verkleinert werden, da bei teilweisen Bindungen nicht ausschließlich die gebundene Koordinate, sondern der gesamte Eintrag gelöscht werden müsste. Darauf hinaus stellt sich die Wahl eines Repräsentanten für die ungebundenen Koordinaten eines teilweisen gebundenen Elements als schwierig dar, da zum Beispiel ein Element mehreren Bindungen zugehörig sein kann.

In dieser Arbeit werden zwei mögliche Ansätze zur Stichprobenreduktion berücksichtigt. Der erste der beiden Ansätze wird im folgenden Verfahren dargelegt. Im Anschluss wird der zweite Ansatz vorgestellt.

5.1.1 Algorithmus (Stichprobenreduktion bzgl. vollständiger Bindungen).

Sei x_1, \dots, x_n eine Stichprobe mit Koordinaten der Form $x_i = (x_{i1}, \dots, x_{id})$.

Zur Stichprobenreduktion bzgl. der vollständigen Bindungen werden die folgenden Schritte durchgeführt:

1. Initialisierung: Es bezeichnen $I^{(k)}$ die Indexmengen der Einträge, die in der k -ten Iteration zur Verfügung stehen und $J^{(k)}$ die Indexmengen der Bindungen und einzigartigen Elemente. Initialisiere $I^{(0)} := \{1, \dots, n\}$.
2. Identifizierung der Bindungen: Für alle $k \in \{1, \dots, n\}$ werden die folgenden Schritte zur Identifizierung der Bindungen durchgeführt.

(i.) Bindungen vom Wert x_k : Setze die Indexmenge $\emptyset \subseteq J^{(k)} \subseteq I^{(k-1)}$ der zu x_k gebundenen und noch nicht verarbeiteten Elemente als

$$J^{(k)} := \{i \in I^{(k-1)} : x_k = x_i\}.$$

(ii.) Indexmenge der noch zu verarbeitenden Elemente: Damit Elemente nicht mehrfach verarbeitet werden, setze die Indexmenge

$$I^{(k)} := I^{(k-1)} \setminus J^{(k)}.$$

Ist $I^{(k)} = \emptyset$, so müssen die Iterationen $k+1, \dots, n$ nicht durchgeführt werden, da $I^{(k)} = I^{(k+1)} = \dots = I^{(n)} = \emptyset$ gilt.

3. Durchführung der Stichprobenreduktion: Für alle $k \in \{1, \dots, n\}$ mit $I^{(k)} \neq \emptyset$ wird ein Repräsentant $i^{(k)} \in I^{(k)}$ ausgewählt. Die Indexmenge für die verbleibenden Elemente nach der Stichprobenreduktion lautet

$$I := \bigcup_k \{i^{(k)}\}.$$

Nach Durchlauf der obigen Schritte wird die bzgl. vollständiger Bindungen reduzierte Stichprobe $(x_i)_{i \in I}$ ausgegeben.

Da Elemente einer Stichprobe mehreren teilweisen Bindungen zugehören können, wäre der Algorithmus für ebendiesen Fall komplexer zu gestalten. In diesem Fall können sich die teilweisen Bindungen überlagern und es sind die Elemente, die als Repräsentant einer Bindung ausgewählt wurden, in der Stichprobe beizubehalten.

Entsprechend dieser Komplexität wird eine alternative Methode zur Stichprobenreduktion gesucht. Um eine Heuristik zur Auswahl der zu löschenen gebundenen Elemente in einer Stichprobe der Größe n und mit d Dimensionen zu entwickeln, sind folgende Entscheidungsregeln zur Auswahl der Elemente naheliegend:

- ① Begrenzung der Stichprobenreduktion: *Die Anzahl der zu löschenen Elemente soll den festgelegten Wert $n^* \in \{1, \dots, n\}$ nicht überschreiten.*
- ② Anzahl der gebundenen Koordinaten als Nutzen: *Bindungen in b Koordinaten sind bei der Lösung den Bindungen in $(b-1)$ oder weniger Koordinaten vorzuziehen, wobei $b \in \{2, \dots, d\}$ sei.*
- ③ Größe der Bindung als Kosten: *Da die Performance der Methoden im Umfang der Stichprobenreduktion abnimmt, sind Bindungen der Größe g gegenüber den Größen $(g+1)$ oder höher zu bevorzugen.*
- ④ Repräsentanten je Bindung: *Ist ein Element als repräsentatives Element einer Bindung ausgewählt worden, so soll es bzgl. weiterer Bindungen nicht entfernt werden.*

Diese Strukturierung kann beispielsweise in Form von *Greedy-Algorithmen* aus der *kombinatorischen Optimierung* umgesetzt werden. Für eine heuristische Lösungsfindung wird über die Anzahlen der gebundenen Koordinaten und der Bindungsgrößen aus ② und ③ eine Gewichtung hergeleitet und darüber optimiert.

Bei einem Greedy-Algorithmus werden die Elemente nach absteigender Gewichtung sortiert und unter angegebenen Einschränkungen, wie z. B. einer maximalen Anzahl wie in ① oder einer Löschbarkeitseinschränkung ④, und der Sortierung ausgewählt, sofern zulässig (vgl. Gritzmann (2013): S. 91-92).

5.1.2 Algorithmus (Stichprobenreduktion über einen Greedy-Algorithmus bzgl. Koordinatenanzahl und Bindungsgrößen, bivariater Fall). *Sei x_1, \dots, x_n eine Stichprobe mit Elementen der Form $x_i = (x_{i1}, x_{i2})$ und bezeichne $b_i \in \{0, 1, 2\}$ die Anzahl der gebundenen Koordinaten und $g_i \in \{1, \dots, n\}$ die maximale Größe von ebendiesen Bindungen von x_i .*

Sei $w_i := w(b_i, g_i)$ die Nutzen-Kosten-Gewichtung zu x_i , die über eine Funktion $w : \mathbb{N}_0 \times \mathbb{N} \rightarrow \mathbb{R}_{\geq}$ wie zum Beispiel $w(b, g) := b/g$ berechnet wird.

Weiter bezeichne $N \in \{1, \dots, n\}$ die maximale Anzahl der zu löschenen Elemente innerhalb der Stichprobe, wobei N fixiert oder als $N = N(n)$ relativ zum Stichprobenumfang $n \in \mathbb{N}$ definiert werden kann.

Der Greedy-Algorithmus zur Auswahl der Elemente wird über folgende Schritte definiert:

1. Pseudostichprobe: Berechne bzgl. der Stichprobe x_1, \dots, x_n die Pseudostichprobe u_1, \dots, u_n mit $u_i = (u_{i1}, u_{i2})$.

2. Berechnung der initialen Gewichtung: Berechne bzgl. u_1, \dots, u_n die Anzahlen der gebundenen Koordinaten b_1, \dots, b_n über

$$b_i = b(u_i | u_1, \dots, u_n) := \sum_{j=1}^2 \left(1 - \prod_{k \neq i} \mathbb{1}_{\{u_{ij} \neq u_{kj}\}} \right)$$

und die maximalen Bindungsgrößen g_1, \dots, g_n als

$$g_i = g(u_i | u_1, \dots, u_n) := \max_{j \in \{1, 2\}} \left\{ \sum_{k=1}^n \mathbb{1}_{\{u_{ij} = u_{kj}\}} \right\}.$$

Die Gewichtungen w_1, \dots, w_n für die Greedy-Auswahl werden über die im Voraus festgelegte Funktion $w_i := w(b_i, g_i)$ für u_i berechnet.

3. Greedy-Sortierung der Stichprobe: Sei $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ die Permutation, welche die Pseudostichprobe in absteigender Gewichtung anordnet, d. h.

$$w_{\pi(1)} \geqq \dots \geqq w_{\pi(n)}.$$

Weiter sei $m \in \{1, \dots, n\}$ derjenige Index, sodass

$$w_{\pi(1)} \geqq \dots \geqq w_{\pi(m)} > 0 = w_{\pi(m+1)} = \dots = w_{\pi(n)}$$

gilt; die Elemente mit Gewicht Null seien Elemente ohne Bindungen.

4. Auswahl der Elemente für die Reduktion und die Repräsentanten je Bindung: Es bezeichnen $I^{(k)}$ und $J^{(k)}$ die Indexmengen der Elemente, die die Bindungen repräsentieren, und der Elemente, die gelöscht werden sollen. Es seien $I^{(0)} := \emptyset$ und $J^{(0)} := \emptyset$.

Für $k' := \pi(k)$, $k \in \{1, \dots, m\}$, führe die folgenden Schritte durch, falls $k' \notin I^{(k-1)} \cup J^{(k-1)}$ ist, d. h. falls das Element in den Iterationen 1 bis $k-1$ noch nicht verarbeitet wurde:

(i.) Handhabe vollständiger Bindungen: Die Indexmenge

$$B_0^{(k)} := \{i \notin I^{(k-1)} \cup J^{(k-1)} : u_i = u_{k'}\}$$

beinhaltet die noch zu verarbeitenden Elemente der vollständigen Bindung bzgl. $u_{k'}$, $k' = \pi(k)$. Falls $B_0^{(k)} \neq \emptyset$ gilt, wird aus dieser Menge zufällig ein Repräsentant $i_0^{(k)} \in B^{(k)}$ ausgewählt.

- (ii.) Handhabe teilweiser Bindungen in der 1. und 2. Koordinate: *Analog zu $B_0^{(k)}$ bezeichnen die beiden Indexmengen*

$$B_1^{(k)} := \{i \notin I^{(k-1)} \cup J^{(k-1)} \cup B_0^{(k)} : u_{i1} = u_{k'1}\}$$

$$B_2^{(k)} := \{i \notin I^{(k-1)} \cup J^{(k-1)} \cup B_0^{(k)} : u_{i2} = u_{k'2}\}$$

die teilweise gebundenen Elemente zu $u_{k'}$, welche nicht vollständig gebunden sind. Aus diesen Indexmengen werden zufällig Elemente als Repräsentanten $i_1^{(k)} \in B_1^{(k)}$ und $i_2^{(k)} \in B_2^{(k)}$ ausgewählt.

- (iii.) Update der Indexmengen: *Es werden die Mengen $I^{(k)} \supseteq I^{(k-1)}$ und $J^{(k)} \supseteq J^{(k-1)}$ konstruiert als*

$$I^{(k)} := I^{(k-1)} \cup \{i_0^{(k)}, i_1^{(k)}, i_2^{(k)}\}$$

und

$$J^{(k)} := J^{(k-1)} \cup ((B_0^{(k)} \cup B_1^{(k)} \cup B_2^{(k)}) \setminus \{i_0^{(k)}, i_1^{(k)}, i_2^{(k)}\})$$

- (iv.) Abgleich mit der maximalen Anzahl der zu löschenenden Elemente: *Gilt $|J^{(k-1)}| \leq N < |J^{(k)}|$ oder ist $k = m$ erreicht, so wird diese Schleife gestoppt und J als $J = J^{(k-1)}$ bzw. $J^{(m)}$ gesetzt.*

5. Durchführung der Stichprobenreduktion: *Die Indexmenge der verbleibenden Elemente I nach der Stichprobenreduktion ist bzgl. J definiert als*

$$I := \{1, \dots, n\} \setminus J.$$

Es wird die reduzierte Stichprobe $(x_i)_{i \in I}$ ausgegeben.

Die Implementierung ist im Appendix D.3.2 zu finden. Zum Greedy-Algorithmus sind noch zwei Bemerkungen notwendig:

- 5.1.3 Bemerkung.** (i.) Parameterwahl: *In dieser Arbeit wird der Parameter N relativ zum Stichprobenumfang n als $N(n) := \lceil 0.25 \cdot n \rceil$ festgelegt. Es soll also höchstens ein Viertel der Stichprobe entfernt werden. Die Wahl des Parameters wurde durch den Autor festgelegt.*

- (ii.) Alternativer Greedy-Ansatz: *Der vorgestellte Greedy-Algorithmus legt die Gewichte zu Beginn fest; diese werden nicht weiter angepasst. Alternativ können die Gewichte nach jeder Iteration neu bestimmt werden. Dieser Ansatz erhöht die Komplexität und damit den Rechenaufwand, weswegen in dieser Arbeit fixierte Gewichte betrachtet werden.*

5.2 Randomisierungsmethoden

In diesem Abschnitt werden die sogenannten Randomisierungsmethoden vorgestellt, die Bindungen über zufällige Anpassungen der Stichprobe oder der Ränge bzw. Pseudostichprobe lösen. Bei diesem Ansatz werden zusätzliche Informationen in den Datensatz eingebracht, wodurch die Datengrundlage verfälscht werden kann. Um die Verfälschung der Datensätze bei Randomisierungen möglichst gering zu halten, sind zwei Vorgehensweisen zielführend:

Erstens können lediglich die Bindungen selbst angepasst werden, sodass der ungebundene Anteil der Stichprobe ohne Änderung bleibt. Die zugrundeliegende Idee ist ein minimalinvasives Lösen der Bindungen. Zweitens können unter Berücksichtigung der ursprünglichen Daten Zufallsdaten, z. B. ein Rauschen, generiert und zu der Stichprobe hinzugefügt werden. Dies ist begründet in der Idee, dass das Rauschen die Bindungen löst; durch die Anpassung bzgl. der ursprünglichen Stichprobe soll die Abhängigkeitsstruktur bei Hinzunahme des Rauschens möglichst erhalten bleiben.

Rangrandomisierung

Der erste Randomisierungsansatz in dieser Arbeit ist die zufällige Rangvergabe. Bei der Rangrandomisierung werden bei jeder Bindung die Ränge zwischen dem minimalen und maximalen Rang zufällig auf die gebundenen Elemente verteilt. Bei einer Pseudostichprobe entspricht die zufällige Rangvergabe dem Überschreiben von gebundenen Pseudo-Beobachtungen durch Pseudo-Beobachtungen bzgl. der Unabhängigkeitscopula.

In der Praxis sind Copulas für Abhängigkeitsmodellierungen abseits der Unabhängigkeit von Interesse. Eine Rangrandomisierung führt dann zu einer Veränderung bzw. Abschwächung der Abhängigkeitsstruktur in der Pseudostichprobe. Dies ist bei der Anwendung zu kritisieren (vgl. Hájek (1999): S. 131-132). Abhängig von der Bindungsstruktur werden die zufälligen Ränge über den folgenden Algorithmus vergeben. Die Vorangehensweise ist an Abschnitt 4.8.1 in Hájek (1999) angelehnt.

5.2.1 Algorithmus (Zufällige Rangvergabe). *Sei $x_1, \dots, x_n \in \mathbb{R}^d$ mit $x_i = (x_{i1}, \dots, x_{id})$ für $i \in \{1, \dots, n\}$ eine beliebige Stichprobe. Die zugehörigen Zufallsränge $r_i = (r_{i1}, \dots, r_{id})$ zu x_i werden komponentenweise generiert über*

$$r_{ij} = r_{\text{rnd}}(x_{ij} | x_{1j}, \dots, x_{nj}),$$

wobei $r_{\text{rnd}}(x_{ij} | x_{1j}, \dots, x_{nj})$ über den folgenden Algorithmus ausgewertet wird.

1. Initialisierung: Es bezeichnen $I^{(k)}$ die Indexmengen der Einträge, die in der k -ten Iteration zur Verfügung stehen und $J^{(k)}$ die Indexmengen der Bindungen und einzigartigen Elementen. Initialisiere $I^{(0)} := \{1, \dots, n\}$.

2. Identifizierung der Bindungen: Für alle $k \in \{1, \dots, n\}$ werden die folgenden Schritte zur Identifizierung der Bindungen durchgeführt:

- (i.) Bindungen vom Wert x_{kj} : Setze die Indexmenge $\emptyset \subseteq J^{(k)} \subseteq I^{(k-1)}$ der zu x_{kj} gebundenen und noch nicht verarbeiteten Elementen als

$$J^{(k)} := \{i \in I^{(k-1)} : x_{kj} = x_{ij}\}.$$

- (ii.) Indexmenge der noch zu verarbeitenden Elemente: Damit Elemente nicht mehrfach verarbeitet werden, setze die Indexmenge

$$I^{(k)} := I^{(k-1)} \setminus J^{(k)}.$$

Ist $I^{(k)} = \emptyset$, so müssen die Iterationen $k+1, \dots, n$ nicht durchgeführt werden, da $I^{(k)} = I^{(k+1)} = \dots = I^{(n)} = \emptyset$ gilt.

3. Zufällige Rangvergabe: Für alle $k \in \{1, \dots, n\}$ mit $I^{(k)} \neq \emptyset$ werden die Ränge der Elemente $(x_{ij} : i \in I^{(k)})$ über die folgenden Schritte vergeben:

- (i.) Menge der möglichen Ränge: Setze für die Bindung $(x_{ij} : i \in I^{(k)})$ die Menge der möglichen Ränge $R^{(k)} := \{r_{\min}^{(k)}, \dots, r_{\max}^{(k)}\}$ über die minimalen und maximalen Ränge

$$\begin{aligned} r_{\min}^{(k)} &= r_{\min}(x_{kj} | x_{1j}, \dots, x_{nj}), \\ r_{\max}^{(k)} &= r_{\max}(x_{kj} | x_{1j}, \dots, x_{nj}), \end{aligned}$$

wobei stets $k \in I^{(k)}$ für $I^{(k)} \neq \emptyset$ existiert.

- (ii.) Zufallspermutation: Sei $\text{Sym}(R^{(k)})$ die symmetrische Gruppe aller Permutationen $\pi : R^{(k)} \rightarrow R^{(k)}$. Für die Rangvergabe wird zufällig eine Permutation $\pi^{(k)} \in \text{Sym}(R^{(k)})$ ausgewählt.

- (iii.) Rangvergabe: Die Zufallsränge $(r_{ij} : i \in I^{(k)})$ für die Elemente $(x_{ij} : i \in I^{(k)})$ werden vergeben über:

$$(r_i : i \in I^{(k)}) = (\pi^{(k)}(r_{\min}^{(k)}), \dots, \pi^{(k)}(r_{\max}^{(k)})),$$

wobei die Indizes in $(r_{ij} : i \in I^{(k)})$ aufsteigend betrachtet werden.

Aufgrund von $I^{(1)} \cup \dots \cup I^{(n)} = \{1, \dots, n\}$ sind alle Elemente verarbeitet worden. Die Zufallsränge der x_{1j}, \dots, x_{nj} sind die generierten r_{1j}, \dots, r_{nj} (vgl. Hájek (1999): S. 131-132).

Die zufällige Rangvergabe kann an verschiedenen Punkten der in Abschnitt 3.3 präsentierten Tests und Methoden ansetzen. So können beispielsweise bei der empirischen Copula oder dem MPL-Schätzer (0, 1)-skalierte Zufallsränge zugrundeliegen. Eine Rangrandomisierung kann über die R-Funktion `rank(x, ties.method="random")` und bei darauf aufbauenden Funktionen über den Parameter `ties.method="random"` aktiviert werden.

Wertrandomisierung – Jittering

Die Jittering-Methode (englisch für *Zittern*) ist eine Randomisierung der gesamten Stichprobe. Es wird ein zufälliges Rauschen überlagert, unabhängig davon, ob ein Element gebunden oder ungebunden ist. Als Grundlage für die Jittering-Definition wird die Studie von Pappadà et al. (2016) genutzt.

5.2.2 Algorithmus (Jittering). *Seien $F : [0, 1]^d \rightarrow [0, 1]$ eine multivariate Verteilungsfunktion und $X_1, \dots, X_n \sim F$ i.i.d. Zufallsvektoren der Form $X_i = (X_{i1}, \dots, X_{id})$. Der Vektor $r = (r_1, \dots, r_d) \in \mathbb{R}_>^d$ definiert die Präzision. Bzgl. der Copula $C_J : [0, 1]^d \rightarrow [0, 1]$ wird wie folgt gejittert:*

1. Rauschen: Generiere für die Zufallsvektoren X_1, \dots, X_n das Rauschen als i.i.d. Zufallsvektoren $U_1, \dots, U_n \sim C_J$ mit $U_i = (U_{i1}, \dots, U_{id})$.
2. Jittering: Durch Überlagerung des bzgl. r skalierten Rauschens entsteht die gejitterte Stichprobe $\tilde{X}_1, \dots, \tilde{X}_n$ mit Elementen

$$\tilde{X}_i = (\tilde{X}_{i1}, \dots, \tilde{X}_{id}), \quad \tilde{X}_{ij} := X_{ij} + r_j \cdot U_{ij}$$

für $(i, j) \in \{1, \dots, n\} \times \{1, \dots, d\}$.

Nach dem Jittering wird die Stichprobe $\tilde{X}_1, \dots, \tilde{X}_n$ ausgegeben; diese kann für die Copula-Methoden genutzt werden (vgl. Pappadà et al. (2016): S. 2484).

5.2.3 Bemerkung. (i.) Präzisionsvektor: Der Vektor $r = (r_1, \dots, r_d)$ ist idealerweise aus dem Sachzusammenhang bekannt oder wird über die empirische Präzision bestimmt.

(ii.) Auswirkungen auf den ungebundenen Anteil: Da das Jittering auf die gesamte Stichprobe angewandt wird, sind die Auswirkungen auf den ungebundenen Anteil zu beachten. Gilt für die Präzision

$$r = (r_1, \dots, r_d) : \quad r_j \leq \min \{ |X_{ij} - X_{kj}| : X_{ij} \neq X_{kj}, i \neq k \},$$

so folgt wegen $0 \leq U_{ij} \leq 1$ für alle (i, j) , dass

$$X_{ij} < X_{kj} \implies \tilde{X}_{ij} := X_{ij} + r_j \cdot U_{ij} \leq X_{kj} \leq \tilde{X}_{kj}.$$

Unter der Stetigkeit von C_J folgt $\mathbb{P}(U_{ij} \in \{0, 1\}) = 0$, sodass fast sicher $\tilde{X}_{ij} < \tilde{X}_{kj}$ gilt. Bei der Anwendung von rangbasierten Methoden bleibt die Anordnung für ungebundene Elemente innerhalb der Pseudostichprobe erhalten.

Im Algorithmus 5.2.2 wird das Rauschen über die Copula C_J modelliert. Pappadà et al. (2016) bezeichnen die Auswahl der Copula als *Randomisierungsstrategie*. Dort werden drei Kandidaten für die Copula betrachtet: Die Unabhängigkeitscopula Π , die bivariate obere Frechét-Hoeffding-Schranke M sowie eine gemischte Copula bzgl. Π und M (vgl. Pappadà et al. (2016): S. 2485). In dieser Arbeit wird das Jittering bzgl. der Frechét-Hoeffding-Schranke nicht berücksichtigt. Das Rauschen bzgl. der Unabhängigkeitscopula stellt die einfachste Randomisierungsstrategie dar. Die empirisch angepasste gemischte Copula hat in einigen der dort präsentierten Simulationsstudien die besten Ergebnisse erzielt (vgl. Pappadà et al. (2016): S. 2488-2495). Die beiden verbleibenden Strategien werden als zusätzliche Annahmen an den Algorithmus 5.2.2 konkretisiert:

5.2.4 Annahme (Jittering mit unabhängigem Rauschen). *Für die Jittering-Copula $C_J : [0, 1]^d \rightarrow [0, 1]$ wird die Unabhängigkeitscopula*

$$C_J(u) := \Pi(u) = \prod_{j=1}^d u_j \quad (u \in [0, 1]^d)$$

gewählt. Es gilt dann $U_{ij} \stackrel{i.i.d.}{\sim} \text{Unif}_{[0,1]}$ für alle $(i, j) \in \{1, \dots, n\} \times \{1, \dots, d\}$.

Für eine bivariate Stichprobe kann ebenfalls eine Randomisierungsstrategie mit einer gemischten Copula, die über das empirische Kendall'sche Tau gewichtet wird, formuliert werden.

5.2.5 Annahme (Jittering mit empirisch angepasstem Rauschen). *Sei $\hat{\tau}_n$ das empirische Kendall'sche Tau der bivariaten Stichprobe X_1, \dots, X_n . Für die Jittering-Copula $C_J : [0, 1]^d \rightarrow [0, 1]$ wird die gemischte Copula*

$$C_J(u) := \text{mix}_{\hat{\tau}_n}(\Pi, M)(u) := (1 - \hat{\tau}_n)\Pi(u) + \hat{\tau}_n M(u) \quad (u \in [0, 1]^2)$$

bzgl. der Copulas Π und M mit Gewichten $((1 - \hat{\tau}_n), \hat{\tau}_n)$ gewählt.

Die Jittering-Methoden sind im Appendix D.3.3 und D.3.4 beigelegt.

5.3 Intervall-Zensur nach Li et al. (2020)

Bei der Rangrandomisierung aus Abschnitt 5.2 werden, wie im Algorithmus 5.2.1 gezeigt, die Ränge bei Bindungen über zufällige Permutationen auf den jeweiligen Mengen der möglichen Ränge $\{r_{\min}, \dots, r_{\max}\} \subset \mathbb{N}$ vergeben.

Statt diskreter Werte wird bei der Intervall-Zensur für jedes gebundene Element ein Intervall $[r_{\min}, r_{\max}] \subset \mathbb{R}$ betrachtet; dieses Intervall umfasst die

zugehörigen möglichen Ränge. Beim Umgang mit Pseudostichproben und einer Nutzung von ebendiesen Intervallen statt einzelner willkürlich gesetzter Elemente wird von einer Intervall-zensierten Pseudostichprobe gesprochen: Die Position innerhalb der Pseudostichprobe ist nur bzgl. des Intervalls scharf bestimmbar. Bei Bindungen kann kein präziserer Wert bestimmt werden, ohne weitere Annahmen an die Rangvergabe oder andere Informationen heranzuziehen (vgl. Lee (2003): S. 3-5, *Type III Censoring*).

5.3.1 Definition (Intervall-zensierte Pseudostichproben). *Es seien Zufallsvektoren $X_1, \dots, X_n \sim F$ von Vektoren der Form $X_i = (X_{i1}, X_{i2})$ gegeben.*

Um die Intervall-zensierte Pseudostichprobe für X_1, \dots, X_n zu definieren, sind die Mengen der möglichen Positionen zu bestimmen. Für jedes Element X_i werden koordinatenweise die Grenzen U_{ij}^\wedge und U_{ij}^\vee der Zensur-Intervalle $[U_{ij}^\wedge, U_{ij}^\vee]$ bestimmt als die $(0, 1)$ -skalierten minimalen und maximalen Ränge

$$U_{ij}^\wedge := \frac{r_{\min}(X_{ij} | X_{1j}, \dots, X_{nj})}{n+1}, \quad U_{ij}^\vee := \frac{r_{\max}(X_{ij} | X_{1j}, \dots, X_{nj})}{n+1} \quad (5-1)$$

für $(i, j) \in \{1, \dots, n\} \times \{1, 2\}$. Setze $U_i^\wedge := (U_{i1}^\wedge, U_{i2}^\wedge)$ und $U_i^\vee := (U_{i1}^\vee, U_{i2}^\vee)$.

Für ungebundene Elementen gilt die Gleichung $U_i^\wedge = U_i^\vee =: U_i$, sodass in diesem Fall statt der Intervalle ein einzelner Punkt vorliegt und nicht zensiert wird. Bei gebundenen Elementen werden die drei Zensierungen

$$\{U_{i1}\} \times [U_{i2}^\wedge, U_{i2}^\vee] \quad \text{oder} \quad [U_{i1}^\wedge, U_{i1}^\vee] \times \{U_{i2}\} \quad \text{oder} \quad [U_{i1}^\wedge, U_{i1}^\vee] \times [U_{i2}^\wedge, U_{i2}^\vee] \quad (5-2)$$

durchgeführt.

Die Intervall-zensierte Pseudostichprobe besteht aus den Punktmengen $\{U_i\}$ bzw. der Punkte U_i , falls das jeweilige Element X_i ungebunden ist, und ansonsten aus den Zensierungen in (5-2) (vgl. Li et al. (2020): S. 105).

5.3.2 Bemerkung. Bei den Zensierungen in (5-2) sind die ersten beiden Zensierungen bei teilweisen Bindungen, d. h. für $U_i^\wedge \leq U_i^\vee$ mit $U_i^\wedge \neq U_i^\vee$, durchzuführen. Die dritte Zensierung findet bei vollständigen Bindungen, also im Fall $U_i^\wedge < U_i^\vee$, statt.

5.3.3 Beispiel (Intervall-zensierte Pseudostichproben). *Es sei eine neunelementige gebundene Stichprobe x_1, \dots, x_{10} mit den Werten*

$$\begin{pmatrix} 1 \\ 3 \end{pmatrix}^\top, \quad \begin{pmatrix} 6 \\ 5 \end{pmatrix}^\top, \quad \begin{pmatrix} 2 \\ 9 \end{pmatrix}^\top, \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix}^\top, \quad \begin{pmatrix} 4 \\ 2 \end{pmatrix}^\top, \quad \begin{pmatrix} 9 \\ 1 \end{pmatrix}^\top, \quad \begin{pmatrix} 3 \\ 4 \end{pmatrix}^\top, \quad \begin{pmatrix} 4 \\ 4 \end{pmatrix}^\top, \quad \begin{pmatrix} 2 \\ 9 \end{pmatrix}^\top$$

gegeben, für die die Pseudostichprobe berechnet werden soll. Die Pseudostichproben unter minimaler und maximaler Rangvergaben lauten:

$$\text{Min: } \begin{pmatrix} 0.1 \\ 0.4 \end{pmatrix}^\top, \begin{pmatrix} 0.8 \\ 0.7 \end{pmatrix}^\top, \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}^\top, \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}^\top, \begin{pmatrix} 0.6 \\ 0.3 \end{pmatrix}^\top, \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix}^\top, \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}^\top, \begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix}^\top, \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}^\top$$

$$\text{Max: } \begin{pmatrix} 0.1 \\ 0.4 \end{pmatrix}^\top, \begin{pmatrix} 0.8 \\ 0.7 \end{pmatrix}^\top, \begin{pmatrix} 0.4 \\ 0.9 \end{pmatrix}^\top, \begin{pmatrix} 0.4 \\ 0.2 \end{pmatrix}^\top, \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix}^\top, \begin{pmatrix} 0.9 \\ 0.2 \end{pmatrix}^\top, \begin{pmatrix} 0.5 \\ 0.6 \end{pmatrix}^\top, \begin{pmatrix} 0.7 \\ 0.6 \end{pmatrix}^\top, \begin{pmatrix} 0.4 \\ 0.9 \end{pmatrix}^\top$$

Die Intervall-zensierte Pseudostichprobe besteht aus den folgenden Mengen:

$$\begin{aligned} & \{0.1\} \times \{0.4\}, & \{0.8\} \times \{0.7\}, & [0.2, 0.4] \times [0.8, 0.9], \\ & [0.2, 0.4] \times [0.1, 0.2], & [0.6, 0.7] \times \{0.3\}, & \{0.9\} \times [0.1, 0.2], \\ & \{0.5\} \times [0.5, 0.6], & [0.6, 0.7] \times [0.5, 0.6], & [0.2, 0.4] \times [0.8, 0.9]. \end{aligned}$$

Auf Basis einer Intervall-zensierten Pseudostichprobe kann weiter eine Maximum-Pseudo-Likelihood-Schätzung für die Copula unter Annahme der Familienzugehörigkeit durchgeführt werden. Bei der zu maximierenden Pseudo-Likelihood-Funktion werden Fallunterscheidungen bzgl. der Zensierungen in Formel (5–2) betrachtet. Um diese Definition zu vereinfachen, gelte für eine differenzierbare bivariate Copula $C : [0, 1]^2 \rightarrow [0, 1]$ die folgende abkürzende Schreibweise

$$\partial_j C_\vartheta(u_1, u_2) := \left. \frac{\partial C_\vartheta(v_1, v_2)}{\partial v_j} \right|_{\substack{v_1 = u_1 \\ v_2 = u_2}}$$

für die Auswertung der j -ten partiellen Ableitung, $j \in \{1, 2\}$.

5.3.4 Definition (Intervall-zensierte MPLE für bivariate Copulafamilien nach Li et al. (2020)). Sei $C := C_{\vartheta_0} : [0, 1]^2 \rightarrow [0, 1]$ eine Copula aus der Familie $\mathcal{C}_0 := \{C_\vartheta : \vartheta \in \Theta\}$ mit unbekanntem Parameter $\vartheta_0 \in \Theta$. Die Dichte der Copula $C_\vartheta \in \mathcal{C}_0$ lautet $c_\vartheta(u) := c(u; \vartheta)$.

Für eine stetige Verteilungsfunktion $F : \mathbb{R}^2 \rightarrow [0, 1]$ mit $F = C(F_1, \dots, F_d)$ sei $X_1, \dots, X_n \sim F$ eine i.i.d. Stichprobe. Für alle $i \in \{1, \dots, n\}$ bezeichnen $U_i^\wedge = (U_{i1}^\wedge, U_{i2}^\wedge)$ und $U_i^\vee = (U_{i1}^\vee, U_{i2}^\vee)$ die begrenzenden Vektoren zur Intervall-Zensur aus (5–1).

Die Log-Pseudo-Likelihood-Funktion $\mathcal{L} : \Theta \rightarrow \mathbb{R}$, $\vartheta \mapsto \mathcal{L}(\vartheta; \dots)$ für die Intervall-zensierte MPLE ist bzgl. der Werte $u_1^\wedge, u_1^\vee, \dots, u_n^\wedge, u_n^\vee$ definiert als

$$\mathcal{L}(\vartheta; u_1^\wedge, u_1^\vee, \dots, u_n^\wedge, u_n^\vee) := \sum_{i=1}^n \log L(u_i^\wedge, u_i^\vee; \vartheta),$$

wobei für die Likelihood-Funktion die Fallunterscheidung

$$L(\vartheta; u_i^\wedge, u_i^\vee) := \begin{cases} c(u; \vartheta), & \text{falls } u = u_i^\wedge = u_i^\vee, \\ \partial_1 C_\vartheta(u_{i1}^\wedge, u_{i2}) - \partial_1 C_\vartheta(u_{i1}^\vee, u_{i2}), & \text{falls } u_i^\wedge \leq u_i^\vee \text{ und} \\ & u_{i2} = u_{i2}^\wedge = u_{i2}^\vee, \\ \partial_2 C_\vartheta(u_{i1}, u_{i2}^\wedge) - \partial_2 C_\vartheta(u_{i1}, u_{i2}^\vee), & \text{falls } u_i^\wedge \leq u_i^\vee \text{ und} \\ & u_{i1} = u_{i1}^\wedge = u_{i1}^\vee, \\ \text{Vol}|_{C_\vartheta}([u_{i1}^\vee, u_{i1}^\wedge] \times [u_{i2}^\vee, u_{i2}^\wedge]), & \text{falls } u_i^\wedge < u_i^\vee \end{cases}$$

bzgl. der Punkte $u_i^\wedge = (u_{i1}^\wedge, u_{i2}^\wedge)$ und $u_i^\vee = (u_{i1}^\vee, u_{i2}^\vee)$ gelte. Der Intervallzensierte MPL-Schätzer $\hat{\vartheta}_n$ für den unbekannten Parameter $\vartheta_0 \in \Theta$ wird über die Maximierung von \mathcal{L} auf der zensierten Pseudostichprobe bestimmt als

$$\hat{\vartheta}_n = \hat{\vartheta}_n(U_1^\wedge, U_1^\vee, \dots, U_n^\wedge, U_n^\vee) = \arg \max_{\vartheta \in \Theta} \mathcal{L}(\vartheta; U_1^\wedge, U_1^\vee, \dots, U_n^\wedge, U_n^\vee).$$

Die Intervall-zensierte MPLE stimmt bei ungebundenen Daten mit der MPLE aus Definition 3.3.11 überein (vgl. Li et al. (2020): S. 105).

In Li et al. (2020) werden neben einer Parameterschätzung ebenso das Konfidenzintervall approximiert und ein Test der Anpassungsgüte formuliert; diese sind in der nächsten Bemerkung zusammengefasst.

5.3.5 Bemerkung. (i.) Konfidenzintervalle: Zum Niveau $\alpha \in (0, 1)$ kann das $(1 - \alpha)$ -Konfidenzintervall des Intervall-zensierten MPL-Schätzers $\hat{\vartheta}_n$ über einen Bootstrap-Ansatz approximiert werden. Es bezeichnet $k \in \{1, \dots, N\}$ den Index der N -vielen Bootstrap-Iterationen, $N \in \mathbb{N}$. In jeder Iteration werden i.i.d. Zufallsvektoren $U_1^{(k)}, \dots, U_n^{(k)} \sim C_{\hat{\vartheta}_n}$ realisiert. Eine ITS bzgl. der empirischen Marginale zu den originalen Daten überträgt die Bindungsstruktur auf die Bootstrap-Stichprobe. Auf dieser Grundlage wird die zensierte Schätzung $\hat{\vartheta}_n^{(k)}$ für ϑ_0 durchgeführt. Das Konfidenzintervall wird über die $(\alpha/2)$ - und $(1 - \alpha/2)$ -Quantile von $\hat{\vartheta}_n^{(1)}, \dots, \hat{\vartheta}_n^{(N)}$ berechnet (vgl. Li et al. (2020): S. 106)

(ii.) Test der Anpassungsgüte: Li et al. (2020) formuliert einen Test der Anpassungsgüte für die Intervall-zensierte Modellanpassung. Dieser basiert auf dem Algorithmus 5.4.4 nach Kojadinovic (2017), wobei in den Bootstrap-Iterationen jeweils eine Intervall-Zensur zur Modellanpassung durchgeführt wird (vgl. Li et al. (2020): S. 106-107).

Die Intervall-zensierte MPLE ist im R-Paket Censorcopula von Li et al. (2016) implementiert. Die Log-Pseudo-Likelihood-Funktion und die Modellanpassung sind in diesem Paket über die Funktionen `Newloglik2` und `intervalFitb` aufrufbar. Zur Implementierung des Tests der Anpassungsgüte wurde die Intervall-zensierte MPLE und der Test erneut implementiert. Die Implementierungen sind in D.3.5 und D.3.6 zu finden.

5.4 Bootstrap-Methoden für Datensätze mit Bindungen nach Kojadinovic (2017)

In diesem Abschnitt werden die Methoden aus Kojadinovic (2017) präsentiert. Diese stellen Alternativen zu den in Abschnitt 3.3 behandelten Tests auf Radial-Symmetrie, Austauschbarkeit und Anpassungsgüte dar. Im Gegensatz zu den Algorithmen 3.3.3, 3.3.8 und 3.3.15 wurden die Kojadinovic-Algorithmen mit dem Ziel der Robustheit unter Bindungen konzipiert.

Die Bootstrap-Iterationen in den Algorithmen haben das Ziel die Stichproben derart künstlich anzupassen, dass die Verteilung der Teststatistik unter der angenommenen Nullhypothese approximiert wird.

Tests auf Radial-Symmetrie und Austauschbarkeit

Der Test auf Radial-Symmetrie entspricht *Procedure 4.1* in Kojadinovic (2017). In dieser Arbeit wird eine bivariate Version des Algorithmus präsentiert.

5.4.1 Algorithmus (Rangbasierter Test auf Radial-Symmetrie mit Bootstrap zur Handhabe von gebundenen Stichproben nach Kojadinovic (2017)).
Seien $X_1, \dots, X_n \sim F$ i.i.d. Zufallsvektoren mit Komponenten $X_i = (X_{i1}, X_{i2})$ und $F : \mathbb{R}^2 \rightarrow [0, 1]$ eine Verteilungsfunktion. Der Parameter $N \in \mathbb{N}$ legt die Anzahl der Bootstrap-Iterationen fest.

Der Test auf Radial-Symmetrie inklusive der Approximation der p-Werte wird mit folgendem Algorithmus durchgeführt:

1. Pseudostichprobe: Berechne für die Stichprobe X_1, \dots, X_n die $(0, 1)$ -skalierten Mid-Ränge U_1, \dots, U_n mit Komponenten $U_i = (U_{i1}, U_{i2})$.
2. Berechnung der Teststatistik: Berechne \mathcal{S}_n aus Formel (3–12), d. h.

$$\mathcal{S}_n := \sum_{i=1}^n (\hat{C}_n(U_i) - \hat{\hat{C}}_n(U_i))^2,$$

wobei \hat{C}_n und $\hat{\hat{C}}_n$ bzgl. der U_1, \dots, U_n konstruiert sind.

3. Empirische Marginale: Konstruiere bzgl. X_1, \dots, X_d die empirischen Marginale $\hat{F}_{n,j} : \mathbb{R} \rightarrow [0, 1]$ als

$$\hat{F}_{n,j}(x) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_{ij} \leq x\}} \quad (x \in \mathbb{R})$$

für $j \in \{1, 2\}$ und es bezeichnet $\hat{F}_{n,j}^\leftarrow$ die zugehörige Quantilfunktion.

4. Bootstrap-Approximation der Verteilung von \mathcal{S}_n unter Annahme der Radial-Symmetrie: Für jede Iteration $k \in \{1, \dots, N\}$ werden die folgenden Berechnungen durchgeführt:

- (i.) Zufällige Spiegelung: Die Zufallsvariablen $B_1^{(k)}, \dots, B_n^{(k)}$ i.i.d. $\sim \text{Bin}_{1,1/2}$ mit $B_1^{(k)}, \dots, B_n^{(k)} \perp\!\!\!\perp U_1, \dots, U_n$ werden zur Steuerung der Spiegelung genutzt. Die Spiegelung resultiert in den Zufallsvektoren

$$V_i^{(k)} := B_i^{(k)} \cdot U_i + (1 - B_i^{(k)}) \cdot (1 - U_i)$$

für $i \in \{1, \dots, n\}$.

- (ii.) Abbildung der ursprünglichen Struktur: Die Vektoren $V_1^{(k)}, \dots, V_n^{(k)}$ werden koordinatenweise bzgl. der Marginale aus Schritt 3. quantiltransformiert zu den Zufallsvektoren $W_1^{(k)}, \dots, W_n^{(k)}$ mit

$$W_i^{(k)} := (\hat{F}_{n,1}^\leftarrow(V_{i1}^{(k)}), \hat{F}_{n,2}^\leftarrow(V_{i2}^{(k)})).$$

Die Bootstrap-Pseudostichprobe $U_1^{(k)}, \dots, U_n^{(k)}$ besteht aus den $(0, 1)$ -skalierten Mid-Rängen mit den Koordinaten

$$U_{ij}^{(k)} := \frac{1}{n+1} r_{\text{mid}}(W_{ij}^{(k)} | W_{1j}^{(k)}, \dots, W_{nj}^{(k)})$$

für $(i, j) \in \{1, \dots, n\} \times \{1, 2\}$.

- (iii.) Bootstrap-Version $\mathcal{S}_n^{(k)}$ der Teststatistik \mathcal{S}_n : Die Teststatistik wird berechnet als

$$\mathcal{S}_n^{(k)} := \sum_{i=1}^n (\hat{C}_n^{(k)}(U_i^{(k)}) - \hat{\bar{C}}_n^{(k)}(U_i^{(k)}))^2,$$

wobei $\hat{C}_n^{(k)}$ und $\hat{\bar{C}}_n^{(k)}$ bzgl. der $U_1^{(k)}, \dots, U_n^{(k)}$ konstruiert sind.

5. Berechnung des p-Wertes: Der für die Teststatistik \mathcal{S}_n approximierte p-Wert $p(\mathcal{S}_n) \in (0, 1)$ lautet

$$p(\mathcal{S}_n) = \frac{1}{N+1} \sum_{k=1}^N (\mathbb{1}_{\{\mathcal{S}_n^{(k)} > \mathcal{S}_n\}} + 0.5).$$

Der Test auf Austauschbarkeit ist *Procedure 3.1* in Kojadinovic (2017).

5.4.2 Algorithmus (Rangbasierter Test auf Austauschbarkeit mit Bootstrap zur Handhabe von gebundenen Stichproben nach Kojadinovic (2017)). Seien $X_1, \dots, X_n \sim F$ i.i.d. Zufallsvektoren mit Komponenten $X_i = (X_{i1}, X_{i2})$ und $F : \mathbb{R}^2 \rightarrow [0, 1]$ eine Verteilungsfunktion. Der Parameter $N \in \mathbb{N}$ legt die Anzahl der Bootstrap-Iterationen fest.

Der Test auf Austauschbarkeit inklusive der Approximation der p-Werte wird mit folgendem Algorithmus durchgeführt:

1. Pseudostichprobe: Berechne für die Stichprobe X_1, \dots, X_n die $(0, 1)$ -skalierten Mid-Ränge U_1, \dots, U_n mit Komponenten $U_i = (U_{i1}, U_{i2})$.
2. Berechnung der Teststatistik: Berechne \mathcal{A}_n aus Formel (3-15), d. h.

$$\mathcal{A}_n := \sum_{i=1}^n (\hat{C}_n(U_{i1}, U_{i2}) - \hat{C}_n(U_{i2}, U_{i1}))^2,$$

wobei die empirische Copula \hat{C}_n bzgl. der U_1, \dots, U_n konstruiert ist.

3. Empirische Marginale: Konstruiere bzgl. X_1, \dots, X_d die empirischen Marginale $\hat{F}_{n,j} : \mathbb{R} \rightarrow [0, 1]$ als

$$\hat{F}_{n,j}(x) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_{ij} \leq x\}} \quad (x \in \mathbb{R})$$

für $j \in \{1, 2\}$ und es bezeichnet $\hat{F}_{n,j}^{-1}$ die zugehörige Quantilfunktion.

4. Bootstrap-Approximation der Verteilung von \mathcal{S}_n unter Annahme der Austauschbarkeit: Für jede Iteration $k \in \{1, \dots, N\}$ werden die folgenden Berechnungen durchgeführt:

- (i.) Zufälliges Austauschen: Die Zufallsvariablen $B_1^{(k)}, \dots, B_n^{(k)}$ $\stackrel{i.i.d.}{\sim}$ $\text{Bin}_{1,1/2}$ mit $B_1^{(k)}, \dots, B_n^{(k)}$ $\perp\!\!\!\perp$ U_1, \dots, U_n werden zur Steuerung

des Austausches genutzt. Die Zufallsvektoren $V_1^{(k)}, \dots, V_n^{(k)}$ mit $V_i^{(k)} := (V_{i1}^{(k)}, V_{i2}^{(k)})$ werden konstruiert über

$$V_{i1}^{(k)} := B_i^{(k)} \cdot U_{i1} + (1 - B_i^{(k)}) \cdot U_{i2},$$

$$V_{i2}^{(k)} := (1 - B_i^{(k)}) \cdot U_{i1} + B_i^{(k)} \cdot U_{i2}$$

für $i \in \{1, \dots, n\}$ konstruiert.

- (ii.) Abbildung der ursprünglichen Struktur: Die Vektoren $V_1^{(k)}, \dots, V_n^{(k)}$ werden koordinatenweise bzgl. der Marginalen aus Schritt 3. quantiltransformiert zu den Zufallsvektoren $W_1^{(k)}, \dots, W_n^{(k)}$ mit

$$W_i^{(k)} := (\hat{F}_{n,1}^{\leftarrow}(V_{i1}^{(k)}), \hat{F}_{n,2}^{\leftarrow}(V_{i2}^{(k)})).$$

Die Bootstrap-Pseudostichprobe $U_1^{(k)}, \dots, U_n^{(k)}$ besteht aus den $(0, 1)$ -skalierten Mid-Rängen mit den Koordinaten

$$U_{ij}^{(k)} := \frac{1}{n+1} r_{\text{mid}}(W_{ij}^{(k)} | W_{1j}^{(k)}, \dots, W_{nj}^{(k)})$$

für $(i, j) \in \{1, \dots, n\} \times \{1, 2\}$.

- (iii.) Bootstrap-Version $\mathcal{A}_n^{(k)}$ der Teststatistik \mathcal{A}_n : Die Teststatistik wird berechnet als

$$\mathcal{A}_n^{(k)} := \sum_{i=1}^n (\hat{C}_n^{(k)}(U_{i1}^{(k)}, U_{i2}^{(k)}) - \hat{C}_n^{(k)}(U_{i2}^{(k)}, U_{i1}^{(k)}))^2,$$

wobei $\hat{C}_n^{(k)}$ bzgl. der $U_1^{(k)}, \dots, U_n^{(k)}$ konstruiert ist.

5. Berechnung des p-Wertes: Der für die Teststatistik \mathcal{A}_n approximierte p-Wert $p(\mathcal{A}_n) \in (0, 1)$ lautet

$$p(\mathcal{A}_n) = \frac{1}{N+1} \sum_{k=1}^N (\mathbb{1}_{\{\mathcal{A}_n^{(k)} > \mathcal{A}_n\}} + 0.5).$$

5.4.3 Bemerkung. Zu den Algorithmen 5.4.1 und 5.4.2 sind die folgenden Punkte anzumerken:

- (i.) Multivariate Version: Wie beim Algorithmus 3.3.3 kann der Test auf Radial-Symmetrie 5.4.1 ebenso für den multivariaten Fall formuliert werden. Die Originalfassung des Algorithmus in Procedure 4.1 in Kojadinovic (2017) berücksichtigt multivariate Copulas.

- (ii.) Notationen: Die ITS in Schritt 4.(ii.) wird in der Originalfassung über die Ordnungsstatistiken der Ränge und Permutationen zur Strukturierung gelöst (vgl. Kojadinovic (2017): S. 7, 10, Procedure 4.1).
- (iii.) Vereinfachte Algorithmen: Kojadinovic (2017) betrachtet ebenso vereinfachte Versionen der Algorithmen ohne Abbildung der ursprünglichen (Bindungs-)Strukturen. Im vereinfachten Algorithmus sind die Schritte 3. und 4.(ii.) wegzulassen. Die Bootstrap-Pseudostichprobe $U_1^{(k)}, \dots, U_n^{(k)}$ ist auf Basis der $V_1^{(k)}, \dots, V_n^{(k)}$ zu berechnen.

Die Tests für Radial-Symmetrie und Austauschbarkeit können über die Funktionen `radSymTest(x, ties=TRUE)` und `exchTest(x, ties=TRUE)` aus dem R-Paket Copula von Hofert et al. (2020) durchgeführt werden; es wird keine separate Implementierung im Appendix bereitgestellt.

Test der Anpassungsgüte

Der Test der Anpassungsgüte in Kojadinovic (2017) verfolgt einen ähnlichen Ansatz wie der Test 3.3.15 aus Genest et al. (2009). Es wird erneut die Bindungsstruktur über Transformationen eingebracht.

5.4.4 Algorithmus (Rangbasierter Test der Anpassungsgüte mit Bootstrap zur Handhabe von gebundenen Stichproben nach Kojadinovic (2017)). Seien $X_1, \dots, X_n \sim F$ i.i.d. Zufallsvektoren mit Komponenten $X_i = (X_{i1}, \dots, X_{id})$ und $F : \mathbb{R}^d \rightarrow [0, 1]$ eine Verteilungsfunktion. Sei $C = C_{\vartheta_0} \in \mathcal{C}_0$ die zu F gehörige Copula aus der Familie $\mathcal{C}_0 = \{C_\vartheta : \vartheta \in \Theta\}$. Der Parameter $N \in \mathbb{N}$ legt die Anzahl der Bootstrap-Iterationen fest.

Der Test der Anpassungsgüte des MPL-Schätzers inklusive der Approximation der p -Werte wird mit folgendem Algorithmus durchgeführt:

1. Pseudostichprobe: Berechne für die Stichprobe X_1, \dots, X_n die $(0, 1)$ -skalierten mittleren Ränge U_1, \dots, U_n und maximalen Ränge $\tilde{U}_1, \dots, \tilde{U}_n$ mit Komponenten $U_i = (U_{i1}, \dots, U_{id})$ bzw. $\tilde{U}_i = (\tilde{U}_{i1}, \dots, \tilde{U}_{id})$.
2. Rangbasierte Copula-Schätzer: Sei $\hat{\vartheta}_n := \hat{\vartheta}(U_1, \dots, U_n)$ der rangbasierte MPL-Schätzer für den Parameter $\vartheta_0 \in \Theta$. Der zugehörige MPL-Copula-Schätzer ist $C_{\hat{\vartheta}_n}$. Konstruiere die empirische Copula \hat{C}_n bzgl. der Max-Ränge $\tilde{U}_1, \dots, \tilde{U}_n$.
3. Berechnung der Teststatistik: Berechne \mathcal{G}_n aus Formel (3–17) bzgl. der Max-Ränge, d. h.

$$\mathcal{G}_n := \sum_{i=1}^n (\hat{C}_n(\tilde{U}_i) - C_{\hat{\vartheta}_n}(\tilde{U}_i))^2.$$

4. Empirische Marginale: Konstruiere bzgl. X_1, \dots, X_d die empirischen Marginale $\hat{F}_{n,j} : \mathbb{R} \rightarrow [0, 1]$ als

$$\hat{F}_{n,j}(x) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_{ij} \leq x\}} \quad (x \in \mathbb{R})$$

für $j \in \{1, \dots, d\}$ und es bezeichnet $\hat{F}_{n,j}^\leftarrow$ die zugehörige Quantilfunktion.

5. Bootstrap-Approximation der Verteilung von \mathcal{G}_n : Für jede Iteration $k \in \{1, \dots, N\}$ werden die folgenden Berechnungen durchgeführt:

(i.) Zufallsaspekt je Iteration: Generiere bzgl. der Copula $C_{\hat{\vartheta}_n}$ die Stichprobe $V_1^{(k)}, \dots, V_n^{(k)}$ i.i.d. $C_{\hat{\vartheta}_n}$.

(ii.) Abbildung der ursprünglichen Struktur: Die Vektoren $V_1^{(k)}, \dots, V_n^{(k)}$ werden koordinatenweise bzgl. der Marginale aus Schritt 3. quantiltransformiert zu den Zufallsvektoren $W_1^{(k)}, \dots, W_n^{(k)}$ mit

$$W_i^{(k)} := (\hat{F}_{n,1}^\leftarrow(V_{i1}^{(k)}), \hat{F}_{n,2}^\leftarrow(V_{i2}^{(k)})) .$$

Berechne bzgl. $W_1^{(k)}, \dots, W_n^{(k)}$ die Bootstrap-Pseudostrichproben der $(0, 1)$ -skalierten mittleren Ränge $U_1^{(k)}, \dots, U_n^{(k)}$ und maxima- len Ränge $\tilde{U}_1^{(k)}, \dots, \tilde{U}_n^{(k)}$.

(iii.) Rangbasierte Copula-Schätzer: Sei $\hat{\vartheta}_n^{(k)} := \hat{\vartheta}(U_1^{(k)}, \dots, U_n^{(k)})$ der rangbasierte MPL-Schätzer für den Parameter $\vartheta_0 \in \Theta$. Der zugehörige MPL-Copula-Schätzer ist $C_{\hat{\vartheta}_n^{(k)}}$. Konstruiere die empirische Copula $\hat{C}_n^{(k)}$ bzgl. der Max-Ränge $\tilde{U}_1^{(k)}, \dots, \tilde{U}_n^{(k)}$.

(iv.) Berechnung der Bootstrap-Teststatistik: Berechne $\mathcal{G}_n^{(k)}$ aus Formel (3–17) bzgl. der Max-Ränge, d. h.

$$\mathcal{G}_n^{(k)} := \sum_{i=1}^n (\hat{C}_n^{(k)}(\tilde{U}_i^{(k)}) - C_{\hat{\vartheta}_n^{(k)}}(\tilde{U}_i^{(k)}))^2 .$$

6. Berechnung des p-Wertes: Der für die Teststatistik \mathcal{G}_n approximierte p-Wert $p(\mathcal{G}_n) \in (0, 1)$ lautet

$$p(\mathcal{G}_n) = \frac{1}{N+1} \sum_{k=1}^N (\mathbf{1}_{\{\mathcal{G}_n^{(k)} > \mathcal{G}_n\}} + 0.5) .$$

Der Test der Anpassungsgüte ist im R-Paket Copula von Hofert et al. (2020) als Funktion `gofCopula(copula, x, ties=TRUE)` implementiert, wobei der Parameter `copula` die Copulafamilie der Nullhypothese beinhaltet. Es wird keine separate Implementierung im Appendix bereitgestellt.

Kapitel 6

Simulationsstudie

6.1 Rahmenbedingungen der Simulationen

Die Simulationen sind in der Programmiersprache R vom R Core Team (2021), in den Versionen 3.6.1 bis 4.0.1, implementiert. Die Durchführung der Simulationen fand auf dem Hochleistungs-Rechencluster *HILBERT* der Heinrich-Heine-Universität (2021) statt. Es war die Installation der *GNU Scientific Library* von Galassi et al. (2021) und der folgenden R-Pakete notwendig: Copula von Hofert et al. (2020), Censorcopula von Li et al. (2016), und Optparse von Davis (2020). Aufgrund eines Installationskonflikts auf dem Cluster HILBERT wurde der Quellcode des Pakets Censorcopula herangezogen.

Die betrachteten Copulas stammen aus der Gauß'schen Copulafamilie, t-Copulafamilie mit fixiertem Freiheitsgrad $\nu = 4$, Clayton-Copulafamilie oder Gumbel-Copulafamilie. Für die Studie zur Austauschbarkeit werden die zugehörigen Khoudraji-Copulafamilien $kho_s(C_\vartheta, \Pi)$ hinzugenommen. Der Formparameter $s = (s_1, s_2)$ wird wie in Kojadinovic (2017) als $s = (0.95, 0.20)$ gewählt (vgl. Kojadinovic (2017): S. 8-9, Table 3). Alle genannten Copulafamilien sind mit Kendall'schen Taus $\tau \in \{0.25, 0.50, 0.75\}$ modelliert.

In der Simulation finden jeweils $N = 500$ bzw. $N = 250$ Iterationen mit Stichproben der Größe $n = 250$ bzw. $n = 150$ statt. Die größeren Simulationen ($N = 500, n = 250$) werden für die Tests auf Radialsymmetrie und Austauschbarkeit sowie für die Modellanpassung durchgeführt. Die Tests der Anpassungsgüte werden aufgrund des hohen Rechenaufwands mit bis zu $N = 250$ Iterationen der Größe $n = 150$ durchgeführt; die Laufzeit wird auf 60 Stunden pro Kombination aus Copula, Kendall'schem Tau und Modellanpassung beschränkt.

Für die Rahmenbedingungen für Finanzmärkte werden die Annahmen 4.4.3 genutzt. Zur Vereinfachung der Simulationsstudie werden beide Koordinatenachsen auf den Intervall $[0, 1]$ beschränkt.

naten gleichermaßen angepasst. Diese Anpassungen werden nicht gleichzeitig umgesetzt, d. h. es wird entweder die Präzision oder die Nullsetzung oder die empirische Struktur modelliert.

Vorgehensweise zur Bemessung der Güte und zum Vergleich der statistischen Testmethoden

Da die Vorgehensweisen für verschiedene Fragestellungen und Hypothesen zusammengefasst dargestellt werden sollen, werden diese in verallgemeinerter Form beschrieben. Bezeichne dazu $(\mathcal{X}, \mathcal{B}, \{P_\vartheta : \vartheta \in \Theta\})$ ein statistisches Modell mit dem Stichprobenraum \mathcal{X} , der σ -Algebra der Ereignisse \mathcal{B} und einer Familie $\{P_\vartheta : \vartheta \in \Theta\}$ von Wahrscheinlichkeitsmaßen. $\Theta = \Theta|_{\mathcal{H}_0} \cup \Theta|_{\mathcal{H}_1}$ bezeichnet den Parameterraum und dessen Zerlegung bzgl. der Hypothesen \mathcal{H}_0 und \mathcal{H}_1 (vgl. Bücher (2019): S. 4-5).

Um die Güte eines (randomisierten) statistischen Testverfahrens zu bewerten, werden für die Null- und Alternativhypotenosen \mathcal{H}_0 und \mathcal{H}_1 die Fehler erster Art und zweiter Art betrachtet. Der Fehler erster Art beschreibt eine Testentscheidung gegen \mathcal{H}_0 , obwohl \mathcal{H}_0 wahr ist. Der Fehler zweiter Art beschreibt eine Bestätigung von \mathcal{H}_0 , obwohl in Wahrheit \mathcal{H}_1 zugrundeliegt.

Als Schranke der Testentscheidung wird das Fünfprozentniveau $\alpha = 5\%$ gewählt. Unter angenommener Nullhypothese \mathcal{H}_0 wird für den Test $\varphi \in \Phi$ aus der Menge der betrachteten statistischen Tests Φ die Testentscheidung

$$\varphi(x) = \begin{cases} 1 \\ 0 \end{cases} \iff p(\varphi(x)) \begin{cases} \leq \\ > \end{cases} \alpha = 0.05$$

durchgeführt, wobei $p(\varphi(x))$ den p-Wert des Tests $\varphi(x)$ auf einer Stichprobe $x \in \mathcal{X}$ bezeichnet. Für die Bewertung und den Vergleich der einzelnen Testmethoden φ wird die Güte bestimmt über die empirischen Ablehnungsraten

$$\hat{\beta}_\varphi := \frac{1}{N} \sum_{k=1}^N \varphi(x_k) := \frac{1}{N} \sum_{k=1}^N \mathbf{1}_{\{p(\varphi(x_k)) \leq \alpha\}}.$$

Es bezeichnen $x_1, \dots, x_N \in \mathcal{X}$ die in den Iterationen generierten Stichproben. Erfüllt der Test φ in der Simulationsstudie die Abschätzung

$$\hat{\beta}_\varphi|_{\mathcal{H}_0} \leq \alpha \tag{6-1}$$

für die entsprechenden \mathcal{H}_0 -Copulas, so liegt empirisch gesehen ein Niveau- α -Test ($\varphi \in \Phi_\alpha$) vor. Ist zusätzlich in der Simulationsstudie

$$\hat{\beta}_\varphi|_{\mathcal{H}_1} \geq \alpha \tag{6-2}$$

zu beobachten, so ist der Test im Zusammenhang mit der Simulation als unverfälscht ($\varphi \in \Phi_{\alpha\alpha}$) zu bezeichnen. Aus den (unverfälschten) Niveau- α -Tests $\varphi \in \Phi_\alpha$ bzw. $\varphi \in \Phi_{\alpha\alpha}$ sind die Tests mit der stärksten Ablehnung im \mathcal{H}_1 -Fall auszuwählen. Diese Vorgehensweise basiert auf dem Konzept der UMP(U)-Tests (Uniformly Most Powerful (Unbiased) Tests) und Maximin-Tests, welche definiert sind als

$$\begin{aligned}\varphi^* \in \Phi_\alpha \text{ ist UMP} &\iff \beta_{\varphi^*}(\vartheta) = \sup_{\varphi \in \Phi_\alpha} \beta_\varphi(\vartheta) \quad (\vartheta \in \Theta|_{\mathcal{H}_1}), \\ \varphi^* \in \Phi_{\alpha\alpha} \text{ ist UMPU} &\iff \beta_{\varphi^*}(\vartheta) = \sup_{\varphi \in \Phi_{\alpha\alpha}} \beta_\varphi(\vartheta) \quad (\vartheta \in \Theta|_{\mathcal{H}_1}), \\ \varphi^* \in \Phi_\alpha \text{ ist Maximin} &\iff \inf_{\vartheta \in \Theta|_{\mathcal{H}_1}} \beta_{\varphi^*}(\vartheta) = \sup_{\varphi \in \Phi_\alpha} \inf_{\vartheta \in \Theta|_{\mathcal{H}_1}} \beta_\varphi(\vartheta).\end{aligned}$$

β_{φ^*} und β_φ sind die Gütfunktionen zu φ^* und φ . Diese Tests beschreiben die optimalen Tests für einen Sachverhalt, wobei Maximin-Tests schwächere Voraussetzungen zur Klassifikation als UMP- und UMPU-Tests haben (vgl. Rüschenhof (2014): S. 182; Bücher (2019): S. 78-81).

Folgend bezeichnet die Notation Φ_α ($\Phi_{\alpha\alpha}$) die Menge der Testmethoden, die aus empirischer Begründung den (unverfälschten) Niveau- α -Tests zugeordnet werden. Die Methoden werden bzgl. der Simulationsresultate den folgenden Kategorien zugeordnet.

1. Niveau- α -Tests: *Die Simulationsresultate für die Testmethode $\varphi \in \Phi$ erfüllen bzgl. \mathcal{H}_0 -Copulas die Ungleichung (6-1), sodass empirisch ein Niveau- α -Test vorliegt, d. h. $\varphi \in \Phi_\alpha$.*
2. Unverfälschte Niveau- α -Tests: *Die Resultate der Testmethode $\varphi \in \Phi_\alpha$ erfüllen im \mathcal{H}_1 -Fall die Ungleichung (6-2), sodass empirisch ein unverfälschter Niveau- α -Test vorliegt, d. h. $\varphi \in \Phi_{\alpha\alpha}$.*
3. Unverfälschte Tests: *Die Resultate für die Testmethode $\varphi \in \Phi$ erfüllen im \mathcal{H}_0 -Fall die Ungleichungen (6-1) und (6-2) zu einem höheren (liberalerem) Niveau $\tilde{\alpha} > \alpha$, sodass $\varphi \in \Phi_{\tilde{\alpha}\tilde{\alpha}}$ gilt.*

Die Testmethoden, die unter diese drei Kategorien fallen, werden überprüft, ob sie Maximin-Tests oder UMP- bzw. UMPU-Tests beschreiben. Falls dies erfüllt ist, kann eine Methodenempfehlung ausgesprochen werden.

Vorgehensweise zur Bemessung der Güte und zum Vergleich der Schätzverfahren zur Modellanpassung

Bei der Schätzung des Parameters ϑ_0 der jeweiligen Copula $C = C_{\vartheta_0}$ unter angenommener Zugehörigkeit zur Familie $C_{\vartheta_0} \in \mathcal{C}_0 := \{C_\vartheta : \vartheta \in \Theta\}$ wird

die Güte eines Schätzers $\hat{\vartheta}_n$ über den Bias und die mittlere quadratische Abweichung (MSE) bemessen. Die formalen Definitionen ebendieser lauten

$$\text{Bias}(\hat{\vartheta}_n|\vartheta_0) := \mathbb{E}[\hat{\vartheta}_n - \vartheta_0|\vartheta_0], \quad \text{MSE}(\hat{\vartheta}_n|\vartheta_0) := \mathbb{E}[(\hat{\vartheta}_n - \vartheta_0)^2|\vartheta_0]$$

und die zugehörigen empirischen Versionen unter bekanntem wahren Parameter ϑ_0 sind definiert als

$$\begin{aligned}\widehat{\text{Bias}}_N(\hat{\vartheta}_n|\vartheta_0) &:= \frac{1}{N} \sum_{k=1}^N (\hat{\vartheta}_n(x_k) - \vartheta_0), \\ \widehat{\text{MSE}}_N(\hat{\vartheta}_n|\vartheta_0) &:= \frac{1}{N} \sum_{k=1}^N (\hat{\vartheta}_n(x_k) - \vartheta_0)^2,\end{aligned}$$

wobei $x_1, \dots, x_N \in \mathcal{X}$ die in den Iterationen generierten Stichproben der Größe n bezeichnen. Das erste Kriterium zur Auswahl der Methoden ist die Unverzerrtheit, die über $\text{Bias}(\hat{\vartheta}_n|\vartheta_0) = 0$ charakterisiert wird. Entsprechend ist eine Methode mit möglichst geringer Verzerrung, d. h. $\hat{\vartheta}_n$ mit $\|\widehat{\text{Bias}}_N(\hat{\vartheta}_n|\vartheta_0)\| \in [0, \infty)$ minimal, auszuwählen.

Gleichzeitig ist auch der MSE minimal zu halten, da hierüber die Fehler der Schätzung gesteuert werden. Bezeichne

$$D_\varepsilon := \{\hat{\vartheta}_n \mid \widehat{\text{Bias}}_N(\hat{\vartheta}_n|\vartheta_0) \in (-\varepsilon, +\varepsilon)\}$$

die Menge der Schätzmethoden $\hat{\vartheta}_n : \mathcal{X} \rightarrow \Theta$ mit einem empirischen Bias im Intervall $(-\varepsilon, +\varepsilon)$ für ein $\varepsilon > 0$. Im Vergleich aller Methoden in D_ε untereinander ist die Schätzmethode mit minimalem MSE auszuwählen.

Diese Vorgehensweise basiert auf den UMVU-Schätzern (Uniformly Minimum Variance Unbiased). Sei D_ϑ eine Menge von unverzerrten quadratintegrierbaren Schätzern für den Parameter $\vartheta \in \Theta$. Ein Schätzer $\hat{\vartheta}_n \in D_\vartheta$, welcher

$$\text{MSE}(\hat{\vartheta}_n|\vartheta) = \inf_{\hat{\theta}_n \in D_\vartheta} \text{MSE}(\hat{\theta}_n|\vartheta) \quad \forall \vartheta \in \Theta$$

erfüllt, wird UMVU-Schätzer genannt. UMVU-Schätzer beschreiben die optimalen Schätzer (vgl. Bücher (2019): S. 14, 16-17). Im Folgenden wird statt dem MSE der RMSE betrachtet, dieser wird aus dem MSE berechnet als

$$\text{RMSE}(\dots) := \sqrt{\text{MSE}(\dots)}$$

und dies gilt auch für die empirischen Varianten.

<i>Übersicht der Methoden zur Mitigation von Bindungen bei den rangbasierten Tests auf Radial-Symmetrie und auf Austauschbarkeit. Für jede Methode wird aus den rechtsstehenden Modifikationen der (Pseudo-)Stichproben, den MPL-Schätzverfahren und den Testverfahren ausgewählt. Haken (✓) geben die Auswahl an.</i>	Modifikationen			<i>Tests nach Genest et al. (2012) u. Genest und Nešlehová (2013) Test nach Kojadinovic (2017)</i>
	<i>Reduktion bzgl. voll. Bindungen</i> <i>Reduktion mit Greedy-Heuristik</i>	<i>Zufällige Rangvergabe</i> <i>Jittering mit i.i.d. Rauschen</i>	<i>Jittering, empirisch angepasst</i>	
Referenzmethode (<i>Reference</i>)	- -	- - -	- - -	✓ -
Reduktion bzgl. vollst. Bindungen (<i>V-Reduct.</i>)	✓ -	- - -	- - -	✓ ✓ -
Reduktion mit Greedy-Heuristik (<i>G-Reduct.</i>)	- ✓	- - -	- - -	✓ ✓ -
Zufällige Rangvergabe (<i>Rnd-Ranks</i>)	- -	✓ - -	- - -	✓ ✓ -
Jittering mit i.i.d. Rauschen (<i>Π-Jitter</i>)	- -	- ✓ -	- - -	✓ ✓ -
Jittering, empirisch angepasst ($\text{mix}_{\hat{\tau}_n}\text{-Jitter}$)	- -	- - -	✓ - -	✓ - -
Bootstrap nach Kojadinovic (2017) (<i>Kojadinovic</i>)	- -	- - -	- - -	✓
<i>V-Reduct.</i> & <i>Rnd-Ranks</i> *	✓ -	✓ - -	- - -	✓ -
<i>G-Reduct.</i> & <i>Rnd-Ranks</i> *	- ✓	✓ - -	- - -	✓ -

Tabelle 6.1: Übersicht der Methoden zur Mitigation von Bindungen bei den rangbasierten Tests auf Radial-Symmetrie und auf Austauschbarkeit zur Anwendung in der Simulationsstudie. Haken (✓) bedeuten den Einbezug der Modifikation. *: Nach der Reduktion werden die Ränge der verbleibenden Bindungen zufällig verteilt. Quelle: Eigene Erstellung.

6.2 Simulation zum Methodenvergleich bzgl. Radial-Symmetrie und Austauschbarkeit

Die Simulationsstudie zur Radial-Symmetrie betrachtet die Hypothesen

$$\mathcal{H}_{0,\text{RS}} : C(u_1, u_2) = \bar{C}(u_1, u_2) \quad \text{und} \quad \mathcal{H}_{1,\text{RS}} : C(u_1, u_2) \neq \bar{C}(u_1, u_2)$$

für $u = (u_1, u_2) \in [0, 1]^2$. Für die Radial-Symmetrie-Hypothese \mathcal{H}_0 werden die Gaus'sche Copula $C_\vartheta^{\text{Gauß}}$ und die t-Copula $C_{\vartheta,4}^t$ betrachtet. Die Hypothese \mathcal{H}_1 wird über die Clayton-Copula C_ϑ^{CL} und der Gumbel-Copula C_ϑ^{GH} abgedeckt. Bzgl. der Austauschbarkeit lauten die beiden Hypothesen

$$\mathcal{H}_{0,\text{EX}} : C(u_1, u_2) = C(u_2, u_1) \quad \text{und} \quad \mathcal{H}_{1,\text{EX}} : C(u_1, u_2) \neq C(u_2, u_1).$$

Die $\mathcal{H}_{0,\text{EX}}$ -Copulas gehören zu den Clayton- und Gumbel-Copulafamilien. Für die Hypothese $\mathcal{H}_{1,\text{EX}}$ werden die Khoudraji-Versionen der Clayton- und Gumbel-Copulas betrachtet.

Die Simulationsstudie für beide statistische Tests umfasst die in Tabelle 6.1 aufgelisteten Methoden und Methodenkombinationen. Diese werden auf unmodifizierte und modifizierte Daten angewandt. Die unmodifizierten Daten werden bzgl. der gewählten Copulas generiert. Bei den modifizierten Daten wurden die Rahmenbedingungen der Finanzmärkte berücksichtigt und künstlich Bindungen herbeigeführt.

Die Simulationsresultate der Tests über Radial-Symmetrie sind in den Abbildungen 6.1 bis 6.4 sowie in den Tabellen A.1 und A.2 im Anhang aufbereitet. Dort sind die Ablehnungsraten zum Fünfprozentniveau dargestellt und die Tests sind bzgl. der jeweiligen Copula und dem Kendall'schen Tau sowie der Anpassungen im datengenerierenden Algorithmus gruppiert.

Bzgl. der Austauschbarkeit sind die Ablehnungsraten zum Niveau $\alpha = 5.0\%$ in den Abbildungen 6.5 bis 6.8 bzw. in den Tabellen A.3 bis A.4 im Anhang aufbereitet. Auch hier sind die Tests bzgl. der jeweiligen Copula und dem Kendall'schen Tau sowie der Anpassungen im datengenerierenden Algorithmus gruppiert.

Bei den Tests auf Radialsymmetrie ist initial zu sagen, dass alle Methoden auf unmodifizierten Stichproben unverfälschte Niveau- α -Tests zum Niveau $\alpha = 5\%$ darstellen. Für unmodifizierte Daten und dem Kendall'schen Tau $\tau = 0.25$ stellen die Tests auf Austauschbarkeit unverfälschte Tests zum Niveaus zwischen 3.3 % und 3.7 % dar. Für $\tau \in \{0.25, 0.50\}$ sind diese Methoden unverfälschte Niveau- α -Tests, $\alpha = 5.0\%$. Bei beiden Sachverhalten generiert die *Kojadinovic*-Methode die höchsten Ablehnungsraten. Die Methoden *Reference*, *V-Reduct.*, *G-Reduct.* und *Rnd-Ranks* sowie die kombinierten Methoden haben niedrigere Ablehnungsraten und diese stimmen auf ungebundenen Daten quasi überein.

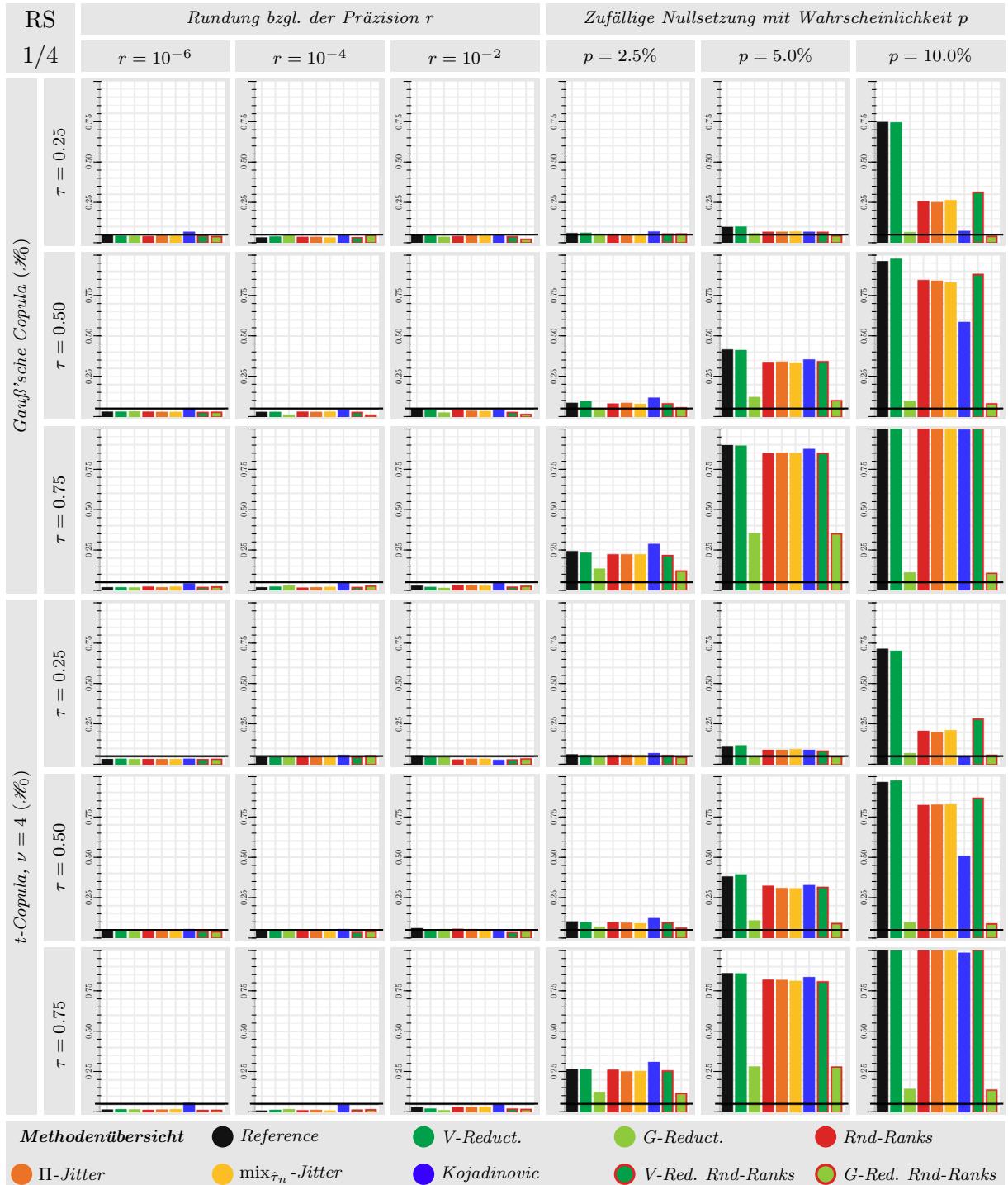


Abbildung 6.1: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Radial-Symmetrie, Teil 1 von 4: Aufbereitete Testentscheidungen der Tests auf Radial-Symmetrie für die \mathcal{H}_0 -Copulas (Gauß'sche Copula, t-Copula) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0\%$. Quelle: Eigene Erstellung.

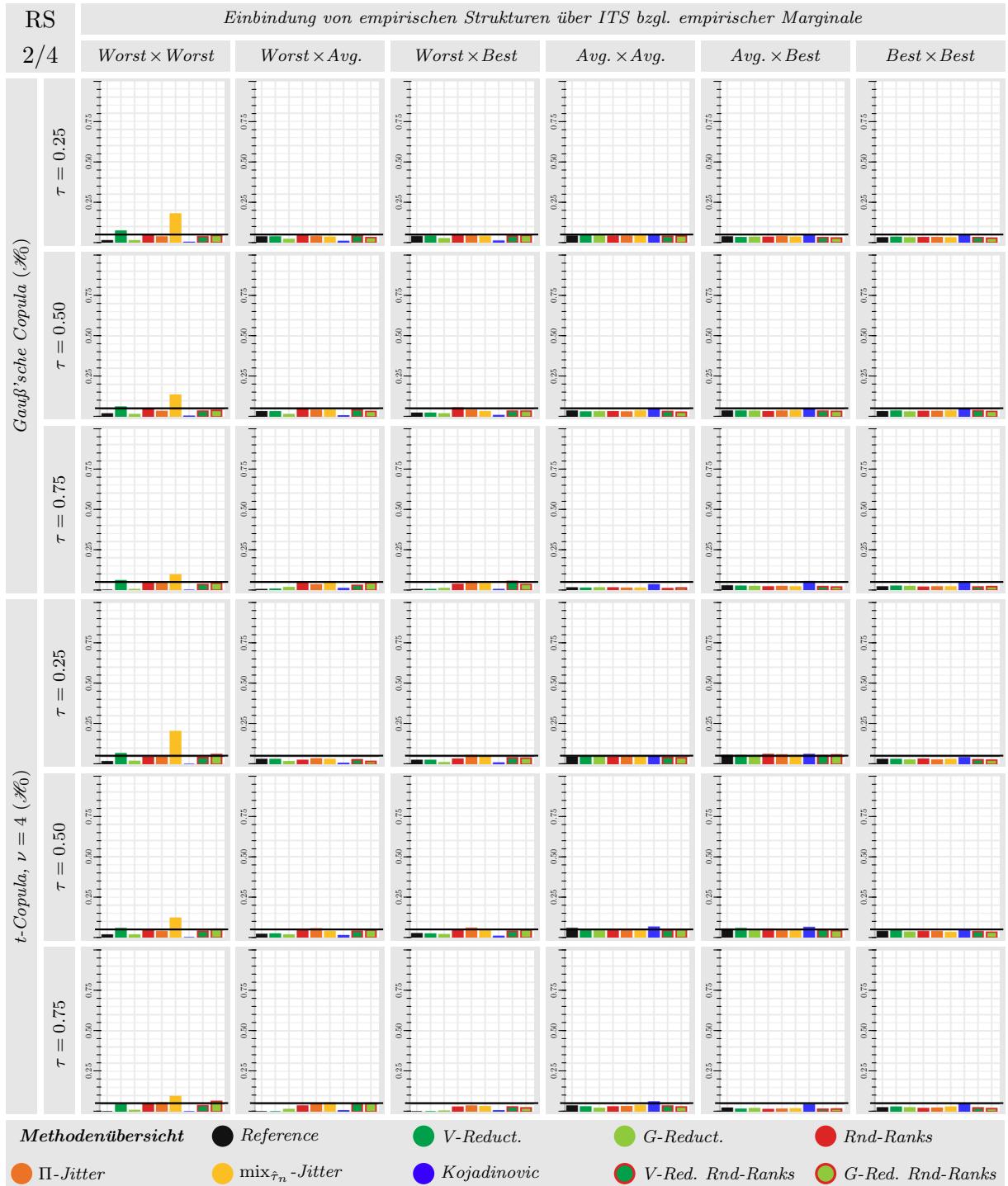


Abbildung 6.2: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Radial-Symmetrie, Teil 2 von 4: Aufbereitete Testentscheidungen der Tests auf Radial-Symmetrie für die \mathcal{H}_0 -Copulas (Gauß'sche Copula, t-Copula) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0 \%$. Quelle: Eigene Erstellung.

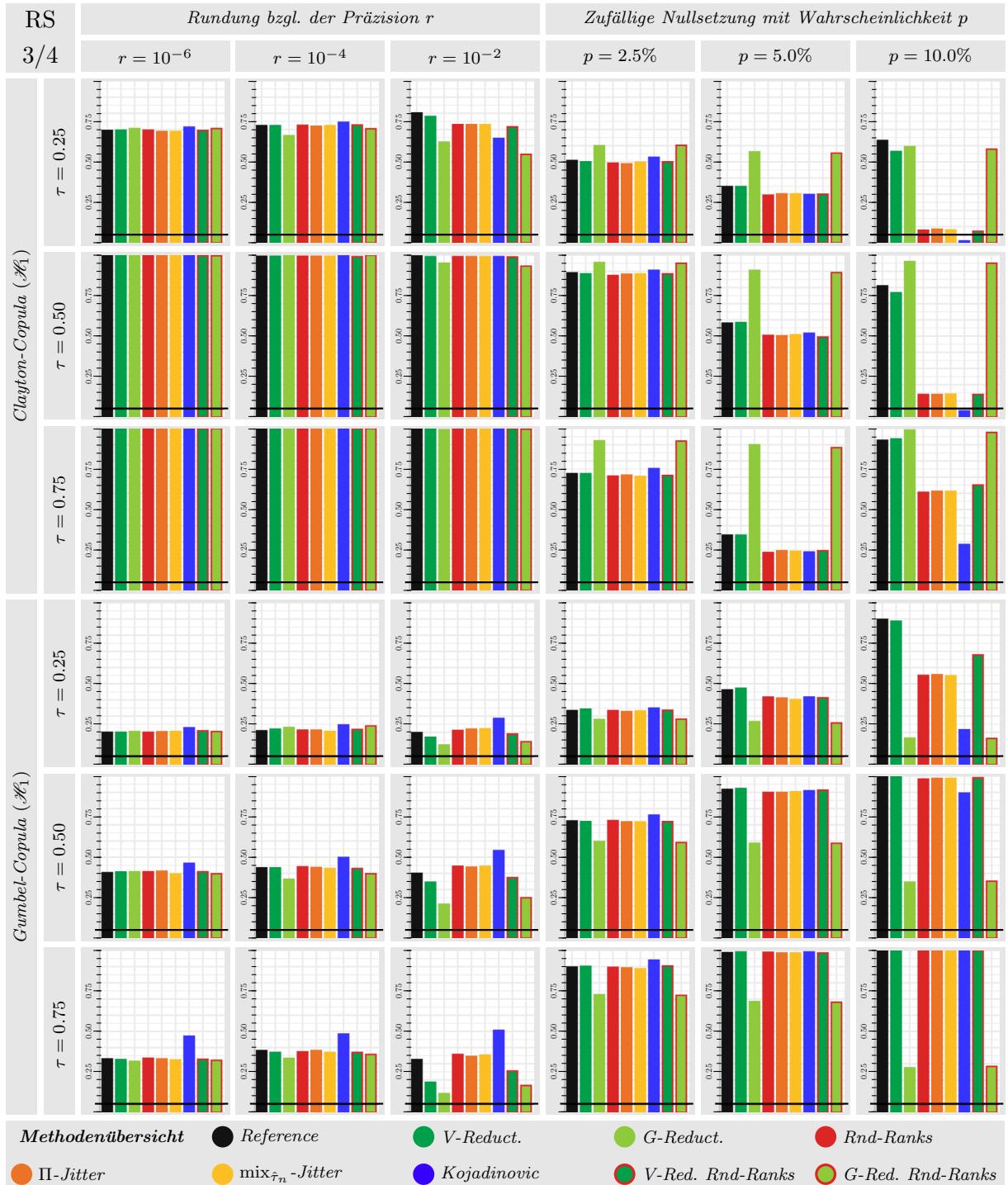


Abbildung 6.3: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Radial-Symmetrie, Teil 3 von 4: Aufbereitete Testentscheidungen der Tests auf Radial-Symmetrie für die \mathcal{H}_1 -Copulas (Clayton-Copula, Gumbel-Copula) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0\%$. Quelle: Eigene Erstellung.

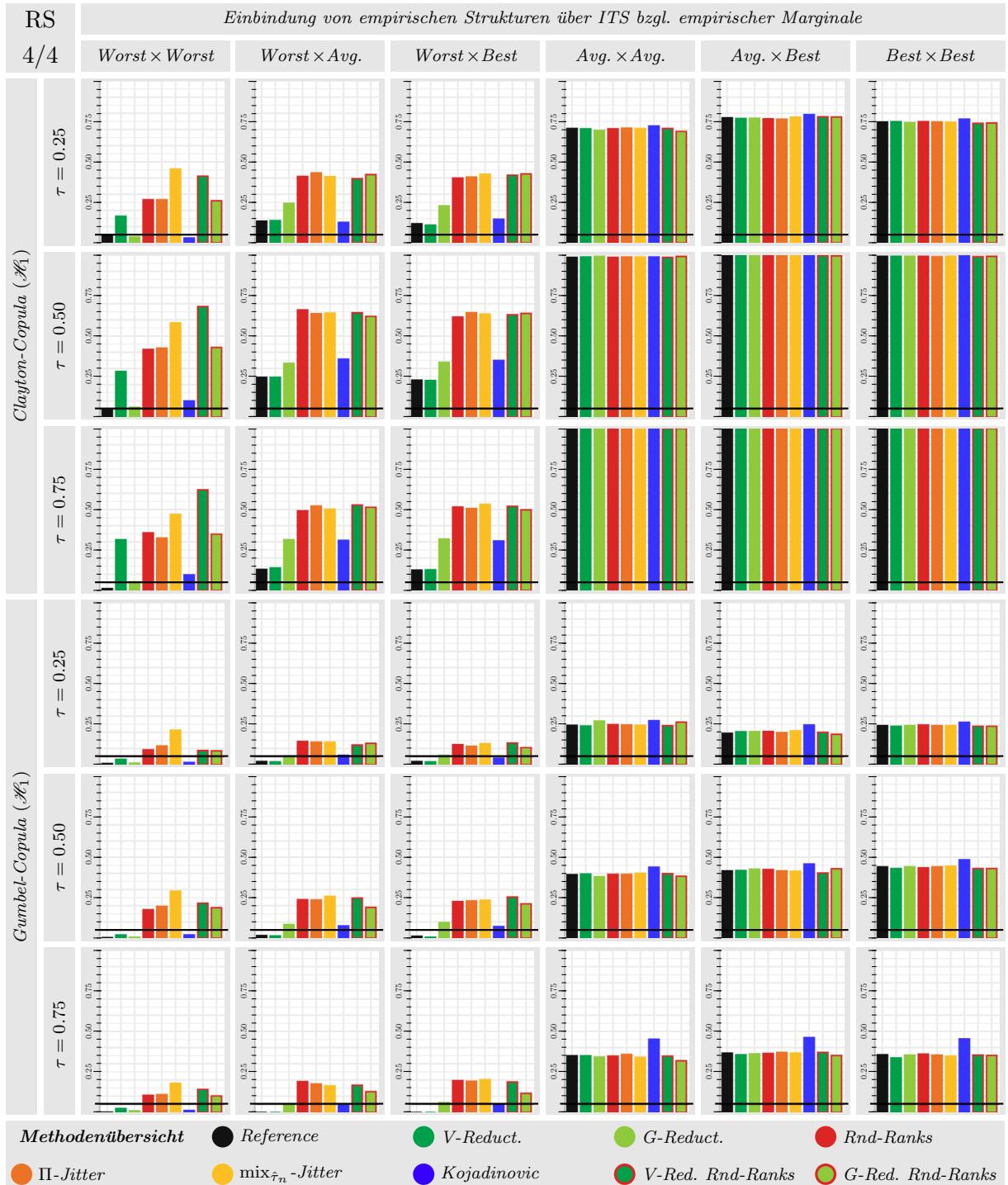


Abbildung 6.4: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Radial-Symmetrie, Teil 4 von 4: Aufbereitete Testentscheidungen der Tests auf Radial-Symmetrie für die \mathcal{H}_1 -Copulas (Clayton-Copula, Gumbel-Copula) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0 \%$. Quelle: Eigene Erstellung.

Simulationsergebnisse bzgl. Radial-Symmetrie für die Anwendung bei Daten mit unbekannten Modifikationen

Es werden zu Beginn die Methoden unter Berücksichtigung aller Datenmodifikationen zum Niveau $\alpha = 5.0\%$ geprüft. Für die \mathcal{H}_0 -Copulas mit $\tau = 0.25$ sind die Methoden *G-Reduct.*, *Kojadinovic* und *G-Reduct. & Rnd.-Ranks* unverfälschte Niveau- α -Tests. Für $\tau = 0.50$ gilt diese Aussage für *G-Reduct.* und im Fall $\tau = 0.75$ stellt keine der Methoden für $\alpha = 5\%$ einen Niveau- α -Tests über Radial-Symmetrie dar. Die beiden Methoden *G-Reduct.* und *G-Reduct. & Rnd.-Ranks* sind für alle betrachteten Taus unverfälschte Tests zu den höheren Niveaus 6.5 % und 7.5 %. Die anderen Methoden sind unverfälscht, d. h. es gilt $\hat{\beta}|_{\mathcal{H}_0} < \tilde{\alpha} \leq \hat{\beta}|_{\mathcal{H}_1}$ für ein $\tilde{\alpha}$. Die Niveaus dieser Methoden liegen zwischen 18.8 % und 19.7 % bei Berücksichtigung aller Kendall'schen Taus.

Simulationsergebnisse bzgl. Radial-Symmetrie für die Anwendung bei Daten mit modifizierter Messpräzision

Bei den \mathcal{H}_0 -Copulas unter Anpassung der Messpräzision sind alle Methoden unverfälschte Niveau- α -Tests für $\alpha = 6.4\%$. In 150 der 162 durchgeführten Simulationen gilt diese Aussage auch für das Fünfprozentniveau; die Methoden dieser Simulationen umfassen *V-Reduct.*, *mix _{$\hat{\tau}_n$} -Jitter*, und *V-Reduct. & Rnd-Ranks*. Aus diesen Methoden sind diejenigen mit der höchsten Ablehnungsrate im \mathcal{H}_1 -Fall zu wählen. Die minimalen Ablehnungsraten dieser drei Methoden im \mathcal{H}_1 -Fall lauten 17.0 %, 20.6 % und 20.5 %. Es ist anzumerken, dass bei der Clayton-Copula weitaus höhere Raten, genauer 69.4 %, 68.8 % und 69.4 %, gemessen wurden. Somit stellt die Methode *mix _{$\hat{\tau}_n$} -Jitter* einen Maximin-Test zur Anwendung bei Daten mit angepasster Messpräzision dar.

Simulationsergebnisse bzgl. Radial-Symmetrie für die Anwendung bei Daten mit zufälliger Nullsetzung

Bei der zufälligen Nullsetzung der Daten ist zu beobachten, dass die Ablehnungsraten von hohen Werten des Kendall'schen Taus und/oder einer hohen Nullsetz-Wahrscheinlichkeiten p beeinflusst werden. Für $\tau = 0.25$ und $p \in \{2.5\%, 5.0\%, 10.0\%\}$ sind die Ablehnungsraten bis auf wenige Ausnahmen größer als fünf Prozent. Die Methoden *G-Reduct.*, *Kojadinovic* und *G-Reduct. & Rnd.-Ranks* haben dort die niedrigsten Ablehnungsraten unter \mathcal{H}_0 : Mit Ausnahme von einer Rate des Kojadinovic-Ansatzes liegen die Werte unterhalb 7.0 %, und die gemischte Methode *G-Reduct. & Rnd.-Ranks* hält das Niveau 5.8%. Die anderen Methoden haben im Fall $p = 10\%$ Ablehnungsraten von der Größenordnung 20 % bis 75 %. Für die höheren Werte des

Kendall'schen Taus sind die niedrigsten Niveaus erneut bei den Methoden *G-Reduct.* und *G-Reduct. & Rnd.-Ranks* zu finden. Für $\tau = 0.50$ ($\tau = 0.75$) liegen die Ablehnungsraten unterhalb 12.0 % und 10.2% (35.0 % und 28.0 %). Das nächstkleinere Niveau der anderen Methoden ist bei *Kojadinovic* zu finden; das Niveau liegt bei 58.4 % (99.0 %), was unter diese Rahmenbedingung gegen eine Handlungsempfehlung dieser Methoden spricht. Bei den \mathcal{H}_1 -Copulas unter zufälliger Nullsetzung sind folgende Simulationsergebnisse zu nennen: Bei der Clayton-Copula sind die minimalen Ablehnungsraten der Methoden *G-Reduct.* und *G-Reduct. & Rnd.-Ranks* 56.2 % und 55.4 %, wobei für $\tau \in \{0.50, 0.75\}$ die Werte 90.6 % und 89.2 % vorliegen.

Insgesamt liefert die Methode *G-Reduct.* die besten Ergebnisse. Es wird eine allgemeine Empfehlung ausgesprochen; es ist zu beachten, dass die Methode kein konsistentes Verhalten aufzeigt. Die Anwendung auf Daten mit hoher Abhängigkeit und Nullsetzung mit hoher Wahrscheinlichkeit wird nicht empfohlen.

Simulationsergebnisse bzgl. Radial-Symmetrie für die Anwendung bei Daten mit empirischer Transformation

Bei den empirisch transformierten Daten sind folgende Resultate vorliegend: Keine der Testmethoden hält im \mathcal{H}_0 -Fall das Fünfprozentniveau, da bei der t-Copula für $\tau = 0.25$ und empirischer Transformationen vom Typ *Average* \times *Best* die Ablehnungsraten zwischen 5.3 % und 6.0 % liegen. Die Testmethoden *Reference*, *G-Reduct.*, *Rnd-Ranks* und *Π-Jitter* sowie *V-Red.* & *Rnd-Ranks* und *G-Red.* & *Rnd-Ranks* sind Niveau- α -Tests für $\alpha = 6.0\%$.

Davon sind die Methoden *Rnd-Ranks*, $\text{mix}_{\hat{\tau}_n}\text{-Jitter}$, *V-Red.* & *Rnd-Ranks* und *G-Red.* & *Rnd-Ranks* bzgl. $\alpha = 6.0\%$ unverfälscht. Die minimalen Ablehnungsraten im \mathcal{H}_1 -Fall betragen 9.2 % für *Rnd-Ranks*, 13.0 % für $\text{mix}_{\hat{\tau}_n}\text{-Jitter}$ und 8.8 % bzw. 8.6 % für die kombinierten Methoden.

Entsprechend ist zur Anwendung auf empirisch transformierte Daten die Nutzung von $\text{mix}_{\hat{\tau}_n}\text{-Jitter}$ zu empfehlen, das dies der Maximin-Test für diese Rahmenbedingungen darstellt.

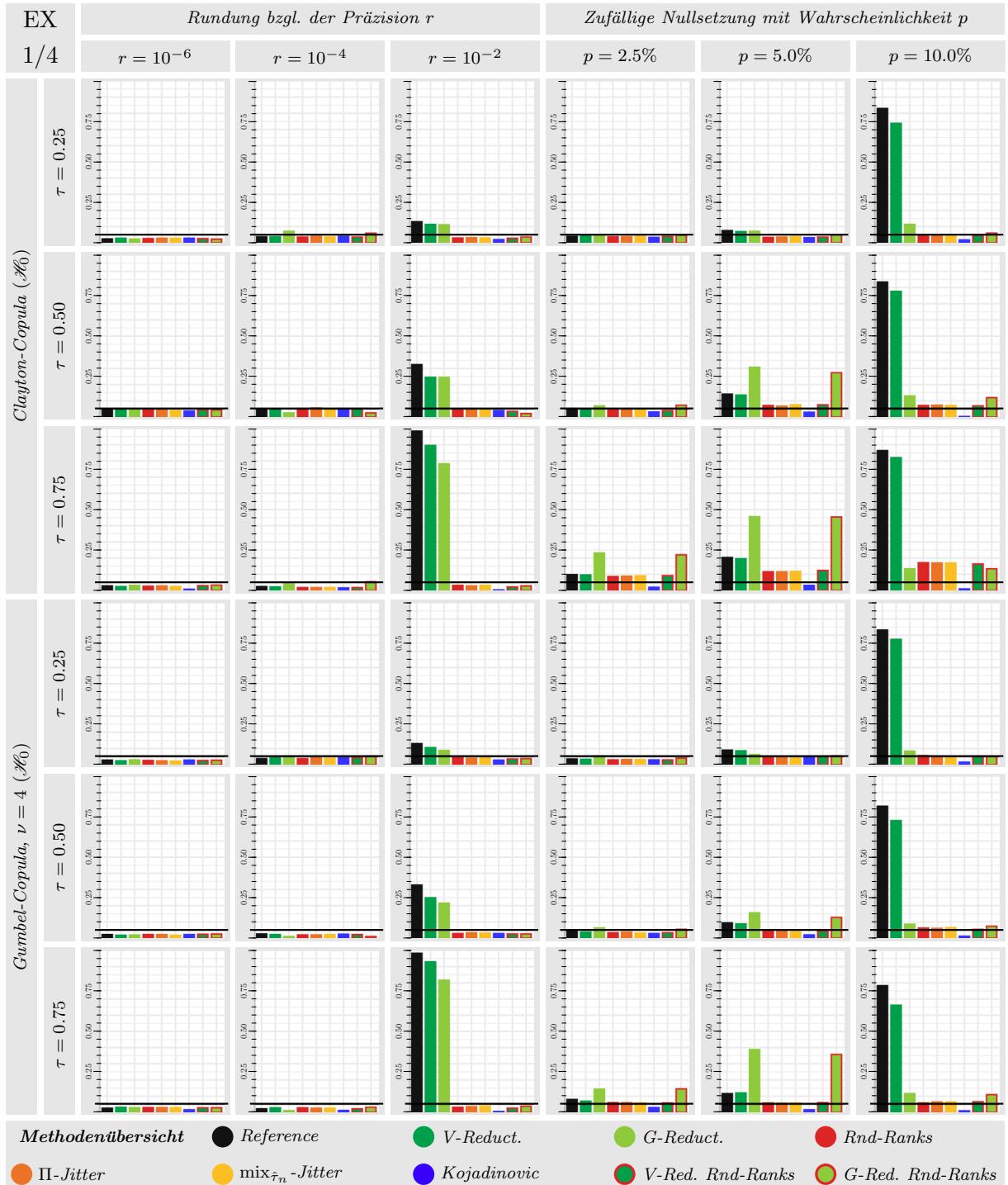


Abbildung 6.5: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Austauschbarkeit, Teil 1 von 4: Aufbereitete Testentscheidungen der Tests auf Austauschbarkeit für die \mathcal{H}_0 -Copulas (Clayton-Copula, Gumbel-Copula) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0 \%$. Quelle: Eigene Erstellung.

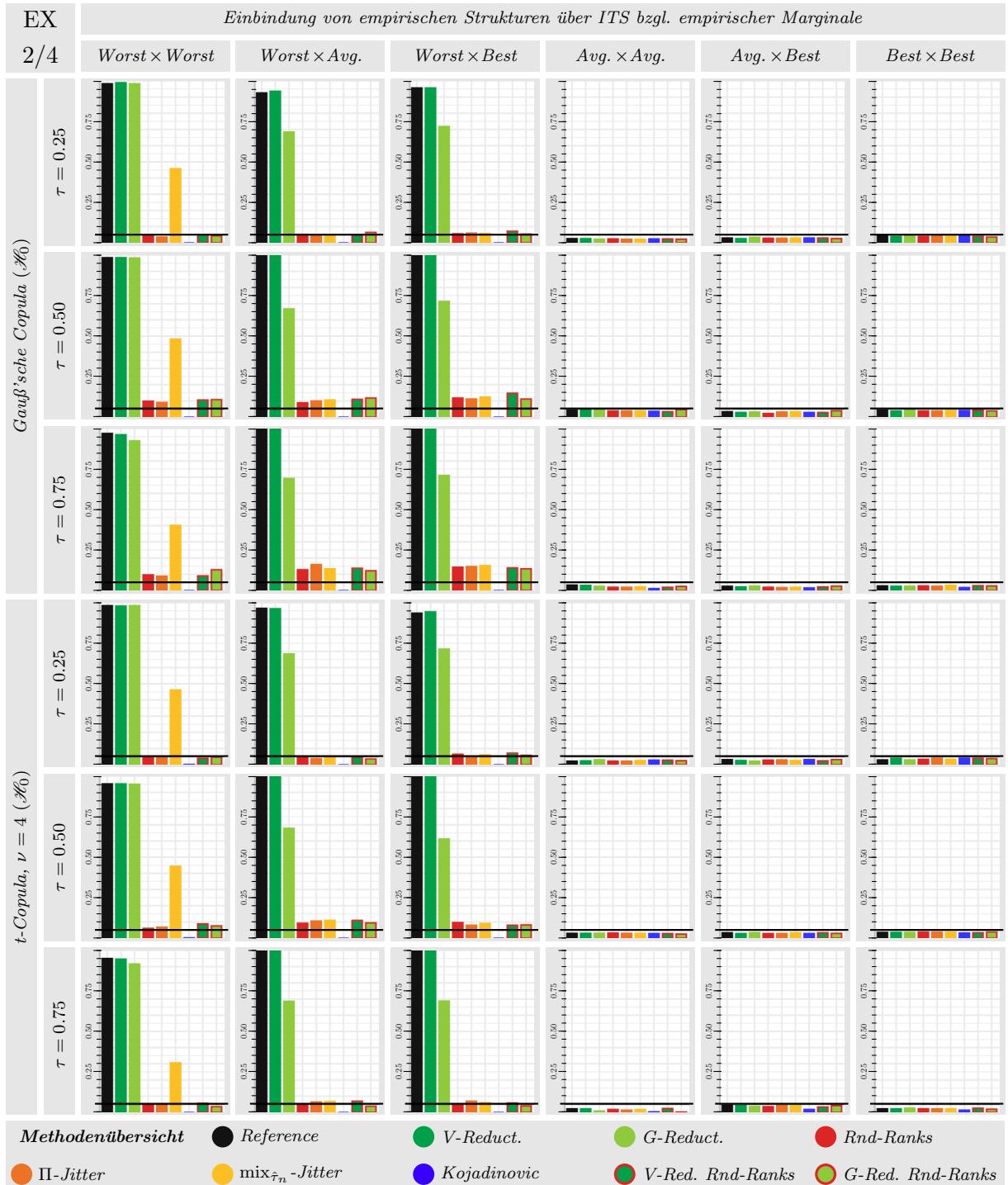


Abbildung 6.6: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Austauschbarkeit, Teil 2 von 4: Aufbereitete Testentscheidungen der Tests auf Austauschbarkeit für die \mathcal{H}_0 -Copulas (Clayton-Copula, Gumbel-Copula) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0 \%$. Quelle: Eigene Erstellung.

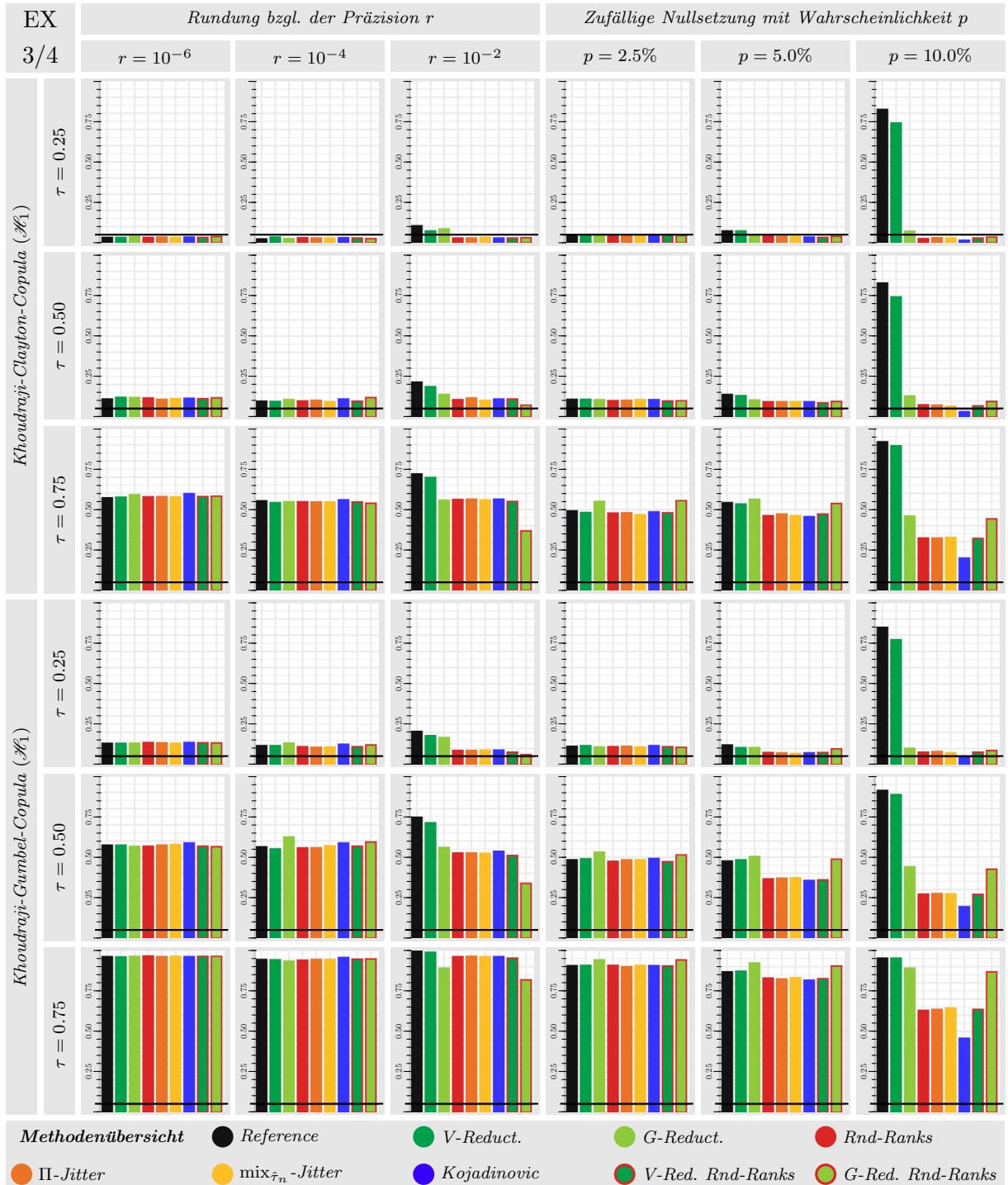


Abbildung 6.7: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Austauschbarkeit, Teil 3 von 4: Aufbereitete Testentscheidungen der Tests auf Austauschbarkeit für die \mathcal{H}_1 -Copulas (Khoudraji-Clayton, Khoudraji-Gumbel) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0 \%$. Quelle: Eigene Erstellung.

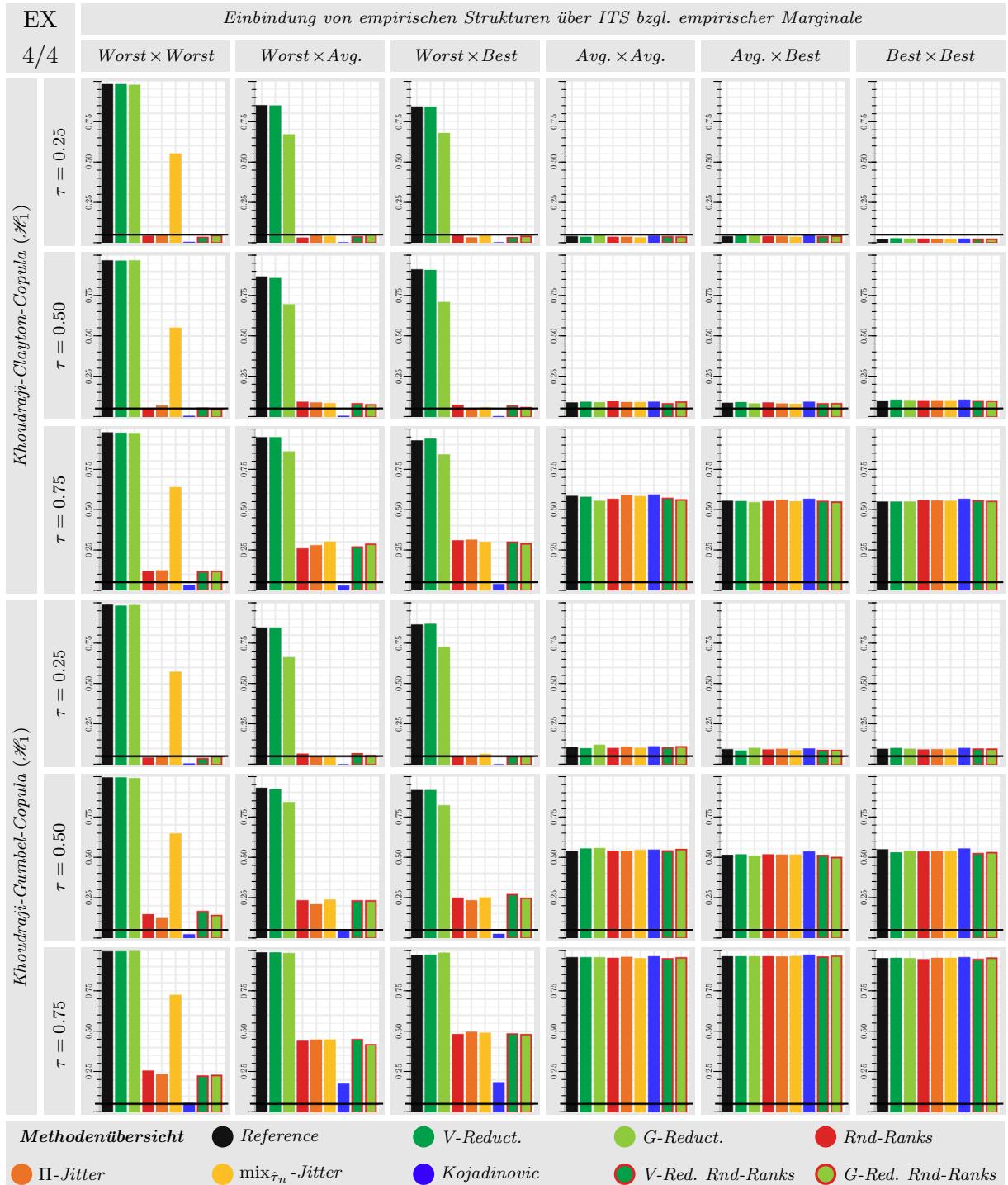


Abbildung 6.8: \mathcal{H}_0 -Ablehnungsraten für die rangbasierten Tests auf Austauschbarkeit, Teil 4 von 4: Aufbereitete Testentscheidungen der Tests auf Austauschbarkeit für die \mathcal{H}_1 -Copulas (Khoudraji-Clayton, Khoudraji-Gumbel) zum Niveau von 5.0 %. Visualisiert sind die in der Simulationsstudie (Iterationen: $N = 500$, Stichprobengröße: $n = 250$) gemessenen Ablehnungsraten. Die horizontale Linie kennzeichnet das Niveau von $\alpha = 5.0 \%$. Quelle: Eigene Erstellung.

Simulationsergebnisse bzgl. Austauschbarkeit für die Anwendung bei Daten mit unbekannten Modifikationen

Es werden zu Beginn die Methoden unter Berücksichtigung aller Datenmodifikationen geprüft. Bei einer Betrachtung aller $\tau \in \{0.25, 0.50, 0.75\}$ existiert kein unverfälschter Niveau- α -Test zum Fünfprozentniveau. Der einzige unverfälschte Niveau- α -Test für alle betrachteten Taus ist mit *Kojadinovic* zum Niveau $\alpha = 2,5\%$ gegeben.

Beim Ausschließen der Khoudraji-Clayton-Copula mit $\tau = 0.25$ aus der Auswertung kommen weitere Tests in Betrachtung. Die *Kojadinovic*-Methode ist mit dieser Einschränkung ein unverfälschter Niveau- α -Test für $\alpha = 5.0\%$, wobei die maximale Ablehnungsrate im \mathcal{H}_0 -Fall bei 2.5% verbleibt. Die minimale Ablehnungsrate unter \mathcal{H}_1 wächst von 4.3% auf 10.7%. Die Methoden *Rnd-Ranks*, *Π-Jitter*, *mix _{$\hat{\tau}_n$} -Jitter* und *V-Reduct. & Rnd-Ranks* ergeben nun unverfälschte Niveau- α -Tests mit den respektiven maximalen Ablehnungsralten 7.4%, 7.6%, 10.1% und 7.4% im \mathcal{H}_0 -Fall sowie den minimalen Ablehnungsralten 8.6%, 8.7%, 12.8% und 8.6% im \mathcal{H}_1 -Fall.

Insgesamt ist eine Empfehlung für den Test auf Austauschbarkeit nach *Kojadinovic* auszusprechen. Dieser Test ist der einzige für alle Copulas und Taus unverfälschte Test. Bzgl. der Einschränkung liegt ein unverfälschter Niveau- α -Test zum Fünfprozentniveau vor.

Simulationsergebnisse bzgl. Austauschbarkeit für die Anwendung bei Daten mit modifizierter Messpräzision

Wie im allgemeinen Fall stellt die Khoudraji-Clayton-Copula mit $\tau = 0.25$ einen Sonderfall in der Auswertung dar, da die Ergebnisse für diesen Fall keine unverfälschten Tests aufzeigen.

Ausgenommen der Khoudraji-Clayton-Copula mit $\tau = 0.25$ sind die Testmethoden *Rnd-Ranks* und *mix _{$\hat{\tau}_n$} -Jitter* unverfälschte Niveau- α -Tests für $\alpha = 5.0\%$. Die minimalen Ablehnungsralten im \mathcal{H}_1 -Fall betragen 8.8% und 9.0%. Die Methoden *Π-Jitter*, *Kojadinovic* und *V-Reduct. & Rnd-Ranks* sind unverfälschte Tests mit den Niveaus 5.6%, 5.2% und 5.2%. Die minimalen Ablehnungsralten unter \mathcal{H}_1 betragen jeweils 8.8%, 9.0% und 7.80%.

Insgesamt ist also die Methode *mix _{$\hat{\tau}_n$} -Jitter* zur Anwendung auf Daten mit angepasster Messpräzision zu empfehlen, da diese Testmethode der Maximin-Test zum Fünfprozentniveau ist.

Simulationsergebnisse bzgl. Austauschbarkeit für die Anwendung bei Daten mit zufälliger Nullsetzung

Für die Khoudraji-Clayton-Copula mit $\tau = 0.25$ stellt keine der Methoden einen unverfälschten Test dar, weswegen dieser Fall im Folgenden ausgelassen wird. In einer Gesamtbetrachtung konnten keine unverfälschten Tests festgestellt werden. Entsprechend werden folgend die Paare aus einer \mathcal{H}_0 -Copula und der Khoudraji-Version einzeln betrachtet. Für (Khoudraji-)Clayton-Copulas mit $\tau \in \{0.50, 0.75\}$ ist der Test auf Austauschbarkeit nach *Kojadinovic* der einzige unverfälschte Niveau- α -Test für $\alpha = 3.2\%$. Für die (Khoudraji-)Gumbel-Copula mit $\tau \in \{0.25, 0.50, 0.75\}$ sind fünf unverfälschte Testmethoden identifizierbar. *Kojadinovic* ist der einzige unverfälschte Niveau- α -Test für $\alpha = 5.0\%$. Die Methoden *Rnd-Ranks*, *Π-Jitter*, $\text{mix}_{\hat{\tau}_n}$ -*Jitter* sind unverfälschte Niveau- α -Tests zu den jeweiligen Niveaus 6.2%, 6.4%, 6.4% und 6.6%. Hierbei hat die Methode *Rnd-Ranks* die größte Differenz $7.6\% - 6.2\% = 1.4\%$ zwischen der minimalen Ablehnungsrate im \mathcal{H}_1 -Fall und der maximalen Ablehnungsrate im \mathcal{H}_0 -Fall.

Insgesamt ist hier die *Kojadinovic*-Methode zu empfehlen, wobei die Anwendung von *Rnd-Ranks* bzgl. der (Khoudraji-)Gumbel-Copula plausibel ist.

Simulationsergebnisse bzgl. Austauschbarkeit für die Anwendung bei Daten mit empirischer Transformation

Bei den Simulationen mit empirisch transformierten Daten existieren vier unverfälschte Niveau- α -Tests: *Rnd-Ranks*, *Π-Jitter*, *V-Reduct.* & *Rnd-Ranks* und *G-Reduct.* & *Rnd-Ranks*, mit den jeweiligen maximalen Ablehnungsraten 6.4%, 6.0%, 7.2% und 6.4% unter \mathcal{H}_0 , und minimalen Ablehnungsraten 26.6%, 26.6%, 39.6% und 26.0% unter \mathcal{H}_1 . *Π-Jitter* ist zum gewählten niedrigsten Niveau 6.0% Maximin-optimal.

Beim Ausschluss des Typs *Worst* \times *Worst* für die empirischen Transformationen gilt dies ebenso für die Methoden $\text{mix}_{\hat{\tau}_n}$ -*Jitter* und *Kojadinovic*. Zum gewählten Niveau $\alpha = 6.0\%$ ist $\text{mix}_{\hat{\tau}_n}$ -*Jitter* Maximin-optimal mit der maximalen Ablehnungsrate 6.0% unter \mathcal{H}_0 und der minimalen Ablehnungsrate 40.8% unter \mathcal{H}_1 . Der zweitplazierte optimale Test ist *Π-Jitter*, wobei hier die minimale Ablehnungsrate unter \mathcal{H}_1 mit 40.6% leicht unterliegt. Zum Niveau $\alpha = 6.5\%$ ist der Maximin-Test über *G-Reduct.* & *Rnd-Ranks* mit den entsprechenden Raten 6.4% und 42.2% gegeben.

Auf Grundlage der Auswertungen wird die Methode *Π-Jitter* empfohlen. Die Anwendung von *Rnd-Ranks* und $\text{mix}_{\hat{\tau}_n}$ -*Jitter* hat eine ähnliche Güte und ist somit für die Praxisanwendung und ggf. unter Einschränkung auf Modellierungen außerhalb der *Worst Case* Datensätze plausibel.

<p><i>Übersicht der Methoden zur Mitigation von Bindungen bei der Modellanpassung über MPLE-Methoden und den zugehörigen rangbasierten Tests der Anpassungsgüte. Je Methode wird aus den Modifikationen der (Pseudo-)Stichproben, den MPL-Schätzern und den Testverfahren gewählt. Haken (✓) geben die Auswahl an.</i></p>	Modifikationen				MPLE	Tests			
	Reduktion bzgl. voll. Bindungen	Reduktion mit Greedy-Heuristik	Zufällige Rangvergabe	Jittering mit i.i.d. Rauschen	Jittering, empirisch angepasst	Unmodifizierter Schätzer	Intervall-zensierter Schätzer	Test nach Genest et al. (2009)	Test nach Kojadinovic (2017)
Referenzmethode (<i>Reference</i>)	- -	- - -	- - -	- - -	✓	-	✓	✓	-
Reduktion bzgl. vollst. Bindungen (<i>V-Reduct.</i>)	✓ -	- -	- - -	- - -	✓ ✓	-	✓	✓	-
Reduktion mit Greedy-Heuristik (<i>G-Reduct.</i>)	- ✓	-	- - -	- - -	✓ ✓	-	✓	✓	-
Zufällige Rangvergabe (<i>Rnd-Ranks</i>)	- -	✓	- - -	- - -	✓ ✓	-	✓	✓	-
Jittering mit i.i.d. Rauschen (<i>II-Jitter</i>)	- -	-	- ✓	- -	✓ ✓	-	✓	✓	-
Jittering, empirisch angepasst (<i>mix_{τ̂n}-Jitter</i>)	- -	-	- -	- ✓	✓ ✓	-	✓	✓	-
Intervall-Zensur nach Li et al. (2020) (<i>Censoring</i>)	- -	-	- -	- -	- ✓	✓	✓	✓	-
Bootstrap nach Kojadinovic (2017)* (<i>Kojadinovic</i>)	- -	-	- -	- -	✓	-	-	-	✓
<i>V-Reduct.</i> und <i>Rnd-Rank</i> †	✓ -	✓	- -	- -	✓	-	✓	✓	-
<i>G-Reduct.</i> und <i>Rnd-Rank</i> †	- ✓	✓	- -	- -	✓	-	✓	✓	-
<i>V-Reduct.</i> und <i>Censoring</i> †	✓ -	-	- -	- -	- ✓	✓	✓	✓	-
<i>G-Reduct.</i> und <i>Censoring</i> †	- ✓	-	- -	- -	- ✓	✓	✓	✓	-

Tabelle 6.2: Übersicht der Methoden zur Mitigation von Bindungen bei den rangbasierten Tests auf Radial-Symmetrie und Austauschbarkeit zur Anwendung in der Simulationsstudie. Haken (✓) bedeuten den Einbezug der Modifikation. *: Die MPLE nach Kojadinovic (2017) nutzt Mid-Ränge statt Max-Ränge. †: Zuerst wird die Stichprobenreduktion und dann eine Intervall-Zensierung oder Rangrandomisierung durchgeführt. Quelle: Eigene Erstellung.

6.3 Simulation zum Methodenvergleich bzgl. Modellanpassung und Anpassungsgüte

Bei der Simulationsstudie über die Modellanpassung und den Tests der Anpassungsgüte werden die Methoden und Methodenkombinationen aus Tabelle 6.2 berücksichtigt. Es werden die Hypothesen

$$\mathcal{H}_0 : C \in \mathcal{C}_0 \quad \text{und} \quad \mathcal{H}_1 : C \notin \mathcal{C}_0,$$

bzgl. der Copulafamilie $\mathcal{C}_0 = \{C_\vartheta : \vartheta \in \Theta\}$ betrachtet. Für \mathcal{C}_0 werden die Gauß'sche, t-, Clayton- und Gumbel-Familien eingesetzt.

Aufgrund des Rechenaufwands wird – anders als in Abschnitt 6.2 – bei der Modellanpassung und dem Test der Anpassungsgüte lediglich die Anwendung auf modifizierte Datensätze durchgeführt. Die Simulationsresultate der Modellanpassungen sind in den Abbildungen 6.9 bis 6.12 sowie in den Tabellen B.1 bis B.12 im Anhang aufbereitet. Dort sind der Bias und RMSE dargestellt und die Ergebnisse sind bzgl. der jeweiligen Copula und dem Kendall'schen Tau sowie der Anpassungen im datengenerierenden Algorithmus gruppiert. Die Ergebnisse zu den Tests der Anpassungsgüte sind in den Tabellen C.1 bis C.4 aufbereitet. Eine Visualisierung wird ausgelassen, da dies zu komplex in der Darstellung wäre.

Im Voraus ist anzumerken, dass die MPLE-Algorithmen in einzelnen Iterationen nicht konvergiert sind, wodurch die nutzbaren Stichproben der Simulationsergebnisse geringer als die angegebenen $N = 500$ Iterationen sein können. Des Weiteren ist die Modellanpassung unter gewissen Rahmenbedingungen fehlgeschlagen. Sofern notwendig, ist dies in den Tabellen im Anhang gekennzeichnet. Bei der Modellanpassung wurde der Parameter $\vartheta_0 \in \Theta$ geschätzt. Die Schätzung des Kendall'schen Taus nutzt die Formeln zur Berechnung des Taus bei gegebenem Modellparameter ϑ .

Simulationsergebnisse bzgl. der Modellanpassung für die Anwendung bei Daten mit unbekannten Modifikationen

Die Simulationsstudien zur Parameterschätzung unter Berücksichtigung aller Datenmodifikationen zeigen folgende Resultate:

Bei einer MPLE-Modellanpassung bzgl. der Gauß'schen Copula oder der t-Copula ist die Methode *G-Reduct. & Censoring* gegenüber den anderen Methoden zu bevorzugen. *G-Reduct. & Censoring* resultiert in den Schätzungen mit gleichzeitig niedrigen Werten für Bias ($|\dots| \leq 0.0414$) und RMSE (≤ 0.0609). Für niedrige Werte des Kendall'schen Taus, d. h. $\tau = 0.25$, generiert die Methode *Censoring* leicht bessere Resultate für Bias ($|\dots| \leq 0.0090$) und RMSE (≤ 0.0603).

Bzgl. der Clayton-Copula sind die Methoden *V-Reduct.*, *G-Reduct.* und *Censoring* sowie deren Kombinationen geeignete Kandidaten. Genauer gilt: Für das niedrige Kendall'sche Tau $\tau = 0.25$ hat *V-Reduct.* die gleichzeitig niedrigsten Werte für Bias (+0.0008) und RMSE (0.0721). Beim mittleren Kendall'schen Tau $\tau = 0.50$ generiert *V-Reduct. & Censoring* die besten Ergebnisse (Bias: +0.0200, RMSE: 0.0465), und für $\tau = 0.75$ ist die Kombination *G-Reduct. & Censoring* zielführend (Bias: -0.0239, RMSE: 0.0700).

Bei der Modellanpassung für die Gumbel-Copula ist für $\tau = 0.25$ keine Methode gleichzeitig in Bias und RMSE minimal. Die Methoden *V-Reduct.*

(Bias: -0.0225 , RMSE: 0.0650) und *Kojadinovic* (*Reference* mit Mid-Rängen; Bias: -0.0037 , RMSE: 0.0768) minimieren je einen dieser Werte. Beim Kendall'schen Tau $\tau = 0.50$ generiert *V-Reduct.* & *Censoring* die besten Ergebnisse für den Bias ($+0.0293$) und RMSE (0.0459), und für $\tau = 0.75$ ist *G-Reduct.* & *Censoring* (Bias: -0.0068 , RMSE: 0.0377) optimal.

Insgesamt ist für die MPLE die Methode *Censoring* zu empfehlen, wobei für steigende Werte des Kendall'schen Taus die Stichprobenreduktionen *V-Reduct.* und *G-Reduct.* vorher angewendet werden sollen.

Simulationsergebnisse bzgl. der Modellanpassung für die Anwendung bei Daten mit modifizierter Messpräzision

Bei Stichproben bzgl. der Gauß'schen Copula mit einer angepassten Präzision $r \in \{10^{-6}, 10^{-4}\}$ haben alle Methoden einen identischen Bias ($-0.0002, -0.0006$) und RMSE ($0.0157, 0.0156$). Es ist auf Grundlage der Abbildungen anzumerken, dass sich die Methoden im IQR unterscheiden; die Methoden *G-Reduct.*, *mix _{$\hat{\tau}_n$} -Jitter*, *V-Reduct.* & *Rnd-Ranks* und *G-Reduct.* & *Censoring* haben nach Abbildung 6.9 den geringsten IQR. Für die Präzision $r = 10^{-2}$ und die Taus $\tau \in \{0.25, 0.75\}$ minimiert die Methode *Censoring* gleichzeitig den Bias ($|\dots| \leq 0.0004$) und RMSE (≤ 0.0160). Alternativ kommen *Reference* bzw. *Kojadinovic* mit einem höheren Bias -0.0007 und einem leicht niedrigerem RMSE 0.0159 in Betracht. Die Methode *V-Reduct.* & *Censoring* hat gleichzeitig minimalen Bias (-0.0007) und RMSE (0.0286) für $r = 10^{-2}$ und $\tau = 0.50$.

Bei der t-Copula für $\tau = 0.25$ und sofern die Ergebnisse existieren, wird der minimale Bias und niedrige bzw. minimale RMSE durch die Methoden *Censoring*, *V-Reduct.* & *Censoring* und *G-Reduct.* & *Censoring* realisiert. Für $\tau \in \{0.50, 0.75\}$ haben *Rnd-Ranks*, *Π -Jitter* und *mix _{$\hat{\tau}_n$} -Jitter* sowie *Reference* und *Kojadinovic* minimalen Bias ($|\dots| \leq 0.0038, 0.0011$) und RMSE ($\leq 0.0341, 0.0187$). Die *Censoring*-Methoden konvergieren für $r = 10^{-2}$ nicht; es ist alternativ die Methode *V-Reduct.* & *Rnd-Ranks* (Bias: -0.0017 , RMSE: 0.0415) zu empfehlen.

Bei der Clayton-Copula mit $\tau = 0.25$ konvergiert *Censoring* nur für $r = 10^{-2}$ und besitzt den viertniedrigsten Bias ($+0.0018$) und minimalen RMSE (0.0374). Sofern *Censoring* nicht konvergiert, sind die Randomisierungsmethoden, *Reference* und *Kojadinovic* zu wählen (Bias: $|\dots| \leq 0.0072$, RMSE: ≤ 0.0204). Die Resultate der Gumbel-Copula, $\tau = 0.25$, stimmen mit den Resultaten der Clayton-Copula für dasselbe Tau überein. Für $\tau = 0.50$ besitzt *V-Reduct.* gleichzeitig minimalen Bias ($|\dots| \leq 0.0052$) und RMSE (≤ 0.0323). Für $\tau = 0.75$ sind bzgl. $r \in \{10^{-6}, 10^{-4}\}$ alle Methoden gleich-

wertig, wobei *Censoring* für dieses Kendall'sche Tau zu empfehlen ist, da es für $r = 10^{-2}$ minimalen Bias (0.0001) und RMSE (0.0180) besitzt.

Insgesamt ist die Empfehlung der Methode *Censoring* und kombinierte Methoden mit *Censoring* zu bestätigen, sofern konvergent. Als Alternative sind die Randomisierungsmethoden zu nennen.

Simulationsergebnisse bzgl. der Modellanpassung für die Anwendung bei Daten mit zufälliger Nullsetzung

Bei der Modellanpassung unter zufälliger Nullsetzung sind die Resultate für die Gauß'sche Copula und die t-Copula eindeutig: Bei der Gauß'schen Copula ist die Methode *G-Reduct. & Censoring* minimal in Bias und RMSE. Es ist anzumerken, dass das wahre Tau unterschätzt wird und bei steigendem Tau gleichzeitig der Bias und RMSE wachsen: Für $\tau = 0.25, 0.50, 0.75$ ist ($|\text{Bias}|, \text{RMSE}$) jeweils kleiner als (0.0173, 0.0500), (0.0496, 0.0687) und (0.1082, 0.1274). Für niedrigere Wahrscheinlichkeiten $p \in \{2.5\%, 5.0\%\}$ des Nullsetzens ist ebenso eine Anwendung von *G-Reduct.* sinnvoll. Bei der t-Copula sind die Resultate der Methode *G-Reduct. & Rnd-Ranks* optimal. Der absolute Bias ist kleiner als 0.0348 und der RMSE verbleibt unterhalb 0.0575; das Tau wird in diesem Fall unterschätzt. Die Methode *G-Reduct.* generiert für $\tau \in \{0.50, 0.75\}$ ebenso valide Werte; die Methode überschätzt niedrige Werte des Taus stark, was z. B. im maximalen Bias \pm RMSE von 0.0765 ± 0.0979 erkennbar ist. Die Methode *G-Reduct. & Censoring* stellt eine dritte Alternative dar, da in fünf der neun Simulationen minimale Werte für Bias und/oder RMSE generiert wurden. Der absolute Bias für $\tau = 0.25, 0.50, 0.75$ ist dabei kleiner als 0.0118, 0.318 bzw. 0.0477 und der RMSE ist stets kleiner als 0.0654.

Die Auswertung der MPLE zu den Clayton- und Gumbel-Copulas ist durch teilweise fehlende Resultate der *Censoring*-Methoden erschwert. Falls existent, besitzen die Methoden *Censoring*, *V-Reduct. & Censoring* oder *G-Reduct. & Censoring* minimalen Bias und RMSE. Dieser Methoden sind für $\tau = 0.25$ nicht auswertbar; die nächstbesten Resultate liegen bei *Reference*, *V-Reduct.* und *G-Reduct.* sowie *G-Reduct. & Rnd-Ranks* vor.

Insgesamt ist bei der Modellanpassung unter zufälliger Nullsetzung die Methode *G-Reduct. & Censoring*, sofern anwendbar, zu empfehlen. Als Alternative ist *G-Reduct. & Rnd-Ranks* zu nennen.

Simulationsergebnisse bzgl. der Modellanpassung für die Anwendung bei Daten mit empirischer Transformation

Bei den empirisch transformierten Daten wird zwischen den Fällen mit und ohne *Worst Case* Datensätze unterschieden. Liegt ein *Worst Case* Datensatz vor, so sind die Methoden *Censoring*, *V-Reduct.* & *Censoring* und *G-Reduct.* & *Censoring* in allen Szenarien – ausgenommen der Clayton- und Gumbel-Copulas mit $\tau = 0.25$ – optimal. Bei der Gauß'schen Copula, $\tau = 0.25$, haben andere Methoden vereinzelt bessere Güte als die *Censoring*-Methoden; *Censoring* ist als einzige Methode optimal für den Fall *Worst* × *Worst*.

Für die Kombinationen aus *Average Case* und *Best Case* Datensätze haben alle Methoden ähnliche Werte für Bias und RMSE, wobei für $\tau = 0.75$ die Methode *G-Reduct.* & *Censoring* oft optimale Resultate zeigt. Die *Censoring*-Methoden konnten bzgl. der Clayton- und Gumbel-Copula bei *Average Case* und *Best Case* Datensätzen teilweise nicht ausgewertet werden.

Für *Worst Case* Datensätze sind die Methoden *Censoring*, *V-Reduct.* & *Censoring* und *G-Reduct.* & *Censoring* zu empfehlen. Für *Average Case* und *Best Case* kann keine sinnvolle Empfehlung ausgesprochen werden.

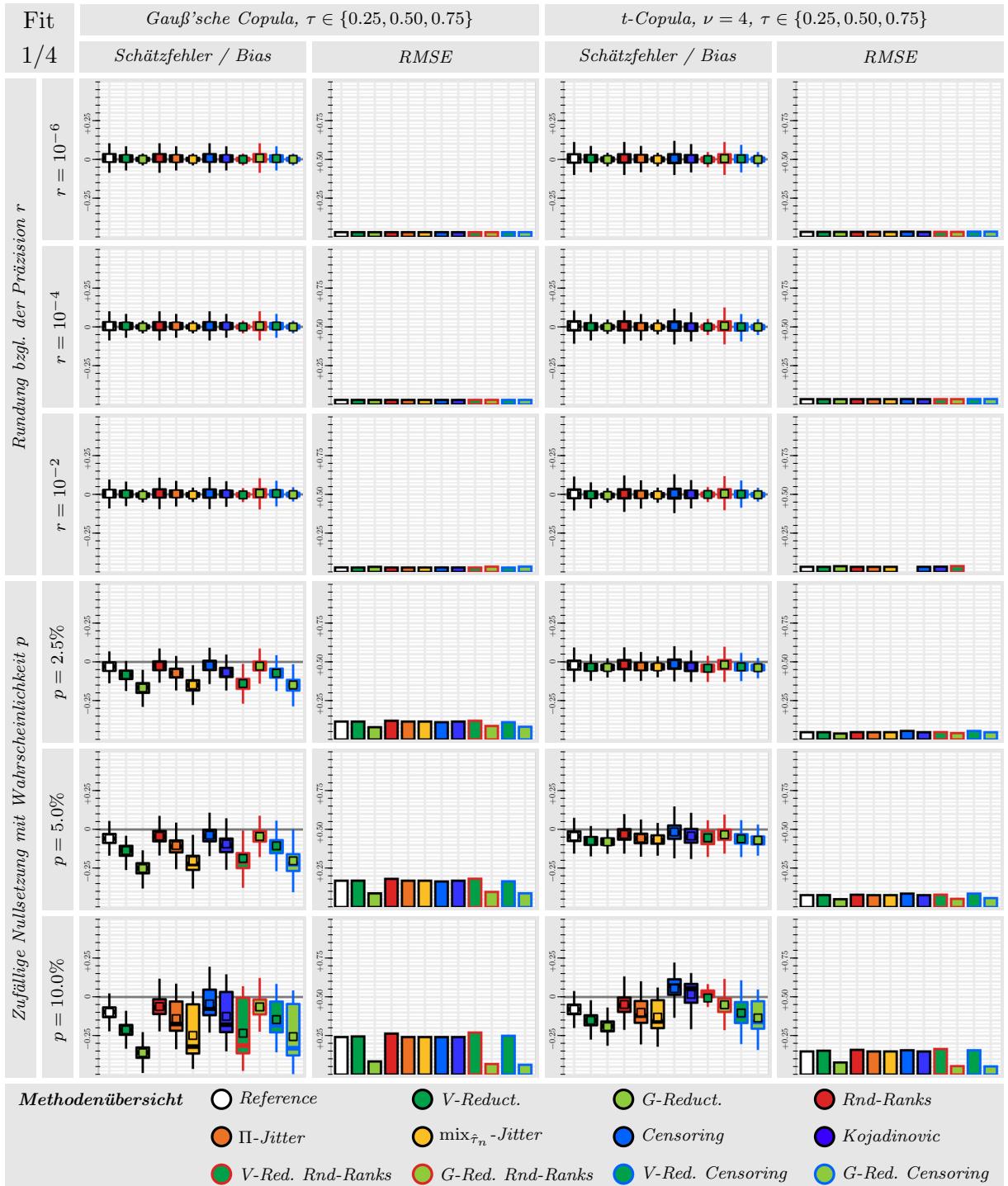
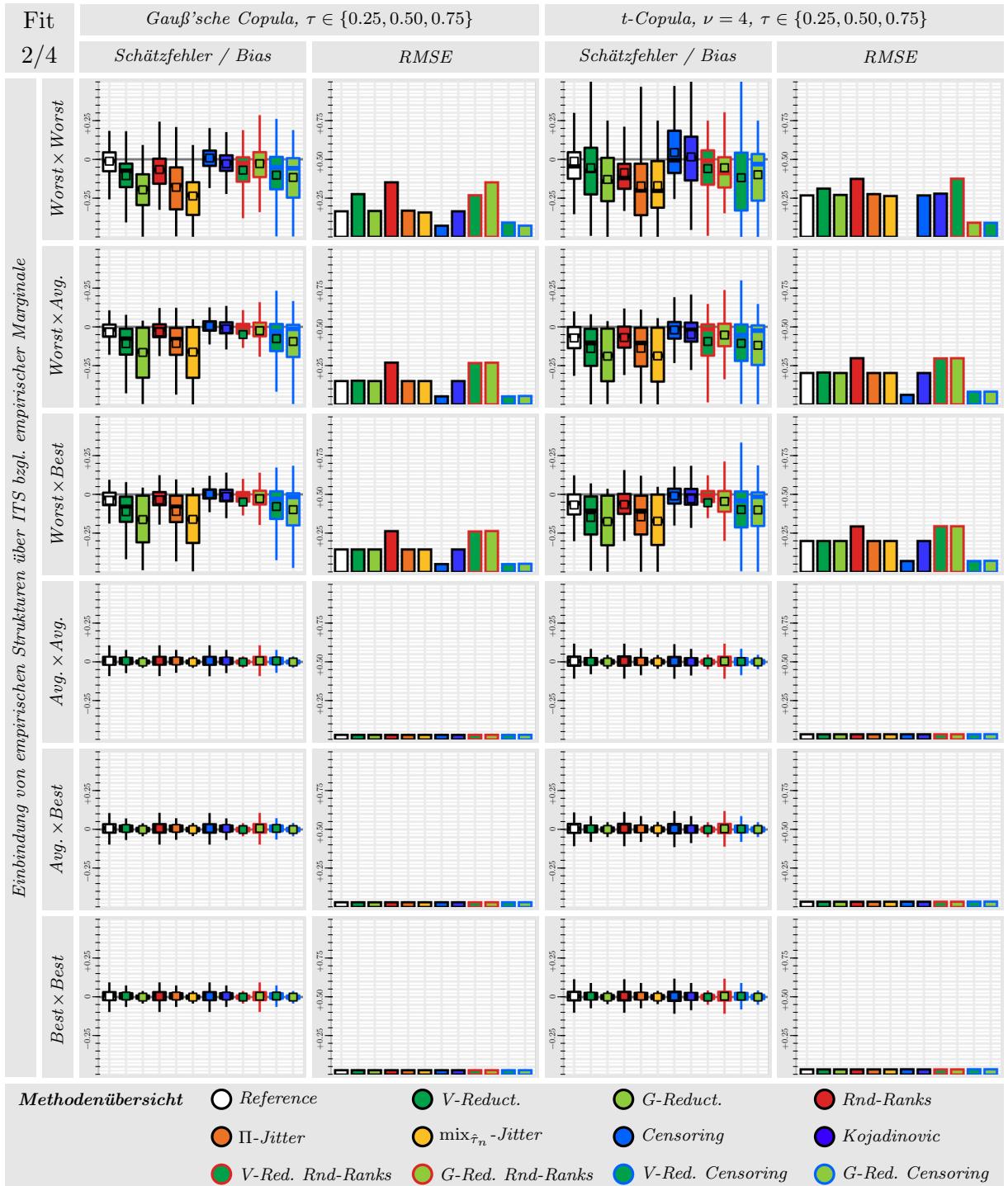


Abbildung 6.9: Bias und RMSE der Modellanpassungen für die Gauß'sche Copula und die t -Copula, Teil 1 von 4: Box-Plots zur Visualisierung der Schätzfehler sowie Balkendiagramme für den RMSE der MPLE-Modellanpassung. Die Simulationsstudie umfasst $N = 500$ Iterationen mit Stichproben der Größe $n = 250$. Die Quadrate in den Boxplots geben den empirische Bias an; Ausreißer sind zur Übersichtlichkeit ausgeblendet. Quelle: Eigene Erstellung.



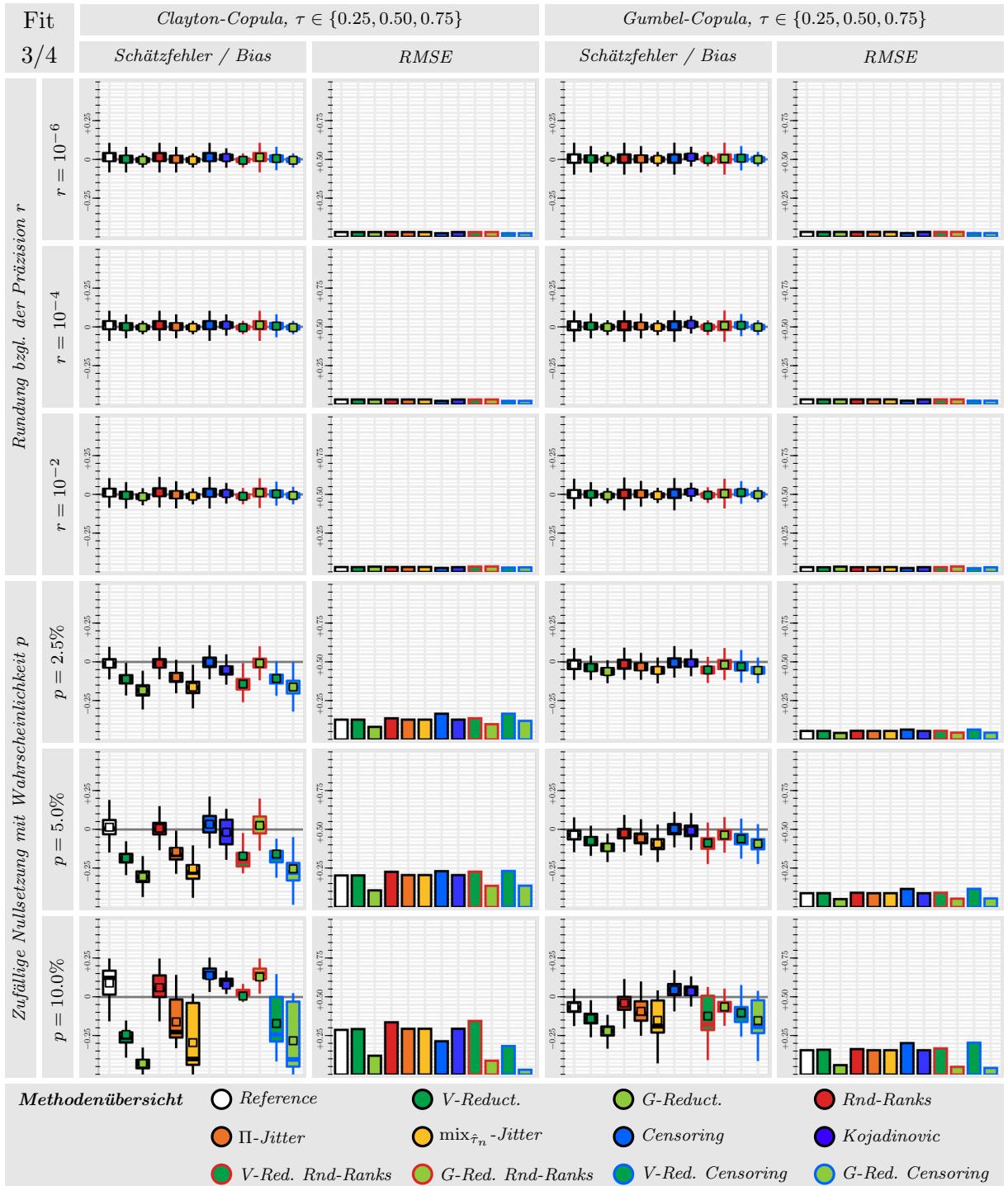


Abbildung 6.11: Bias und RMSE der Modellanpassungen für die Clayton-Copula und die Gumbel-Copula, Teil 3 von 4: Box-Plots zur Visualisierung der Schätzfehler sowie Balkendiagramme für den RMSE der MPLE-Modellanpassung. Die Simulationsstudie umfasst $N = 500$ Iterationen mit Stichproben der Größe $n = 250$. Die Quadrate in den Boxplots geben den empirische Bias an; Ausreißer sind zur Übersichtlichkeit ausgeblendet. Quelle: Eigene Erstellung.

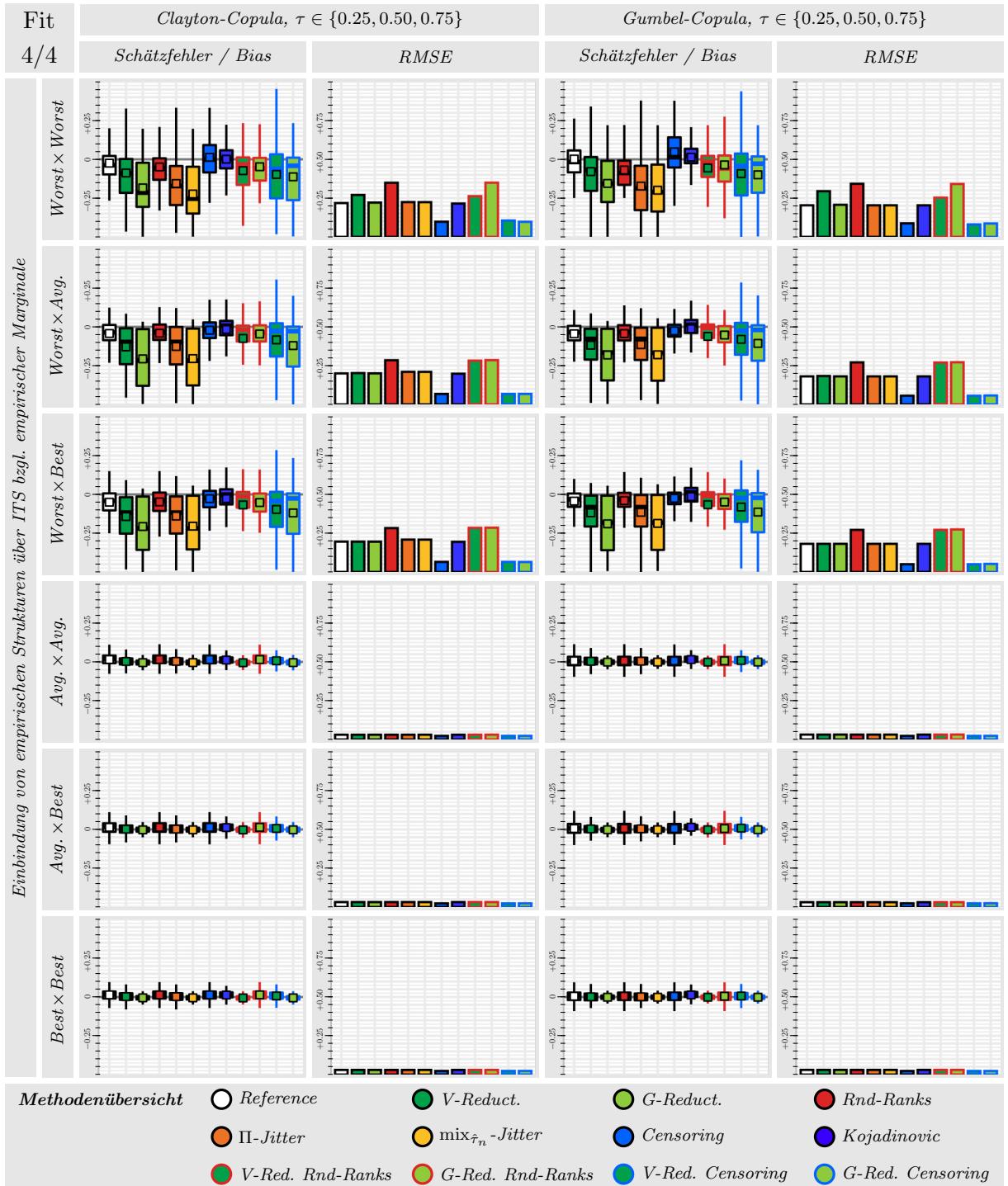


Abbildung 6.12: Bias und RMSE der Modellanpassungen für die Clayton-Copula und die Gumbel-Copula, Teil 4 von 4: Box-Plots zur Visualisierung der Schätzfehler sowie Balkendiagramme für den RMSE der MPLE-Modellanpassung. Die Simulationsstudie umfasst $N = 500$ Iterationen mit Stichproben der Größe $n = 250$. Die Quadrate in den Boxplots geben den empirische Bias an; Ausreißer sind zur Übersichtlichkeit ausgeblendet. Quelle: Eigene Erstellung.

Simulationsergebnisse bzgl. Anpassungsgüte für die Anwendung bei Daten mit unbekannten Modifikationen

Es ist anzumerken, dass die Auswertungen bzgl. der vier Copulafamilien auf unterschiedlich großen Stichproben basieren (vgl. Tabellen C.1 bis C.4).

Bei Berücksichtigung aller Tests der Anpassungsgüte und aller Datenmodifikationen ist kein unverfälschter Niveau- α -Test zum Fünfprozentniveau feststellbar. Unverfälschte Tests bzgl. aller vier Nullhypotesen sind bei der Methode *V-Reduct.* zu finden; diese hält das Niveau 30.3 %.

Werden die Nullhypotesen bzgl. der t-, Clayton- und Gumbel-Copulas betrachtet, so sind weitere unverfälschte Tests identifizierbar. Alle Methoden außer *Censoring*, *Kojadinovic*, *V-Reduct.* & *Censoring* und *G-Reduct.* & *Censoring* sind unverfälscht. Die drei Methoden *Reference*, *V-Reduct.* und *G-Reduct.* sind unverfälschte Niveau- α -Tests zu den Niveaus 25.5 %, 26.2 % und 19.7 %. Bei *Rnd-Ranks*, *V-Reduct.* & *Rnd-Ranks* und *G-Reduct.* & *Rnd-Ranks* liegen unverfälschte Niveau- α -Tests zu den Niveaus 21.1 %, 21.0 % und 21.3 % vor. *II-Jitter* und *mix_{τ̂ₙ}-Jitter* sind unverfälscht zu den Niveaus 21.3 % und 22.9 %. Die Maximin-Tests unter diesen Rahmenbedingungen sind gegeben mit *V-Reduct.* und *G-Reduct.*, welche minimale \mathcal{H}_0 -Ablehnungsraten von 29.0 % unter \mathcal{H}_1 besitzen. Der nachfolgend optimale Test ist *Reference* mit einem Wert von 28.7 %.

Die Methoden *V-Reduct.* und ebenso *G-Reduct.* und *Reference* erzielen übergreifend die besten Ergebnisse, weshalb diese für die Anwendung als Ansatzpunkte zu empfehlen sind.

Simulationsergebnisse bzgl. Anpassungsgüte für die Anwendung bei Daten mit modifizierter Messpräzision

Bei einer angepassten Messpräzision sind alle Methoden ausgenommen der *Jittering*-Methoden unverfälschte Niveau- α -Tests, wobei die maximalen Ablehnungsraten unter \mathcal{H}_0 zwischen 5.3 % und 7.3 % liegen. Der Wert 5.3 % gehört zur Methode *Kojadinovic*, die eine minimale Ablehnungsrate 10.2 % unter \mathcal{H}_1 besitzt. Zum Niveau $\alpha = 7.3$ existieren zwei Tests mit höheren minimalen Ablehnungsraten im \mathcal{H}_1 -Fall: Die Raten für *Rnd-Ranks* und *V-Reduct.* & *Rnd-Ranks* betragen 10.7 % und 11.1 %. Diese Methoden sind somit Maximin-Tests. Bei der *Kojadinovic*-Methode ist die Differenz zwischen der maximalen Rate unter \mathcal{H}_0 und minimalen Rate unter \mathcal{H}_1 mit 4.2 % größer als die Differenzen 3.4 % und 3.8 % der Maximin-Tests.

Insgesamt ist unter diesen Rahmenbedingungen die Nutzung der *Kojadinovic*-Methode zu empfehlen.

Simulationsergebnisse bzgl. Anpassungsgüte für die Anwendung bei Daten mit zufälliger Nullsetzung

Die Simulationsergebnisse bzgl. der zufälligen Nullsetzung ergeben keine unverfälschten Tests. Entsprechend wird keine Empfehlung ausgesprochen.

Simulationsergebnisse bzgl. Anpassungsgüte für die Anwendung bei Daten mit empirischer Transformation

Bei den empirisch transformierten Daten ist zwischen den Fällen mit und ohne *Worst Case* Daten zu unterscheiden. Bzgl. *Worst Case* Daten konnten keine unverfälschten Tests identifiziert werden, da in mehreren Fällen sehr hohe Ablehnungsraten unter \mathcal{H}_0 oder vereinzelt sehr niedrige Ablehnungsraten unter \mathcal{H}_1 vorliegen. Bei der ausschließlichen Anwendung für die Gauß'sche und Gumbel-Copulas ist die *Kojadinovic*-Methode ein unverfälschter Niveau- α -Test für $\alpha = 7.4\%$. Da die betrachtete Menge an Copulas halbiert wurde, ist darauf aufbauend keine allgemeine Empfehlung möglich.

Für die Transformationen bzgl. *Average Case* und *Best Case* Datensätzen sind mit Ausnahme von *Censoring*, *V-Reduct. & Censoring* und *G-Reduct. & Censoring* alle Methoden unverfälschte Tests mit Niveaus zwischen 5.3 % und 7.6 %. Die Tests mit den niedrigsten Niveaus lauten *V-Reduct. & Rnd-Ranks* und *G-Reduct. & Rnd-Ranks*. Unter \mathcal{H}_0 sind deren maximalen Ablehnungsraten 5.3 % und 5.8 %. Unter \mathcal{H}_1 sind deren minimalen Ablehnungsraten 8.8 % und 10.3 %. Die Maximin-Tests sind gegeben mit *G-Reduct.* gefolgt von *G-Reduct. & Rnd-Ranks*. *G-Reduct.* hat die entsprechenden Ablehnungsraten 6.2 % und 10.5 %.

Insgesamt werden für diese Rahmenbedingungen die Methoden *G-Reduct.* und *G-Reduct. & Rnd-Ranks* für die Anwendung empfohlen.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Beantwortung der Forschungsfragen

Die in der Einleitung vorgestellten Kernfragen werden zusammenfassend beantwortet. Bei der ersten Forschungsfrage

- ① Welche Eigenschaften, Strukturen, Rahmenbedingungen, Datenfehler etc. in empirischen Finanzmarktdaten führen zu Verletzungen der Annahmen? Welcher Anteil der Daten ist davon betroffen?

ist zu beantworten, dass insbesondere zwei Faktoren verantwortlich für das Auftreten von Bindungen in Finanzmarktdaten sind. Regulatorische und börsenspezifische Anforderungen an die Präzision der Finanzmarktdaten führen durch die Rundung der Werte zu Bindungen. Aufgrund der Komplexität der Finanzmärkte ist eine gewisse Fehleranfälligkeit in der Datenverarbeitung gegeben, sodass Datensätze unvollständig bzw. mit Auslassungen bereitgestellt werden. Diese Rahmenbedingungen liegen allen Finanzmarktdaten zugrunde, wobei sich die genauen Anforderungen und Effekte stetig verändern.

Für die Antwort auf die Forschungsfrage

- ② Welche Auswirkungen (bzgl. Bias, Anwendbarkeit, Robustheit oder ähnlichen Kriterien) haben Annahmeverstöße auf die Methoden?

ist heranzuziehen, dass die unmodifizierten Methoden *Reference* nur unter schwachen Datenmodifikationen plausible Ergebnisse generieren und dort die modifizierten Methoden bereits eine bessere Güte innehaben. Für schwere Eingriffe in die Datensätze sind die unmodifizierten Methoden in der Regel nicht mehr nutzbar. Es liegen trotz Datenmodifikation verschiedene Methoden vor, die unverfälschte Niveau- α -Tests darstellen oder die besten Methoden unter gewissen Einschränkungen oder Annahmen sind.

7.2 Anwendungsempfehlungen

Damit die neun bis zwölf Methoden je Sachverhalt in der Praxis effektiv eingesetzt werden können, werden die beiden Forschungsfragen

- ③ Wie groß ist der Nutzen einer Modifikation im Vergleich zu den anderen Methoden für Datensätze mit Bindungen?
- ④ Falls mehrere Methoden für einen Sachverhalt vorliegen: Welche der Methoden ist zur Anwendung zu empfehlen? Welche Methode ist nur unter Einschränkungen oder nur mit der Bereitstellung von zusätzlichen Informationen nutzbar?

beantwortet. Die Abbildungen 7.1 bis 7.4 stellen Entscheidungsdiagramme zur Nutzung der einzelnen Methoden dar. Es werden vier Fälle unterschieden: *Allgemeine Anwendung* umfasst die Nutzung ohne Einbezug weiterer Erkenntnisse an die Finanzmarktdaten. *Angepasste Präzision* bezieht sich auf Daten mit Einschränkungen an die Messpräzision. *Zufälliges Nullsetzen* wird zur Handhabe von fehlerhaften Datenübertragungen betrachtet. *Empirische Transformation* bezieht sich auf synthetisch generierte Daten, die über empirische Datensätze generiert wurden.

In den Entscheidungsdiagrammen sind die Empfehlungen zur Methodennutzung je Rahmenbedingung aufgelistet. Die Empfehlung basiert auf der Güte der Resultate aus den Simulationsstudien. Die Angaben in Klammern oder mit Schrägstrich geben alternative Methoden und zusätzliche Informationen an. Falls keine Empfehlung ausgesprochen werden kann, ist dies ebenso gekennzeichnet und die zusätzlich genannte Methode ist die Methode mit der besten Güte.

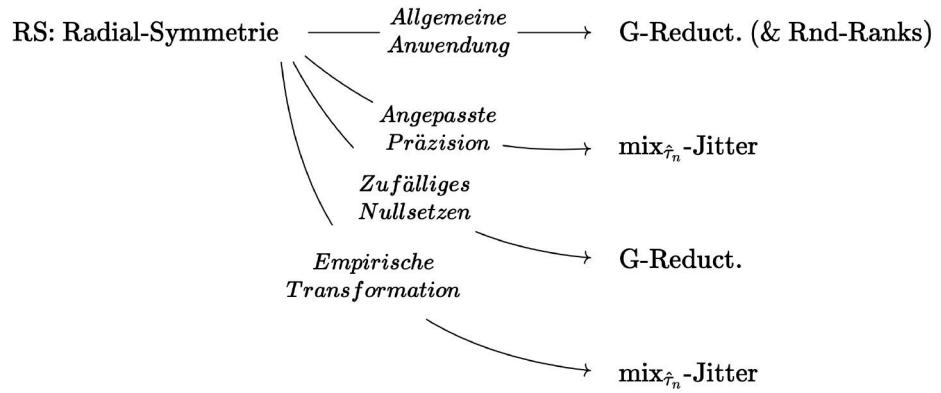


Abbildung 7.1: Schema zur Methodenauswahl für die Tests auf Radial-Symmetrie unter bekannten und unbekannten Rahmenbedingungen. Quelle: Eigene Erstellung.

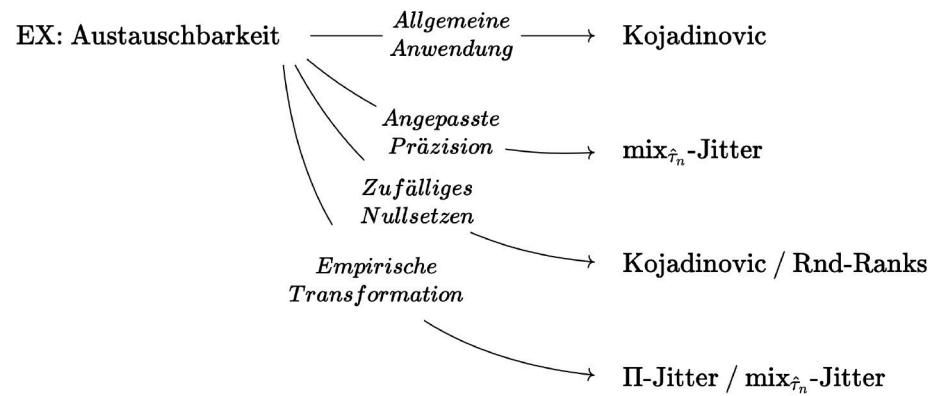


Abbildung 7.2: Schema zur Methodenauswahl für die Tests auf Austauschbarkeit unter bekannten und unbekannten Rahmenbedingungen. Quelle: Eigene Erstellung.

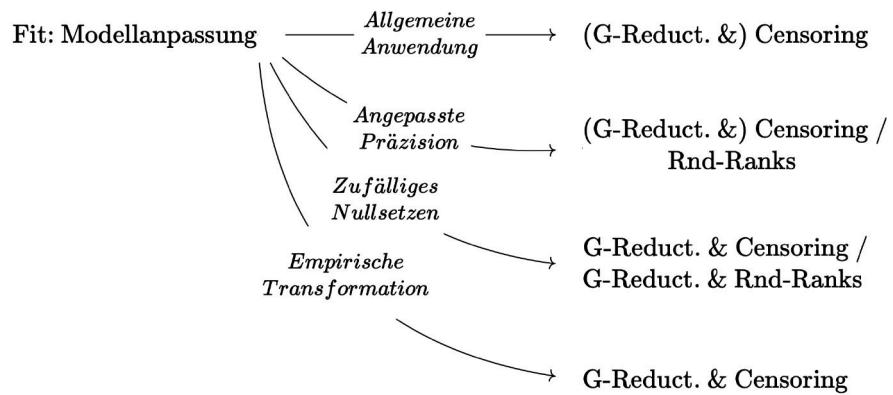


Abbildung 7.3: Schema zur Methodenauswahl für die Modellanpassung unter bekannten und unbekannten Rahmenbedingungen. Quelle: Eigene Erstellung.

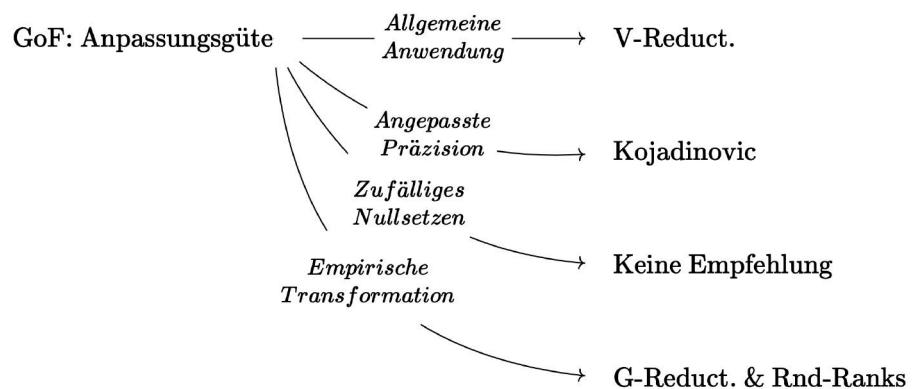


Abbildung 7.4: Schema zur Methodenauswahl für die Tests der Anpassungsgüte unter bekannten und unbekannten Rahmenbedingungen. Quelle: Eigene Erstellung.

7.3 Kritik und Ausblick

Diese Arbeit stellt einen grundlegenden Vergleich der Methoden bzgl. der Resultate unter gewissen Anpassungen der Daten auf. Um die Wirkung von Bindungen und Finanzmarktrahmenbedingungen auf Copula-Methoden besser zu verstehen, sind weitere Analysen über die Strukturen innerhalb der Datensätze notwendig. Auch wären bzgl. der Rahmenbedingungen Schätzungen von Schwellenwerten zur Entscheidung, ob gewisse Methoden angewendet werden können oder nicht, sinnvoll.

Die Laufzeit der Simulationen und Methoden sind robuster zu gestalten und ggf. sind andere Optimierungsalgorithmen in den Modellanpassungen und Tests der Anpassungsgüte zu berücksichtigen. Um die Anwendbarkeit der Methoden auf eine Vielzahl von Szenarien tiefergehend zu validieren, hätten größere Simulationsstudien mit Stichproben unterschiedlichen Umfangs in Betracht gezogen werden müssen. Da einige Simulationen für den Test der Anpassungsgüte 72 Stunden pro 100 Iterationen benötigt haben, ist die zusätzliche Laufzeit ein wichtiger Faktor der berücksichtigt werden muss.

Durch die Untersuchung weiterer rangbasierter Copula-Methoden für die Anwendungsbereiche der Zeitreihenanalyse und Extremwerttheorie ließen sich die hier gewonnenen Ergebnisse ergänzen.

Begründet durch die positiven Resultate des Greedy-Algorithmus zur Stichprobenreduktion ergibt sich weiterer Forschungsbedarf über die Nutzung weiterer Arten von Greedy- bzw. Optimierungsalgorithmen zum effizienten Verarbeiten von Bindungen. Auch von Interesse wären Machine-Learning-Methoden, obwohl hier das Kosten-Nutzen-Verhältnis möglicherweise ineffizient ist und größere Datenmengen in besserer Qualität benötigt würden.

Literatur

- Agresti (März 2010). *Ordinal Categorical Data 2e*. John Wiley Sons. 418 S. ISBN: 0470082895. URL: https://www.ebook.de/de/product/9648681/agresti_ordinal_categorical_data_2e.html.
- Buske, Nils (18. März 2021). „Börsenfeiertage 2021 - An diesen Tagen stehen die Börsen still“. In: *Handelsblatt*. URL: <https://www.handelsblatt.com/finanzen/anlagestrategie/trends/boersenfeiertage-2021-an-diesen-tagen-stehen-die-boersen-still/27015118.html> (Abfrage: 22.05.2021).
- Bücher, Axel (2019). „Mathematische Statistik I - Skriptum zur Vorlesung an der Heinrich-Heine-Universität Düsseldorf“. In:
- Bücher, Axel und Ivan Kojadinovic (2016). „A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing“. In: *Bernoulli* 22.2. DOI: 10.3150/14-bej682.
- Czado, Claudia (2010). „Pair-copula constructions of multivariate copulas“. In: *Copula theory and its applications*. Springer, S. 93–109.
- (Mai 2019). *Analyzing Dependent Data with Vine Copulas*. Springer-Verlag GmbH. 242 S. URL: https://www.ebook.de/de/product/37749557/claudia_czado_analyzing_dependent_data_with_vine_copulas.html.
- Davis, Trevor L (2020). *optparse: Command Line Option Parser*. R package version 1.6.6. URL: <https://CRAN.R-project.org/package=optparse>.
- Dhevi, A. T. Sree (2014). „Imputing missing values using Inverse Distance Weighted Interpolation for time series data“. In: *2014 Sixth International Conference on Advanced Computing (ICoAC)*. IEEE. DOI: 10.1109/icoac.2014.7229721.
- Embrechts, Paul und Marius Hofert (2013). „A note on generalized inverses“. In: *Mathematical Methods of Operations Research* 77.3, S. 423–432. DOI: 10.1007/s00186-013-0436-7.
- EU, Europäische Union (Juni 2014). „Richtlinie 2014/65/EU des Europäischen Parlaments und des Rates vom 15. Mai 2014 über Märkte für Finanzinstrumente sowie zur Änderung der Richtlinien 2002/92/EG und 2011/61/EU“.

- In: *Amtsblatt der Europäischen Union L* 173 57, S. 349 –496. ISSN: 1977-0677. URL: <http://data.europa.eu/eli/dir/2014/65/oj>.
- EU, Europäische Union (März 2017). „Delegierte Verordnung (EU) 2017/588 DER KOMMISSION vom 14. Juli 2016 zur Ergänzung der Richtlinie 2014/65/EU des Europäischen Parlaments und des Rates durch technische Regulierungsstandards für das Tick-Größen-System für Aktien, Aktienzertifikate und börsengehandelte Fonds“. In: *Amtsblatt der Europäischen Union L* 87 60, S. 411 –416. ISSN: 1977-0642. URL: http://data.europa.eu/eli/reg_del/2017/588/oj.
- Fusion Media Ltd. (2021a). *Investing.com*. Quellenreferenz für den Bezug verschiedener Datensätze. URL: <https://www.investing.com>.
- Fusion Media Ltd., Investing.com (2021b). *Investing.com: Risk Disclosure*. Die Risk Disclosure ist im Footer der Internetseite zu finden. URL: <https://www.investing.com> (Abfrage: 22.05.2021).
- (2021c). *Investing.com: Über uns*. URL: <https://de.investing.com/about-us/> (Abfrage: 24.06.2021).
- Galassi, Mark et al. (2021). *GNU scientific library*. 2.7. Bristol: Network Theory. ISBN: 0954161734.
- Genest, Christian und Johanna G. Nešlehová (2013). „On tests of radial symmetry for bivariate copulas“. In: *Statistical Papers* 55.4, S. 1107–1119. DOI: [10.1007/s00362-013-0556-4](https://doi.org/10.1007/s00362-013-0556-4).
- Genest, Christian und Bruno Rémillard (2008). „Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models“. In: *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 44.6. DOI: [10.1214/07-aihp148](https://doi.org/10.1214/07-aihp148).
- Genest, Christian, Bruno Rémillard und David Beaudoin (2009). „Goodness-of-fit tests for copulas: A review and a power study“. In: *Insurance: Mathematics and Economics* 44.2, S. 199–213. DOI: [10.1016/j.insmatheco.2007.10.005](https://doi.org/10.1016/j.insmatheco.2007.10.005).
- Genest, Christian, Johanna Nešlehová und Martin Ruppert (2011). „Discussion: Statistical models and methods for dependence in insurance data“. In: *Journal of the Korean Statistical Society* 40.2, S. 141–148. DOI: [10.1016/j.jkss.2011.03.004](https://doi.org/10.1016/j.jkss.2011.03.004).
- Genest, Christian, Johanna Nešlehová und Jean-François Quessy (2012). „Tests of symmetry for bivariate copulas“. In: *Annals of the Institute of Statistical Mathematics* 64.4, S. 811–834. DOI: [10.1007/s10463-011-0337-6](https://doi.org/10.1007/s10463-011-0337-6).
- Gritzmann, Peter (2013). *Grundlagen der Mathematischen Optimierung*. Springer Fachmedien Wiesbaden. DOI: [10.1007/978-3-8348-2011-2](https://doi.org/10.1007/978-3-8348-2011-2).
- Hafner, Robert (2001). *Nichtparametrische Verfahren der Statistik*. Springer Vienna. DOI: [10.1007/978-3-7091-6273-6](https://doi.org/10.1007/978-3-7091-6273-6).

- Harder, Michael und Ulrich Stadtmüller (2016). „Testing exchangeability of copulas in arbitrary dimension“. In: *Journal of Nonparametric Statistics* 29.1, S. 40–60. DOI: 10.1080/10485252.2016.1253841.
- Heinrich-Heine-Universität (2021). *High-Performance Computing Rechencluster HILBERT*. URL: <https://www.zim.hhu.de/forschung/high-performance-computing> (Abfrage: 01.08.2021).
- Heldt, Cordula (19. Feb. 2018). „Over-the-Counter Market - Ausführliche Definition im Online-Lexikon“. In: URL: <https://wirtschaftslexikon.gabler.de/definition/over-counter-market-45932/version-269218> (Abfrage: 25.05.2021).
- Hlawatsch, Stefan und Peter Reichling (Juli 2010). *Konstruktion und Anwendung von Copulas in der Finanzwirtschaft*. FEMM Working Papers 100016. Otto-von-Guericke University Magdeburg, Faculty of Economics und Management. URL: <https://ideas.repec.org/p/mag/wpaper/100016.html>.
- Hofert, Marius et al. (2018). *Elements of Copula Modeling with R*. Springer International Publishing. ISBN: 3319896342. DOI: 10.1007/978-3-319-89635-9. URL: https://www.ebook.de/de/product/32071108/marius_hofert_ivan_kojadinovic_martin_maechler_jun_yan_elements_of_copula_modeling_with_r.html.
- Hofert, Marius et al. (2020). *copula: Multivariate Dependence with Copulas*. R package version 1.0-0. URL: <https://CRAN.R-project.org/package=copula>.
- Hájek, Jaroslav (1999). *Theory of rank tests*. San Diego, Calif: Academic Press. ISBN: 9780126423501.
- Khoudraji, Abdelhaq (Jan. 1995). „Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées /“. Diss.
- Klenke, Achim (Feb. 2008). *Wahrscheinlichkeitstheorie*. Springer-Verlag GmbH. 624 S. ISBN: 9783540775713. URL: https://www.ebook.de/de/product/25604428/achim_klenke_wahrscheinlichkeitstheorie.html.
- Kojadinovic, Ivan (2017). „Some copula inference procedures adapted to the presence of ties“. In: *Computational Statistics Data Analysis* 112, S. 24–41. ISSN: 0167-9473. DOI: <https://doi.org/10.1016/j.csda.2017.02.006>. URL: <https://www.sciencedirect.com/science/article/pii/S0167947317300269>.
- Krämer Nicole, Schepsmeier Ulf (Dez. 2011). *Introduction to vine copulas*. TU München. URL: <https://www.groups.ma.tum.de/fileadmin/w00ccg/statistics/veranstaltungen/Vines.pdf> (Abfrage: 01.07.2021).
- Lee, Elisa (2003). *Statistical methods for survival data analysis*. Hoboken, N.J: Wiley. ISBN: 0471369977.

- Li, Yan et al. (2016). *censorcopula: Estimate Parameter of Bivariate Copula*. R package version 2.0. URL: <https://CRAN.R-project.org/package=censorcopula>.
- Li, Yan et al. (2020). „Copula modeling for data with ties“. In: *Statistics and Its Interface* 13.1, S. 103–117. DOI: 10.4310/sii.2020.v13.n1.a9.
- Lü, Tian-Jian et al. (2020). „Modeling multivariate distribution of multiple soil parameters using vine copula model“. In: *Computers and Geotechnics* 118, S. 103340. DOI: 10.1016/j.compgeo.2019.103340.
- McNeil, Alexander J., Rüdiger Frey und Paul Embrechts (Mai 2015). *Quantitative Risk Management*. Princeton University Press. 720 S. ISBN: 0691166277. URL: https://www.ebook.de/de/product/23323892/alexander_j_mcneil_rudiger_frey_paul_embrechts_quantitative_risk_management.html.
- NASDAQ (2021a). *Continued Listing Guide*. March 2021. URL: <https://listingcenter.nasdaq.com/assets/continuedguide.pdf> (Abfrage: 22.05.2021).
- (2021b). *Initial Listing Guide*. March 2021 - V2. URL: <https://listingcenter.nasdaq.com/assets/initialguide.pdf> (Abfrage: 22.05.2021).
- Nelsen, Roger B. (2006). *An Introduction to Copulas*. Springer-Verlag GmbH. ISBN: 978-0387-28659-4. URL: https://www.ebook.de/de/product/11428705/roger_b_nelsen_an_introduction_to_copulas.html.
- NYSE (o.J.). *New York Stock Exchange Listed Company Manual*. URL: <https://nyse.wolterskluwer.cloud/listed-company-manual> (Abfrage: 22.05.2021).
- Pappadà, R., F. Durante und G. Salvadori (2016). „Quantification of the environmental structural risk with spoiling ties: is randomization worthwhile?“ In: *Stochastic Environmental Research and Risk Assessment* 31.10, S. 2483–2497. DOI: 10.1007/s00477-016-1357-9.
- Patton, Andrew (2013). „Copula Methods for Forecasting Multivariate Time Series“. In: *Handbook of Economic Forecasting*. Elsevier, S. 899–960. DOI: 10.1016/b978-0-444-62731-5.00016-6.
- Quessy, Jean-François (2021). *J.-F. Quessy's Matlab webpage for copulas*. URL: https://oraprdnt.uqtr.quebec.ca/pls/public/gscw031?owa_no_site=834 (Abfrage: 12.08.2021).
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. 3.6.1–4.0.1. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Refinitiv (2019). *Refinitiv Eikon – Your open, trusted source of financial analysis data and tools*. URL: https://www.refinitiv.com/content/dam/marketing/en_us/documents/brochures/eikon-overview-brochure.pdf (Abfrage: 24.06.2021).

- Rémillard, Bruno (2017). „Goodness-of-Fit Tests for Copulas of Multivariate Time Series“. In: *Econometrics* 5.1, S. 13. DOI: 10.3390/econometrics5010013.
- Rüschedorf, Ludger (2009). „On the distributional transform, Sklar's theorem, and the empirical copula process“. In: *Journal of Statistical Planning and Inference* 139.11, S. 3921–3927. DOI: 10.1016/j.jspi.2009.05.030.
- (2014). *Mathematische Statistik*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-41997-3.
- (2016). *Wahrscheinlichkeitstheorie*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-48937-6.
- SEC, Securities and Exchange Commission (Apr. 2020). *Title 17 Commodity and Securities Exchanges 240 - End*, URL: <https://www.govinfo.gov/content/pkg/CFR-2020-title17-vol4/pdf/CFR-2020-title17-vol4.pdf>.
- Segers, Johan (2012). „Asymptotics of empirical copula processes under non-restrictive smoothness assumptions“. In: *Bernoulli* 18.3. DOI: 10.3150/11-bej387.
- Sklar, A. (1959). „Fonctions de répartition à n dimensions et leurs marges“. In: *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231.
- Smillie, Alan (2008). „New copula models in quantitative finance“. Diss. URL: <https://spiral.imperial.ac.uk/bitstream/10044/1/11326/2/Smillie-A-2008-PhD-Thesis.pdf> (Abfrage: 22.08.2021).
- Thomson Reuters (2021). *Refinitiv Eikon*. Quellenreferenz für den Bezug verschiedener Datensätze. URL: <https://eikon.thomsonreuters.com/index.html>.
- Witting, Hermann (1985). *Mathematische Statistik I : Parametrische Verfahren bei festem Stichprobenumfang*. Wiesbaden: Vieweg+Teubner Verlag Imprint Vieweg+Teubner Verlag. ISBN: 9783322901507.
- Yahoo Inc. (2021a). *Yahoo! Finance*. Quellenreferenz für den Bezug verschiedener Datensätze. URL: <http://finance.yahoo.com>.
- Yahoo Inc., Yahoo! Finance (2021b). *Börsen und Datenanbieter auf Yahoo Finanzen*. URL: https://help.yahoo.com/kb/finance-for-web/SLN2310.html?locale=de_DE (Abfrage: 22.05.2021).

Anhang A

Detaillierte Testresultate für die Tests auf Radial-Symmetrie und Austauschbarkeit

Die Tabellen A.1 und A.2 umfassen die Simulationsergebnisse bzgl. der \mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie. Diese dienen zur Grundlage der Abbildungen 6.1 bis 6.4.

Die Tabellen A.3 und A.4 umfassen die Simulationsergebnisse bzgl. der \mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit. Diese Daten liegen den Abbildungen 6.5 bis 6.8 zugrunde.

In den Tabellen sind die \mathcal{H}_0 -Ablehnungsraten für die jeweiligen \mathcal{H}_1 - und \mathcal{H}_0 -Copulas dargestellt. Rote Einträge geben Werte unterhalb 5.0 % an. Blaue Einträge umfassen die Methoden mit den höchsten Ablehnungsraten je Copula, Kendall'schem Tau und Datenmodifikation.

Tests auf Radial-Symmetrie		<i>Gauß'sche Copula (\mathcal{H}_0)</i> \mathcal{H}_0 -Ablehnungsraten in %									<i>t-Copula, $\nu = 4$ (\mathcal{H}_0)</i> \mathcal{H}_0 -Ablehnungsraten in %								
Tau	Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix. φ_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix. φ_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks
0.25	<i>Unmodifiziert</i>	3.8 10.1	3.7 10.6	3.5 3.7	3.7 6.0	3.6 5.9	3.7 7.1	4.7 4.3	3.7 6.4	3.6 3.9	3.9 10.1	3.9 10.3	3.9 3.7	4.0 5.8	3.8 5.9	3.8 7.3	4.4 4.1	3.9 6.4	3.8 4.2
	$r = 10^{-6}$	4.4 3.0 4.8 4.8 5.6 9.4 74.2	4.0 3.6 4.8 4.6 5.8 9.6 74.0	3.9 4.8 3.4 3.4 5.3 5.6 6.2	3.8 3.4 3.4 4.6 5.2 6.4 25.4	4.4 3.4 3.4 4.6 5.0 6.4 24.8	4.2 4.2 3.0 4.4 5.2 6.6 26.0	6.4 6.4 4.8 4.4 6.6 6.4 7.0	4.4 4.4 3.2 4.2 5.4 3.8 31.2	3.9 3.9 3.2 3.2 3.4 2.2 4.0	3.2 4.4 5.4 5.4 6.0 11.2 71.2	3.4 4.2 4.4 2.8 3.4 3.4 20.6	3.2 4.8 2.8 3.4 3.4 2.6 19.8	3.2 4.8 3.4 3.4 2.6 4.4 21.0	3.4 5.6 5.6 5.8 5.6 6.8 5.2	3.2 5.6 5.6 5.8 5.6 6.8 28.2	3.2 5.6 5.6 5.8 5.8 6.8 5.8		
	$W \times W$	1.2 3.6 3.8 4.0 4.0 3.6 3.0	7.2 3.6 2.0 2.4 4.5 3.4 3.0	1.2 2.0 3.4 4.4 4.0 3.8 3.2	5.0 4.6 3.8 4.4 4.0 3.8 3.2	3.6 3.8 3.4 4.4 4.0 3.4 3.2	17.8 3.4 3.8 3.8 4.4 3.4 3.0	0.2 0.8 1.0 4.2 5.0 4.6 4.2	3.8 4.4 4.4 4.6 4.0 3.2 2.8	1.6 3.0 2.4 2.4 4.6 5.4 3.0	6.6 3.0 1.6 2.4 4.6 5.4 0.6	1.8 2.4 1.0 3.2 4.6 5.6 2.4	4.2 3.4 4.4 4.2 4.2 5.8 2.6	20.2 3.0 0.6 2.8 4.0 5.4 4.2	0.0 0.8 0.8 4.0 5.0 6.0 5.4	4.2 1.8 2.8 3.8 4.0 2.6 2.6			
0.5	<i>Unmodifiziert</i>	3.4 14.5	3.4 14.9	3.3 4.2	3.3 13.5	3.2 13.1	3.2 13.7	4.7 11.3	3.4 13.4	3.4 4.5	3.9 14.8	4.1 15.2	3.9 4.9	3.8 13.8	3.8 13.7	3.9 13.9	5.1 10.8	3.9 13.6	3.9 5.4
	$r = 10^{-6}$	3.0 2.8 5.0 8.4 41.4 95.8	3.0 2.8 4.2 9.4 41.0 97.4	3.2 1.1 2.4 5.0 8.0 9.6	3.0 3.0 4.2 5.0 8.4 84.2	2.8 2.8 3.6 7.8 11.6 83.8	2.8 3.0 3.4 7.8 11.6 82.8	5.2 4.4 4.4 8.2 35.2 58.4	2.8 2.8 2.8 8.2 34.2 88.0	3.0 3.0 5.2 5.7 10.2 8.2	4.0 4.0 4.6 9.4 10.6 9.4	4.4 4.1 4.2 6.7 10.6 81.8	4.6 4.2 5.2 9.4 32.0 82.0	4.4 4.2 5.2 9.2 30.6 82.2	4.2 5.2 5.2 8.8 30.4 50.4	5.2 5.2 5.2 12.0 32.4 86.4	4.2 3.6 3.4 9.4 31.4 8.8	3.9 3.6 3.4 4.2 3.4 3.4	
	$W \times W$	1.8 3.2 2.2 3.6 3.6 3.2	6.0 3.2 1.8 3.1 3.5 3.0	1.4 1.4 1.8 3.2 3.5 3.4	5.0 5.2 4.2 3.0 3.2 3.6	3.4 4.0 4.0 3.2 3.6 3.6	13.4 4.0 5.4 3.0 3.2 3.2	0.4 0.6 0.8 5.2 5.2 4.8	3.6 4.4 4.4 3.6 3.4 3.4	5.6 2.2 1.8 4.4 5.0 4.2	1.6 2.0 2.4 5.6 5.6 4.2	1.6 1.6 1.8 5.8 5.0 <br; style="background-color: #0000ff;">5.0</br;>	5.0 5.4 5.2 4.4 4.4 4.4	12.0 3.8 3.8 4.4 6.4 4.4	0.0 1.2 0.8 0.8 4.4 3.8	4.0 4.0 4.0 4.0 4.0 4.0	5.2 4.8 4.8 5.0 5.0 3.4		
0.75	<i>Unmodifiziert</i>	1.8 18.8	1.8 19.2	1.9 6.5	1.8 19.2	1.8 19.3	1.7 19.7	4.2 20.2	1.8 19.1	1.9 7.5	2.0 18.8	2.0 19.0	2.1 6.1	2.2 19.2	2.0 19.4	2.0 19.8	4.3 20.2	2.0 19.2	2.0 7.2
	$r = 10^{-6}$	1.6 1.6 2.6 24.0 89.4 100.0	1.6 2.0 1.8 23.0 89.0 100.0	1.5 2.7 1.2 22.0 35.0 10.8	2.0 1.4 3.0 22.0 84.4 100.0	1.6 1.6 2.8 22.0 84.6 100.0	2.0 1.8 2.6 22.0 84.4 100.0	4.0 4.4 4.2 28.4 87.0 99.0	1.8 1.8 2.0 21.6 84.6 100.0	2.1 2.7 2.6 12.0 35.0 10.6	1.4 0.8 3.2 26.4 85.6 100.0	1.6 1.2 1.0 12.3 28.0 14.2	1.5 1.6 1.0 26.0 30.6 100.0	1.2 0.8 0.8 25.0 30.8 100.0	5.4 4.4 4.6 2.0 30.8 98.2	1.2 1.4 2.0 11.7 25.6 13.8	1.3 1.6 1.8 4.1 3.4 3.4		
	$W \times W$	0.0 0.4 0.4 1.4 2.6 2.0	5.8 0.6 1.0 1.4 2.4 2.4	0.4 1.6 3.6 4.6 2.2 2.2	0.4 4.8 5.2 4.6 5.2 2.2	4.6 3.4 5.2 5.4 5.2 5.0	9.4 5.2 1.0 3.4 5.2 5.0	0.0 1.0 0.4 1.2 2.0 2.0	3.6 3.0 5.4 4.6 2.2 2.4	4.4 5.2 5.4 3.6 2.2 2.8	0.0 0.0 0.0 1.4 1.6 2.4	0.8 1.2 1.0 3.0 2.0 2.8	5.0 5.2 5.4 3.6 3.2 4.4	5.4 5.0 6.0 6.4 6.2 5.0	0.0 0.6 0.6 4.2 4.2 2.4	3.8 5.2 5.0 4.4 4.2 2.0			

Tabelle A.1: \mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie für die \mathcal{H}_0 -Copulas. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen zwei höchsten Ablehnungsraten je Copula sind blau hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten ($W = \text{Worst}$, $A = \text{Average}$, $B = \text{Best}$). Im Fall *Unmodifiziert* ($N = 6\,000$) werden keine Modifikationen durchgeführt. †: Der Fall *Modifiziert* ($N = 6\,000$) umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Tests auf Radial-Symmetrie		<i>Clayton-Copula (\mathcal{H}_1)</i> \mathcal{H}_0 -Ablehnungsraten in %										<i>Gumbel-Copula (\mathcal{H}_1)</i> \mathcal{H}_0 -Ablehnungsraten in %													
Tau	Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π-Jitter	mix- r_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π-Jitter	mix- r_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks						
0.25	<i>Unmodifiziert</i>	71.8 72.1 71.8 72.2 71.9 72.2 74.3 71.8 71.9	21.5 21.6 21.5 21.4 21.6 21.4 23.7 21.4 21.5	51.9 52.0 51.9 52.6 52.8 54.5 46.0 53.8 56.7	25.2 25.4 16.7 24.9 24.9 25.8 22.0 25.6 18.5																				
	<i>Modifiziert†</i>	69.4 69.6 70.6 69.6 68.8 68.8 71.4 69.4 70.6	20.0 20.0 20.5 20.0 20.4 20.6 22.8 21.0 20.5	72.4 72.4 66.2 72.6 72.0 72.4 74.4 72.8 70.4	21.0 22.0 23.0 21.4 21.4 20.6 24.6 21.8 24.0	80.2 78.0 62.2 73.0 73.0 73.0 64.4 71.6 54.6	19.8 17.0 12.2 21.2 22.0 22.2 28.6 19.0 14.2	50.8 50.0 60.0 49.2 48.8 49.8 52.8 50.0 60.2	33.4 34.4 28.0 33.4 32.8 33.2 35.0 33.6 28.2	34.8 34.8 56.2 29.4 30.2 30.2 29.8 30.0 55.4	46.2 47.2 26.6 41.8 41.2 40.2 41.8 41.4 25.8	89.8 88.6 16.4 55.2 55.6 55.0 21.6 67.8 16.2													
	$r = 10^{-6}$	4.4 16.4 3.6 26.6 26.6 45.6 3.0 41.2 26.0	0.8 3.4 1.0 9.2 11.6 21.4 1.4 8.8 8.6	13.4 13.8 24.4 41.0 43.2 40.8 12.6 39.6 42.2	2.0 1.8 5.4 14.4 14.0 14.0 5.8 12.2 13.2	$r = 10^{-4}$	2.0 1.8 5.8 12.4 11.4 13.0 4.0 13.4 10.6	$r = 10^{-2}$	11.8 11.0 22.8 40.0 40.6 42.4 14.6 41.8 42.6	24.4 24.0 26.9 24.8 24.6 24.4 27.2 24.2 26.3	70.6 70.4 69.4 70.4 70.8 70.6 72.2 70.6 68.8	19.4 20.4 20.5 20.6 19.8 21.0 24.6 20.0 18.8	77.2 76.8 77.0 76.6 76.4 77.6 79.2 77.8 77.7	24.2 23.8 24.2 24.6 24.2 24.2 24.2 26.2 23.8 23.8	$p = 0.025$	50.8 50.0 60.0 49.2 48.8 49.8 52.8 50.0 60.2	88.9 88.4 95.4 87.4 88.2 88.4 90.6 88.4 95.0	$p = 0.050$	58.0 58.4 90.6 50.4 50.2 50.8 51.8 49.4 89.2	81.0 76.8 96.0 14.0 14.0 14.2 3.8 14.0 95.0	$p = 0.100$				
	W×W	4.4 16.4 3.6 26.6 26.6 45.6 3.0 41.2 26.0	0.4 2.0 0.6 17.6 19.6 29.0 2.0 21.6 18.8	W×A	13.4 13.8 24.4 41.0 43.2 40.8 12.6 39.6 42.2	W×B	11.8 11.0 22.8 40.0 40.6 42.4 14.6 41.8 42.6	A×A	70.6 70.4 69.4 70.4 70.8 70.6 72.2 70.6 68.8	A×B	77.2 76.8 77.0 76.6 76.4 77.6 79.2 77.8 77.7	B×B	74.6 74.8 74.2 74.8 74.6 74.4 76.4 73.8 74.0												
0.5	<i>Unmodifiziert</i>	99.4 99.4 99.4 99.4 99.3 99.4 99.5 99.3 99.4	41.6 41.7 41.7 41.8 41.7 41.7 47.3 41.9 41.9	73.0 74.5 76.2 76.4 76.5 77.9 68.6 78.5 84.8	43.0 42.6 32.2 48.3 48.4 49.3 46.8 48.2 36.3																				
	<i>Modifiziert†</i>	99.6 99.6 99.6 99.8 99.6 99.6 99.6 99.6 99.6	40.4 40.8 41.0 41.0 41.4 39.6 46.2 41.0 39.7	99.2 99.2 100.0 99.2 99.2 99.4 99.0 100.0	43.4 43.4 36.4 44.0 43.6 43.0 49.8 43.0 39.8	$r = 10^{-6}$	99.4 99.0 95.0 99.0 99.0 99.0 98.8 93.2	$r = 10^{-4}$	89.0 88.4 95.4 87.4 88.2 88.4 90.6 88.4 95.0	$r = 10^{-2}$	80.0 84.6 21.0 44.4 43.8 44.4 54.0 37.4 25.0	$p = 0.025$	72.4 72.0 59.7 72.6 71.8 71.8 76.0 72.0 59.1	$p = 0.050$	91.8 92.4 58.6 90.0 90.0 90.4 91.0 91.4 58.6	$p = 0.100$	99.8 99.8 34.6 98.2 98.6 98.6 89.6 99.0 35.2								
	W×W	4.4 28.2 6.0 41.8 42.6 58.2 10.0 68.2 43.0	0.4 2.0 0.6 17.6 19.6 29.0 2.0 21.6 18.8	W×A	24.6 24.6 33.2 66.2 63.8 64.2 35.8 64.4 62.2	W×B	22.8 22.6 33.8 61.8 64.4 63.6 35.0 63.2 64.0	A×A	98.6 98.8 99.1 98.6 98.8 98.8 98.8 98.6 99.1	A×B	99.6 99.6 99.5 99.6 99.6 99.6 100.0 99.6 99.5	B×B	99.2 99.2 99.2 99.2 99.0 99.2 99.4 99.0 99.2												
0.75	<i>Unmodifiziert</i>	100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0	35.5 35.4 35.1 35.1 35.1 35.1 46.2 35.2 35.3	68.8 71.5 76.3 74.2 74.4 75.5 66.6 77.3 82.4	41.6 40.3 29.6 46.0 45.8 46.0 48.8 45.1 32.1																				
	<i>Modifiziert†</i>	100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0	33.0 32.6 31.5 33.4 33.0 32.4 47.0 32.8 32.2	100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0	38.2 37.0 33.3 37.4 38.2 37.0 48.4 37.0 35.8	$r = 10^{-6}$	99.2 99.2 99.2 99.2 99.2 99.2 99.2 99.2 99.2	$r = 10^{-4}$	99.2 99.2 99.2 99.2 99.2 99.2 99.2 99.2 99.2	$r = 10^{-2}$	99.2 99.2 99.2 99.2 99.2 99.2 99.2 99.2 99.2	$p = 0.025$	72.2 72.2 92.4 70.6 71.2 70.4 75.2 71.0 92.2	$p = 0.050$	34.2 34.2 90.0 23.4 24.6 24.2 23.8 24.6 88.2	$p = 0.100$	92.8 93.6 99.0 60.6 61.2 61.2 28.4 65.2 97.6								
	W×W	1.2 31.4 5.4 35.6 32.4 47.0 9.6 62.2 34.8	0.2 2.4 1.0 10.6 11.0 18.0 1.2 14.2 10.2	W×A	13.0 14.0 31.4 49.2 52.2 50.2 31.0 52.8 51.4	W×B	12.6 12.8 31.8 51.6 50.6 53.2 30.6 52.0 49.8	A×A	100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0	A×B	100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0	B×B	100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0 100.0												

Tabelle A.2: \mathcal{H}_0 -Ablehnungsraten der Tests auf Radial-Symmetrie für die \mathcal{H}_1 -Copulas. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen zwei höchsten Ablehnungsraten je Copula sind blau hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). Im Fall *Unmodifiziert* ($N = 6\,000$) werden keine Modifikationen durchgeführt. †: Der Fall *Modifiziert* ($N = 6\,000$) umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Tests auf Austauschbarkeit		<i>Clayton-Copula (\mathcal{H}_0)</i> \mathcal{H}_0 -Ablehnungsraten in %									<i>Gumbel-Copula (\mathcal{H}_0)</i> \mathcal{H}_0 -Ablehnungsraten in %									
Tau τ	Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix- τ_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix- τ_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks	
0.25	<i>Unmodifiziert</i> <i>Modifiziert</i> [†]	3.5 34.2	3.3 33.4	3.4 26.9	3.3 3.8	3.4 3.8	3.3 7.3	3.5 2.2	3.4 4.0	3.3 4.3	3.6 34.3	3.5 33.7	3.4 26.4	3.5 4.0	3.4 3.8	3.5 7.5	3.7 2.5	3.4 4.1	3.5 4.2	
		$r = 10^{-6}$	2.4 3.8	2.8 3.8	2.3 7.1	2.6 3.6	2.8 4.0	2.8 3.6	2.6 4.0	2.3 6.0	2.8 3.8	2.4 4.2	3.0 5.3	2.6 3.8	2.4 4.2	2.2 4.0	2.8 4.2	2.6 4.4	2.8 5.3	
		$r = 10^{-4}$	13.0 4.0	11.4 4.6	11.2 5.0	3.0 3.8	3.2 4.0	3.0 3.4	2.0 3.4	2.8 4.0	3.6 4.8	10.6 3.4	8.8 5.3	4.2 3.0	4.0 3.2	5.0 3.2	2.8 3.0	3.6 3.0	3.8 4.1	
		$r = 10^{-2}$	4.0 7.6	4.6 7.0	5.0 7.2	3.8 3.4	3.6 3.6	3.8 3.8	3.2 3.2	3.6 3.6	4.8 5.0	3.4 8.6	5.3 6.2	3.0 4.8	3.2 4.8	4.6 4.6	5.0 5.0	5.2 5.2	5.4 5.4	
		$p = 0.025$	83.0 83.0	73.8 73.8	11.4 11.4	4.4 4.4	4.6 4.6	4.6 4.6	1.8 1.8	4.6 4.6	6.0 6.0	77.4 83.2	8.4 5.6	5.6 5.2	5.2 5.0	1.6 1.6	5.2 5.2	5.2 5.2	5.2 5.2	
		W×W	98.2 92.6	98.8 93.6	98.2 68.4	4.8 5.2	3.6 4.4	45.8 4.2	0.0 0.0	5.0 4.6	4.4 6.4	98.2 96.4	98.0 68.4	4.4 4.8	4.4 3.8	46.2 5.2	0.2 0.0	4.0 4.8	4.8 3.6	
		W×A	95.6 92.6	95.6 93.6	68.4 68.4	5.2 5.2	4.4 4.2	0.0 0.0	7.2 7.2	5.4 5.4	94.4 93.6	71.4 6.4	4.2 4.2	4.0 6.0	0.0 0.0	7.2 7.2	5.8 5.8			
		W×B	92.6 95.6	95.6 95.6	71.8 71.8	5.6 5.6	6.0 6.0	5.8 5.8	0.0 0.0	7.2 5.4	9.6 2.2	2.4 2.4	3.1 2.2	2.2 2.2	2.6 2.6	2.8 2.8	2.5 2.5			
		A×A	2.6 3.0	2.6 2.6	2.3 3.4	2.4 2.8	2.2 3.0	2.4 3.0	2.4 3.0	2.3 2.7	2.3 3.6	2.2 3.2	2.4 3.0	3.1 3.2	2.2 2.6	2.4 3.2	3.2 3.2	2.4 3.2		
		A×B	5.2 3.0	4.8 2.6	6.9 3.4	4.2 2.8	4.4 3.0	4.8 3.0	3.2 3.0	4.0 3.0	7.4 2.7	8.6 4.0	15.6 3.0	4.4 4.0	4.4 4.0	4.0 4.0	12.8 4.0	5.4 3.8		
		B×B	5.2 4.4	4.8 4.0	6.9 4.4	4.0 4.0	4.8 4.8	4.4 4.4	4.4 4.4	4.0 4.0	12.0 3.6	72.6 4.0	8.6 3.0	6.2 3.4	6.0 4.0	6.4 3.4	1.2 4.0	5.4 3.8		
		$p = 0.050$	14.2 83.4	13.6 77.6	30.8 13.0	7.2 7.2	6.8 7.4	7.6 7.2	3.0 0.4	7.6 7.0	27.4 12.0	9.8 72.6	9.8 8.6	4.4 6.2	4.4 6.0	4.0 6.4	2.0 1.2	4.4 5.4	12.8 7.4	
		W×W	98.4 99.8	98.2 99.8	9.8 71.4	9.0 11.8	48.2 11.2	0.0 12.4	10.6 0.0	10.8 11.2	95.2 100.0	95.0 61.2	6.0 9.6	6.8 8.0	44.4 9.0	0.2 0.0	8.8 8.0	7.6 8.2		
		W×A	99.8 99.8	99.6 99.8	66.8 71.4	8.8 11.8	10.0 11.2	10.6 12.4	0.0 0.0	11.0 14.8	99.8 100.0	67.8 61.2	9.2 9.6	10.6 8.0	11.0 9.0	0.0 0.0	11.0 8.0	9.4 8.2		
		W×B	99.8 99.8	99.8 99.8	71.4 71.4	11.8 11.8	11.2 12.4	12.4 12.4	0.0 0.0	14.8 11.2	99.8 100.0	21.6 16.0	2.8 2.8	3.4 3.2	3.2 2.8	2.8 2.8	3.4 3.4	5.4 5.4		
		A×A	4.2 4.2	4.6 4.6	4.7 4.7	3.8 3.8	3.8 3.8	3.8 3.8	3.6 3.6	3.4 2.8	4.7 3.9	3.0 3.4	3.1 2.8	3.2 3.0	3.0 2.8	2.8 2.8	2.5 3.0			
		A×B	5.2 3.4	4.8 2.8	6.9 3.2	4.2 2.2	4.4 3.2	4.8 3.4	3.2 2.8	4.0 2.8	7.4 3.9	8.6 3.6	15.6 3.6	4.4 3.8	4.4 3.8	4.0 3.8	2.8 3.2	3.4 3.8		
		B×B	5.2 4.2	4.8 3.8	6.9 4.2	4.0 3.8	4.8 3.8	4.4 4.0	4.4 4.0	4.2 4.2	12.0 3.8	72.6 3.6	8.6 3.6	6.2 3.8	6.0 3.8	6.4 3.8	1.2 3.2	5.4 3.8		
0.5	<i>Unmodifiziert</i> <i>Modifiziert</i> [†]	3.6 37.9	3.4 36.6	3.5 30.8	3.6 6.0	3.6 6.2	3.7 9.6	3.6 2.5	3.5 6.6	3.7 9.2	2.8 36.5	2.9 34.9	2.8 27.0	2.9 4.6	2.9 4.7	2.8 8.0	2.8 1.9	2.8 4.8	2.8 6.0	
		$r = 10^{-6}$	4.6 4.6	4.8 4.8	4.3 2.6	4.4 5.0	4.6 5.6	4.0 4.6	3.8 4.0	4.6 5.2	4.5 2.6	2.2 2.6	1.8 1.0	2.2 2.0	2.2 2.0	1.8 2.2	2.2 2.4	2.4 2.2	2.6 1.0	
		$r = 10^{-4}$	32.4 5.2	24.6 4.8	24.6 6.9	4.4 4.2	4.6 4.4	4.8 4.8	4.0 3.2	3.6 4.0	2.2 7.4	32.8 4.2	25.0 3.8	21.6 3.4	2.8 3.8	3.4 3.2	3.2 2.8	2.8 3.4	2.6 5.4	
		$r = 10^{-2}$	5.2 14.2	4.8 13.6	6.9 30.8	4.2 7.2	4.4 6.8	4.8 7.6	3.2 3.0	4.0 7.6	7.4 27.4	9.4 8.4	8.8 15.6	4.4 4.4	4.4 4.0	4.0 4.0	2.0 2.0	4.4 4.4	12.8 12.8	
		$p = 0.025$	14.2 83.4	13.6 77.6	30.8 13.0	7.2 7.2	6.8 7.4	7.6 7.2	0.4 0.4	7.0 7.0	12.0 12.0	72.6 86.4	8.6 82.0	13.4 17.2	17.0 17.0	1.0 1.0	16.4 16.4	13.4 13.4	1.2 1.4	5.4 2.0
		W×W	98.4 97.0	98.2 96.2	9.8 9.6	9.0 8.8	48.2 40.2	0.0 0.0	10.6 9.0	10.8 12.8	95.0 100.0	95.0 69.2	6.0 12.8	6.8 16.0	44.4 13.4	0.2 0.0	8.8 7.2	7.6 3.8		
		W×A	99.8 100.0	99.6 100.0	66.8 69.2	8.8 12.8	10.0 16.0	13.4 13.4	0.0 0.0	11.0 13.8	11.8 12.2	99.8 100.0	67.8 68.6	9.2 6.4	10.6 6.8	11.0 0.0	0.0 0.0	11.0 7.0	9.4 3.8	
		W×B	100.0 100.0	100.0 71.0	71.0 14.4	11.6 14.8	11.6 15.4	11.8 0.0	12.4 14.0	12.1 13.4	100.0 100.0	81.6 48.8	3.2 4.4	3.6 7.0	3.8 6.0	0.6 0.0	0.6 0.0	2.6 5.8	3.8 4.0	
		A×A	3.4 3.4	3.2 2.8	2.5 2.0	2.0 2.0	2.2 2.2	1.2 1.2	2.0 2.0	2.5 2.5	2.5 2.5	3.0 2.2	1.1 2.2	1.1 2.2	1.1 2.2	1.1 2.2	1.1 2.2	2.2 2.2	0.3 0.3	
		A×B	2.6 2.6	2.2 2.2	2.7 2.7	2.0 2.0	1.8 2.0	2.0 2.0	1.6 1.6	2.2 2.2	2.7 2.7	4.4 4.4	3.7 4.2	3.6 4.2	4.2 4.0	4.0 4.0	1.8 1.8	3.4 3.4	4.2 4.2	
		B×B	2.8 2.8	2.6 2.6	2.6 2.6	2.6 2.6	3.2 3.2	1.8 1.8	3.0 3.0	2.8 2.8	2.0 2.0	2.2 2.2	2.6 2.2	2.2 2.2	2.2 2.2	2.2 2.2	1.4 1.4	2.6 2.6	2.0 2.0	
		$p = 0.050$	14.2 86.4	13.6 82.0	38.0 13.4	7.4 17.2	7.6 17.0	10.1 17.0	1.1 1.0	7.4 16.4	12.5 13.4	42.0 78.2	40.6 66.2	36.2 11.6	4.0 5.8	4.5 6.4	6.6 6.4	1.0 1.0	4.4 1.0	8.0 11.0
		$r = 10^{-6}$	2.8 2.4	2.4 2.2	3.2 4.1	2.6 1.8	2.8 1.8	2.4 1.8	0.8 1.6	3.0 1.8	3.4 5.4	2.6 2.2	3.2 1.1	2.9 2.8	3.0 2.6	3.0 2.6	1.6 1.2	2.8 2.2	2.9 3.3	
		$r = 10^{-4}$	98.4 98.4	89.6 89.6	78.2 78.2	3.2 3.2	2.8 2.8	3.2 3.2	0.4 0.4	2.2 2.2	2.8 2.8	98.2 80.0	93.0 81.6	81.6 32.2	3.2 3.6	3.8 3.8	0.6 0.6	2.6 2.6	3.8 3.8	
		$r = 10^{-2}$	9.8 9.8	9.6 9.6	23.1 8.6	8.6 8.8	8.8 9.0	9.0 9.2	2.0 9.2	9.2 6.0	22.1 6.0	8.0 7.0	14.3 10.4	6.0 6.0	6.0 5.8	5.8 3.0	3.0 3.0	14.6 14.6	35.8 35.8	
		$p = 0.025$	20.4 9.8	19.6 9.6	45.6 23.1	11.6 8.6	11.6 8.8	11.8 9.0	3.2 2.0	12.4 9.2	45.4 22.1	11.6 8.0	38.8 32.2	5.8 3.2	5.6 3.8	5.8 3.0	1.6 3.0	6.0 5.8	14.6 35.8	
		$p = 0.050$	20.4 9.8	19.6 9.6	45.6 23.1	11.6 8.6	11.6 8.8	11.8 9.0	3.2 2.0	12.4 9.2	45.4 22.1	11.6 8.0	38.8 32.2	5.8 3.2	5.6 3.8	5.8 3.0	1.6 3.0	6.0 5.8	14.6 35.8	
		$p = 0.100$	86.4 86.4	82.0 82.0	13.4 13.4	17.2 17.2	17.0 17.0	17.0 17.0	1.0 1.0	16.4 16.4	13.4 13.4	78.2 78.2	66.2 66.2	11.6 11.6	5.8 5.8	6.4 6.4	6.4 6.4	1.0 1.0	6.6 6.6	11.0 11.0

Tabelle A.3: \mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit für die \mathcal{H}_0 -Copulas. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen zwei höchsten Ablehnungsraten je Copula sind blau hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). Im Fall *Unmodifiziert* ($N = 6\,000$) werden keine Modifikationen durchgeführt. †: Der Fall *Modifiziert* ($N = 6\,000$) umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Tests auf Austauschbarkeit		<i>Khoudraji-Clayton-Copula (\mathcal{H}_1)</i> \mathcal{H}_0 -Ablehnungsraten in %									<i>Khoudraji-Gumbel-Copula (\mathcal{H}_1)</i> \mathcal{H}_0 -Ablehnungsraten in %									
Tau τ	Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix- r_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix- r_n -Jitter	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks	
0.25	<i>Unmodifiziert</i> <i>Modifiziert</i> [†]	3.8 32.2	3.8 31.4	3.7 25.5	3.8 3.5	3.6 3.6	3.8 7.8	4.3 2.6	3.6 3.4	3.7 3.9	9.9 37.7	9.9 36.6	9.9 30.8	9.9 8.6	9.9 8.7	9.9 12.8	10.7 7.6	9.9 8.6	9.9 8.5	
		$r = 10^{-6}$	3.4	3.4	3.9	3.4	3.6	3.4	3.6	3.4	13.2	13.2	13.2	13.8	13.6	13.2	13.8	13.6	13.5	
	$r = 10^{-4}$ $r = 10^{-2}$ $p = 0.025$ $p = 0.050$ $p = 0.100$	2.4	2.4	2.6	3.2	3.0	3.0	3.2	3.0	2.6	11.8	11.8	13.3	11.2	10.8	11.0	12.6	11.2	12.2	
		10.6	7.4	8.6	3.0	3.0	3.4	3.0	3.0	3.2	20.6	18.0	16.8	8.8	8.8	9.0	9.0	7.8	6.2	
		4.8	5.2	5.3	4.8	4.8	5.0	5.0	4.8	4.7	11.4	11.8	11.0	11.2	11.4	11.0	11.8	11.2	10.8	
		7.4	7.4	4.6	4.2	4.0	3.8	3.6	3.4	4.0	12.2	10.6	10.6	7.6	7.4	7.0	7.4	7.6	9.8	
		82.4	74.0	7.2	2.6	3.2	3.0	1.6	3.0	3.6	84.8	77.2	10.2	7.8	8.2	7.4	5.0	7.8	8.8	
		W×W	97.6	97.6	97.2	4.0	5.0	54.8	0.2	3.4	4.6	98.4	97.8	98.2	4.0	5.0	57.0	0.4	3.8	5.0
		W×A	84.6	84.4	66.6	2.8	4.6	3.8	0.0	3.8	5.0	84.2	84.2	66.0	6.4	4.8	4.4	0.0	6.8	5.6
		W×B	83.8	83.6	67.4	4.4	3.0	4.8	0.0	3.2	4.0	86.2	86.6	72.4	4.8	4.6	6.2	0.0	5.0	5.2
		A×A	3.8	3.4	4.5	3.4	3.4	3.0	4.0	3.8	3.6	10.6	9.8	11.9	10.0	10.8	10.2	11.0	10.4	11.1
		A×B	3.8	4.2	4.1	3.8	3.8	3.4	4.6	3.8	4.1	9.2	8.4	10.0	9.0	9.4	8.6	9.6	8.8	8.8
		B×B	1.8	2.4	2.2	2.2	2.0	2.0	2.2	2.2	2.2	9.4	10.0	9.4	9.0	9.2	9.2	10.0	9.6	9.6
0.5	<i>Unmodifiziert</i> <i>Modifiziert</i> [†]	10.2 37.6	10.2 36.6	10.0 30.3	10.1 9.1	10.0 9.0	10.0 12.8	11.1 7.2	10.1 8.8	10.1 8.6	53.1 68.0	53.3 67.5	53.2 62.6	53.2 41.4	53.5 41.0	53.2 45.9	55.1 37.2	53.0 41.5	53.3 40.6	
		$r = 10^{-6}$	11.2	12.2	12.1	11.8	11.0	11.4	11.6	11.4	57.4	57.4	56.7	56.8	57.4	57.4	57.8	56.8	56.5	
		$r = 10^{-4}$	9.8	9.6	10.8	10.0	10.4	9.4	11.2	9.8	56.4	55.2	62.5	55.8	55.8	57.0	58.8	56.8	59.4	
		$r = 10^{-2}$	21.6	18.8	14.0	10.8	11.8	10.4	11.2	11.2	74.6	71.2	56.0	52.6	52.4	53.6	51.0	33.8		
		$p = 0.025$	11.0	11.0	10.8	10.2	10.4	10.8	10.8	10.0	48.4	49.0	53.1	47.4	48.4	48.4	49.2	47.2	51.4	
		$p = 0.050$	14.0	13.2	10.6	9.4	9.4	9.4	9.4	8.8	47.6	48.4	50.4	36.6	37.0	37.2	35.6	36.0	48.8	
		$p = 0.100$	82.8	74.2	13.0	7.6	7.4	6.8	3.4	7.0	9.6	91.2	88.6	44.0	27.2	27.6	27.4	19.4	27.0	42.6
		W×W	96.4	96.2	96.4	5.2	6.8	54.8	0.4	5.4	4.8	98.8	98.8	98.4	14.4	12.0	64.2	2.0	16.4	14.0
		W×A	86.4	85.4	69.2	9.0	8.6	8.2	0.4	8.2	7.6	92.4	91.6	83.6	23.0	20.6	23.4	4.6	23.0	23.0
		W×B	90.8	90.4	70.6	7.2	5.0	5.6	0.2	6.8	5.8	91.0	91.0	81.6	24.6	23.0	24.8	2.2	26.8	24.6
		A×A	8.6	9.0	8.7	9.4	8.8	8.8	9.0	8.2	9.3	53.4	55.0	55.3	53.6	53.6	54.0	54.2	53.8	54.7
		A×B	8.4	8.8	8.0	8.6	8.0	7.8	9.0	8.2	8.3	51.0	51.4	50.5	51.4	51.2	51.2	53.2	51.0	49.8
		B×B	9.8	10.4	10.2	10.0	10.0	10.0	10.4	10.0	9.8	54.4	52.6	53.6	53.2	53.4	53.4	55.0	52.2	52.8
0.75	<i>Unmodifiziert</i> <i>Modifiziert</i> [†]	56.1 69.2	55.5 68.7	55.8 64.1	55.8 44.1	55.7 44.7	55.6 48.7	57.6 38.9	55.9 44.2	55.5 43.5	95.4 95.2	95.3 95.2	95.3 94.7	95.5 77.2	95.6 77.2	95.4 81.4	96.0 69.6	95.4 77.0	95.4 76.5	
		$r = 10^{-6}$	57.2	57.6	59.1	57.8	58.0	57.8	59.8	58.0	96.2	96.0	96.4	96.6	96.2	96.4	96.4	96.4	96.4	
		$r = 10^{-4}$	55.4	54.2	54.8	54.8	54.6	54.6	56.0	54.6	94.4	94.2	93.4	94.0	94.4	94.4	95.6	94.6	94.7	
		$r = 10^{-2}$	72.0	69.8	55.8	56.2	56.4	56.0	56.4	55.0	99.4	98.8	89.0	96.2	96.4	96.2	96.2	95.2	81.8	
		$p = 0.025$	49.2	48.2	55.1	47.8	48.0	46.8	48.6	48.0	90.6	90.8	94.1	90.8	90.0	90.8	90.6	90.4	94.1	
		$p = 0.050$	54.2	53.4	56.3	46.2	47.2	46.2	45.6	47.2	86.8	87.2	92.2	83.0	82.4	83.2	81.6	82.6	90.4	
		$p = 0.100$	91.8	89.4	46.0	32.4	32.4	32.8	20.0	32.2	95.2	95.2	89.0	63.0	63.6	64.4	45.8	63.6	86.8	
		W×W	97.2	97.0	96.8	11.6	12.0	63.4	3.0	11.6	11.8	99.0	99.0	99.2	25.4	23.2	72.2	5.4	22.4	22.8
		W×A	94.2	94.2	85.4	25.6	27.6	29.8	2.6	26.8	28.6	98.4	98.4	98.0	43.8	44.6	44.6	17.4	45.0	41.8
		W×B	92.2	93.4	83.6	30.6	31.0	29.6	3.6	29.8	28.8	96.8	97.0	98.2	48.0	49.4	48.8	18.2	48.4	48.0
		A×A	58.0	57.4	55.0	56.2	58.4	57.8	58.8	56.8	95.4	95.4	95.4	95.0	95.6	94.8	96.0	94.8	95.4	
		A×B	55.0	54.8	54.0	54.8	55.6	54.6	56.2	55.0	96.0	96.0	96.0	95.8	96.2	97.0	95.8	96.5		
		B×B	54.4	54.4	54.4	55.4	55.2	55.0	56.2	55.4	94.8	95.0	94.8	94.2	95.0	95.0	95.4	94.4	95.2	

Tabelle A.4: \mathcal{H}_0 -Ablehnungsraten der Tests auf Austauschbarkeit für die \mathcal{H}_1 -Copulas.

Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen zwei höchsten Ablehnungsraten je Copula sind blau hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). Im Fall *Unmodifiziert* ($N = 6\,000$) werden keine Modifikationen durchgeführt. †: Der Fall *Modifiziert* ($N = 6\,000$) umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Anhang B

Detaillierte Testresultate zur MPLE-Modellanpassung

Die Tabellen B.1 bis B.12 umfassen die Simulationsergebnisse bzgl. der MPLE-Modellanpassung der Copulas unter der Annahme, dass die zugrundeliegenden Copulafamilien bekannt sind.

Mit den Schätzungen $\hat{\tau}_n(x_k) = \tau(\hat{\vartheta}_n(x_k))$ auf Grundlage der $N = 500$ Stichproben x_1, \dots, x_N der Größe $n = 250$ wurden der Bias und RMSE empirisch bestimmt. In den Tabellen sind je Kombination aus Copula, Kendall'schem Tau und Datenmodifikation der Bias und RMSE dargestellt. Die Werte wurden auf vier Nachkommastellen gerundet.

Die Schätzungen wurden unter Wahl der Standardparameter in den Methoden und bei der Intervall-Zensur mit dem Optimierungsalgorithmus *BFGS* durchgeführt. Sind die Methoden nicht konvergiert und konnten auf den Datengrundlagen nicht initialisiert werden, so sind diese als “–” eingetragen.

Rote Einträge geben Ergebnisse mit gleichzeitig minimalem Bias und RMSE an. Blaue Einträge umfassen die Ergebnisse mit minimalen Werten in einer dieser Kennzahlen.

Die Ergebnisse sind die Grundlage der Abbildungen 6.9 bis 6.12.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	Gauß'sche Copula mit Kendall'schem Tau $\tau = 0.25$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	-0.0111 (± 0.0652)	-0.0234 (± 0.0643)	+0.0095 (± 0.0639)	-0.0400 (± 0.0857)	-0.0118 (± 0.0676)	-0.0098 (± 0.0604)	-0.0090 (± 0.0603)	-0.0111 (± 0.0652)	-0.0385 (± 0.0831)	-0.0235 (± 0.0783)	-0.0139 (± 0.0617)	+0.0063 (± 0.0521)
$r = 10^{-6}$	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)	+0.0079 (± 0.0398)
$r = 10^{-4}$	+0.0059 (± 0.0380)	+0.0059 (± 0.0380)	+0.0059 (± 0.0381)	+0.0059 (± 0.0380)	+0.0059 (± 0.0381)	+0.0059 (± 0.0380)	+0.0059 (± 0.0381)					
$r = 10^{-2}$	+0.0068 (± 0.0369)	+0.0045 (± 0.0365)	+0.0076 (± 0.0430)	+0.0054 (± 0.0365)	+0.0048 (± 0.0364)	+0.0048 (± 0.0364)	+0.0064 (± 0.0367)	+0.0068 (± 0.0369)	+0.0030 (± 0.0360)	+0.0079 (± 0.0438)	+0.0038 (± 0.0362)	+0.0084 (± 0.0445)
$p = 0.025$	-0.0302 (± 0.0508)	-0.0303 (± 0.0508)	-0.0082 (± 0.0445)	-0.0333 (± 0.0527)	-0.0302 (± 0.0508)	-0.0302 (± 0.0508)	-0.0314 (± 0.0516)	-0.0302 (± 0.0508)	-0.0336 (± 0.0529)	-0.0185 (± 0.0456)	-0.0315 (± 0.0517)	-0.0173 (± 0.0446)
$p = 0.050$	-0.0561 (± 0.0704)	-0.0569 (± 0.0710)	+0.0094 (± 0.0566)	-0.0632 (± 0.0762)	-0.0561 (± 0.0704)	-0.0561 (± 0.0704)	-0.0595 (± 0.0733)	-0.0561 (± 0.0704)	-0.0647 (± 0.0772)	-0.0165 (± 0.0504)	-0.0607 (± 0.0741)	-0.0159 (± 0.0500)
$p = 0.100$	-0.0913 (± 0.1018)	-0.0991 (± 0.1090)	+0.0749 (± 0.0978)	-0.1058 (± 0.1145)	-0.0913 (± 0.1018)	-0.0913 (± 0.1018)	-0.0986 (± 0.1083)	-0.0913 (± 0.1018)	-0.1147 (± 0.1228)	+0.0058 (± 0.0518)	-0.1096 (± 0.1183)	+0.0061 (± 0.0514)
W×W	+0.0285 (± 0.1407)	-0.1055 (± 0.1291)	+0.0213 (± 0.1429)	-0.1428 (± 0.1639)	+0.0219 (± 0.1541)	+0.0456 (± 0.1120)	+0.0177 (± 0.0891)	+0.0285 (± 0.1407)	-0.1143 (± 0.1435)	-0.1381 (± 0.1618)	-0.0256 (± 0.0875)	+0.0206 (± 0.0903)
W×A	-0.0104 (± 0.0472)	-0.0128 (± 0.0484)	-0.0103 (± 0.0474)	-0.0859 (± 0.1237)	-0.0104 (± 0.0472)	-0.0104 (± 0.0472)	+0.0138 (± 0.0602)	-0.0104 (± 0.0472)	-0.0819 (± 0.1192)	-0.0764 (± 0.1220)	+0.0135 (± 0.0600)	+0.0208 (± 0.0629)
W×B	-0.0132 (± 0.0458)	-0.0132 (± 0.0458)	-0.0131 (± 0.0458)	-0.0871 (± 0.1235)	-0.0132 (± 0.0458)	-0.0132 (± 0.0458)	+0.0110 (± 0.0577)	-0.0132 (± 0.0458)	-0.0881 (± 0.1238)	-0.0809 (± 0.1250)	+0.0110 (± 0.0577)	+0.0180 (± 0.0605)
A×A	+0.0069 (± 0.0389)	+0.0069 (± 0.0389)	+0.0069 (± 0.0390)	+0.0069 (± 0.0389)	+0.0069 (± 0.0392)	+0.0069 (± 0.0392)	+0.0078 (± 0.0392)					
A×B	+0.0068 (± 0.0411)	+0.0068 (± 0.0411)	+0.0068 (± 0.0410)	+0.0068 (± 0.0411)	+0.0068 (± 0.0414)	+0.0068 (± 0.0411)	+0.0077 (± 0.0414)					
B×B	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)	+0.0053 (± 0.0375)

Tabelle B.1: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gauß'sche Copula mit $\tau = 0.25$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

$\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	<i>Gauß'sche Copula mit Kendall'schem Tau</i> $\tau = 0.50$ <i>Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.</i>											
<i>Modifikation der Daten*</i>	<i>Reference</i>	<i>V-Reduct.</i>	<i>G-Reduct.</i>	<i>Rnd-Ranks</i>	<i>Π-Jitter</i>	<i>mix$\hat{\tau}_n$-Jitter</i>	<i>Censoring</i>	<i>Kojadinovic</i>	<i>V-Rd. Rnd-Ranks</i>	<i>G-Rd. Rnd-Ranks</i>	<i>V-Rd. Censoring</i>	<i>G-Rd. Censoring</i>
<i>Modifiziert†</i>	-0.0488 (± 0.0959)	-0.0633 (± 0.1170)	-0.0156 (± 0.0683)	-0.0885 (± 0.1591)	-0.0490 (± 0.0959)	-0.0491 (± 0.0959)	-0.0272 (± 0.0853)	-0.0488 (± 0.0959)	-0.0837 (± 0.1493)	-0.0580 (± 0.1397)	-0.0344 (± 0.0899)	+0.0009 (± 0.0486)
$r = 10^{-6}$	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)	+0.0049 (± 0.0301)
$r = 10^{-4}$	+0.0056 (± 0.0292)	+0.0056 (± 0.0292)	+0.0056 (± 0.0292)	+0.0055 (± 0.0292)	+0.0056 (± 0.0292)	+0.0056 (± 0.0292)	+0.0056 (± 0.0292)	+0.0056 (± 0.0292)	+0.0055 (± 0.0292)	+0.0056 (± 0.0292)	+0.0056 (± 0.0292)	+0.0056 (± 0.0292)
$r = 10^{-2}$	+0.0059 (± 0.0301)	+0.0027 (± 0.0296)	+0.0059 (± 0.0332)	+0.0041 (± 0.0297)	+0.0036 (± 0.0299)	+0.0032 (± 0.0300)	+0.0056 (± 0.0301)	+0.0059 (± 0.0301)	-0.0020 (± 0.0287)	+0.0044 (± 0.0346)	-0.0007 (± 0.0286)	+0.0064 (± 0.0351)
$p = 0.025$	-0.0812 (± 0.0900)	-0.0813 (± 0.0900)	-0.0408 (± 0.0603)	-0.0865 (± 0.0951)	-0.0812 (± 0.0900)	-0.0812 (± 0.0900)	-0.0803 (± 0.0892)	-0.0812 (± 0.0900)	-0.0866 (± 0.0952)	-0.0535 (± 0.0684)	-0.0804 (± 0.0893)	-0.0496 (± 0.0652)
$p = 0.050$	-0.1314 (± 0.1372)	-0.1319 (± 0.1376)	-0.0160 (± 0.0653)	-0.1435 (± 0.1491)	-0.1314 (± 0.1372)	-0.1314 (± 0.1372)	-0.1313 (± 0.1373)	-0.1314 (± 0.1372)	-0.1450 (± 0.1507)	-0.0462 (± 0.0743)	-0.1324 (± 0.1385)	-0.0411 (± 0.0687)
$p = 0.100$	-0.2011 (± 0.2060)	-0.2059 (± 0.2107)	+0.0470 (± 0.0752)	-0.2249 (± 0.2295)	-0.2011 (± 0.2060)	-0.2011 (± 0.2060)	-0.2061 (± 0.2113)	-0.2011 (± 0.2060)	-0.2332 (± 0.2375)	-0.0118 (± 0.0535)	-0.2158 (± 0.2208)	-0.0103 (± 0.0500)
$W \times W$	-0.0543 (± 0.1069)	-0.2165 (± 0.2503)	-0.0594 (± 0.1125)	-0.2773 (± 0.3146)	-0.0543 (± 0.1069)	-0.0543 (± 0.1069)	+0.0207 (± 0.0702)	-0.0543 (± 0.1069)	-0.2100 (± 0.2494)	-0.2763 (± 0.3152)	-0.0474 (± 0.1015)	+0.0216 (± 0.0704)
$W \times A$	-0.0749 (± 0.1089)	-0.0771 (± 0.1114)	-0.0747 (± 0.1091)	-0.1785 (± 0.2367)	-0.0749 (± 0.1089)	-0.0749 (± 0.1089)	+0.0177 (± 0.0534)	-0.0749 (± 0.1089)	-0.1774 (± 0.2367)	-0.1692 (± 0.2360)	+0.0172 (± 0.0532)	+0.0257 (± 0.0564)
$W \times B$	-0.0778 (± 0.1120)	-0.0778 (± 0.1120)	-0.0774 (± 0.1121)	-0.1837 (± 0.2444)	-0.0778 (± 0.1120)	-0.0778 (± 0.1120)	+0.0180 (± 0.0537)	-0.0778 (± 0.1120)	-0.1787 (± 0.2369)	-0.1748 (± 0.2462)	+0.0180 (± 0.0537)	+0.0268 (± 0.0568)
$A \times A$	+0.0056 (± 0.0298)	+0.0056 (± 0.0298)	+0.0058 (± 0.0300)	+0.0056 (± 0.0298)	+0.0056 (± 0.0298)	+0.0056 (± 0.0298)	+0.0056 (± 0.0298)	+0.0056 (± 0.0298)	+0.0056 (± 0.0298)	+0.0070 (± 0.0301)	+0.0056 (± 0.0298)	+0.0070 (± 0.0301)
$A \times B$	+0.0064 (± 0.0293)	+0.0064 (± 0.0293)	+0.0063 (± 0.0292)	+0.0063 (± 0.0293)	+0.0064 (± 0.0293)	+0.0064 (± 0.0293)	+0.0064 (± 0.0293)	+0.0064 (± 0.0293)	+0.0064 (± 0.0293)	+0.0075 (± 0.0295)	+0.0064 (± 0.0293)	+0.0075 (± 0.0294)
$B \times B$	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)	+0.0063 (± 0.0275)

Tabelle B.2: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gauß'sche Copula mit $\tau = 0.50$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\vartheta}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten ($W = \text{Worst}$, $A = \text{Average}$, $B = \text{Best}$). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	Gauß'sche Copula mit Kendall'schem Tau $\tau = 0.75$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	-0.1069 (± 0.1753)	-0.1211 (± 0.1980)	-0.0604 (± 0.1258)	-0.1530 (± 0.2537)	-0.1070 (± 0.1753)	-0.1071 (± 0.1753)	-0.0587 (± 0.1323)	-0.1069 (± 0.1753)	-0.1422 (± 0.2352)	-0.1070 (± 0.2180)	-0.0645 (± 0.1360)	-0.0178 (± 0.0609)
$r = 10^{-6}$	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)	-0.0002 (± 0.0157)
$r = 10^{-4}$	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)	-0.0006 (± 0.0156)
$r = 10^{-2}$	-0.0007 (± 0.0159)	-0.0039 (± 0.0164)	-0.0005 (± 0.0171)	-0.0027 (± 0.0160)	-0.0031 (± 0.0159)	-0.0034 (± 0.0160)	-0.0004 (± 0.0159)	-0.0007 (± 0.0159)	-0.0111 (± 0.0190)	-0.0032 (± 0.0189)	-0.0092 (± 0.0190)	-0.0006 (± 0.0190)
$p = 0.025$	-0.1651 (± 0.1713)	-0.1651 (± 0.1713)	-0.1035 (± 0.1153)	-0.1724 (± 0.1789)	-0.1651 (± 0.1713)	-0.1651 (± 0.1630)	-0.1568 (± 0.1713)	-0.1651 (± 0.1713)	-0.1726 (± 0.1791)	-0.1149 (± 0.1264)	-0.1569 (± 0.1631)	-0.1082 (± 0.1184)
$p = 0.050$	-0.2458 (± 0.2499)	-0.2460 (± 0.2502)	-0.0885 (± 0.1254)	-0.2618 (± 0.2661)	-0.2458 (± 0.2499)	-0.2458 (± 0.2394)	-0.2348 (± 0.2499)	-0.2458 (± 0.2499)	-0.2625 (± 0.2667)	-0.1110 (± 0.1427)	-0.2359 (± 0.2404)	-0.0989 (± 0.1274)
$p = 0.100$	-0.3474 (± 0.3503)	-0.3506 (± 0.3536)	-0.0100 (± 0.0791)	-0.3761 (± 0.3790)	-0.3474 (± 0.3503)	-0.3474 (± 0.3503)	-0.3412 (± 0.3447)	-0.3474 (± 0.3503)	-0.3833 (± 0.3862)	-0.0463 (± 0.0916)	-0.3502 (± 0.3537)	-0.0433 (± 0.0823)
W×W	-0.1840 (± 0.2243)	-0.3466 (± 0.3852)	-0.1841 (± 0.2248)	-0.4524 (± 0.4952)	-0.1840 (± 0.2243)	-0.1840 (± 0.2243)	+0.0128 (± 0.0519)	-0.1840 (± 0.2243)	-0.3091 (± 0.3648)	-0.4514 (± 0.4952)	-0.0380 (± 0.0874)	+0.0129 (± 0.0521)
W×A	-0.1722 (± 0.2323)	-0.1742 (± 0.2347)	-0.1713 (± 0.2323)	-0.2899 (± 0.3827)	-0.1722 (± 0.2323)	-0.1722 (± 0.2323)	+0.0111 (± 0.0411)	-0.1722 (± 0.2323)	-0.2869 (± 0.3789)	-0.2841 (± 0.3823)	+0.0107 (± 0.0411)	+0.0147 (± 0.0418)
W×B	-0.1641 (± 0.2203)	-0.1641 (± 0.2203)	-0.1637 (± 0.2203)	-0.2774 (± 0.3641)	-0.1641 (± 0.2203)	-0.1641 (± 0.2203)	+0.0078 (± 0.0382)	-0.1641 (± 0.2203)	-0.2776 (± 0.3648)	-0.2720 (± 0.3638)	+0.0078 (± 0.0382)	+0.0109 (± 0.0384)
A×A	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	-0.0006 (± 0.0157)	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	-0.0007 (± 0.0156)	+0.0002 (± 0.0156)
A×B	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	-0.0008 (± 0.0166)	+0.0001 (± 0.0165)	-0.0008 (± 0.0165)	±0.0000 (± 0.0165)
B×B	-0.0006 (± 0.0154)	-0.0006 (± 0.0154)	-0.0006 (± 0.0154)	-0.0006 (± 0.0154)	-0.0006 (± 0.0154)	-0.0006 (± 0.0154)	-0.0005 (± 0.0154)	-0.0006 (± 0.0154)	-0.0006 (± 0.0154)	-0.0006 (± 0.0154)	-0.0005 (± 0.0154)	-0.0005 (± 0.0154)

Tabelle B.3: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gauß'sche Copula mit $\tau = 0.75$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$		<i>t-Copula, $\nu = 4$, mit Kendall'schem Tau $\tau = 0.25$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.</i>										
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	-0.0086 (± 0.1067)	-0.0301 (± 0.0747)	+0.0099 (± 0.1081)	-0.0400 (± 0.0905)	-0.0105 (± 0.1104)	-0.0077 (± 0.1042)	-0.0028 (± 0.0548)	-0.0086 (± 0.1067)	-0.0382 (± 0.0870)	-0.0267 (± 0.0878)	-0.0004 (± 0.0612)	+0.0046 (± 0.0583)
$r = 10^{-6}$	+0.0055 (± 0.0402)	+0.0055 (± 0.0402)	+0.0055 (± 0.0402)	+0.0055 (± 0.0402)	+0.0055 (± 0.0402)	+0.0055 (± 0.0402)	+0.0033 (± 0.0410)	+0.0055 (± 0.0402)	+0.0055 (± 0.0402)	+0.0055 (± 0.0402)	+0.0033 (± 0.0410)	+0.0033 (± 0.0410)
$r = 10^{-4}$	+0.0063 (± 0.0439)	+0.0063 (± 0.0439)	+0.0064 (± 0.0439)	+0.0063 (± 0.0439)	+0.0063 (± 0.0439)	+0.0063 (± 0.0445)	+0.0037 (± 0.0445)	+0.0063 (± 0.0439)	+0.0063 (± 0.0439)	+0.0063 (± 0.0438)	+0.0037 (± 0.0445)	+0.0040 (± 0.0445)
$r = 10^{-2}$	+0.0053 (± 0.0423)	+0.0006 (± 0.0418)	+0.0058 (± 0.0485)	+0.0039 (± 0.0420)	+0.0035 (± 0.0420)	+0.0034 (± 0.0420)	- (-)	+0.0053 (± 0.0423)	-0.0017 (± 0.0415)	+0.0044 (± 0.0491)	- (-)	- (-)
$p = 0.025$	-0.0212 (± 0.0473)	-0.0213 (± 0.0473)	-0.0023 (± 0.0451)	-0.0223 (± 0.0477)	-0.0212 (± 0.0473)	-0.0212 (± 0.0473)	-0.0224 (± 0.0472)	-0.0212 (± 0.0473)	-0.0224 (± 0.0477)	-0.0111 (± 0.0447)	-0.0224 (± 0.0472)	-0.0118 (± 0.0442)
$p = 0.050$	-0.0427 (± 0.0620)	-0.0434 (± 0.0624)	+0.0157 (± 0.0603)	-0.0457 (± 0.0637)	-0.0427 (± 0.0620)	-0.0427 (± 0.0620)	-0.0474 (± 0.0667)	-0.0427 (± 0.0620)	-0.0466 (± 0.0643)	-0.0092 (± 0.0512)	-0.0474 (± 0.0667)	-0.0114 (± 0.0505)
$p = 0.100$	-0.0742 (± 0.0864)	-0.0821 (± 0.0931)	+0.0765 (± 0.0979)	-0.0817 (± 0.0921)	-0.0742 (± 0.0864)	-0.0742 (± 0.0864)	-0.1003 (± 0.1076)	-0.0742 (± 0.0864)	-0.0900 (± 0.0993)	+0.0122 (± 0.0541)	-0.1003 (± 0.1076)	+0.0064 (± 0.0519)
W×W	+0.1072 (± 0.3080)	-0.1346 (± 0.1537)	+0.1001 (± 0.3099)	-0.1566 (± 0.1787)	+0.0865 (± 0.3232)	+0.1194 (± 0.2975)	- (-)	+0.1072 (± 0.3080)	-0.1169 (± 0.1506)	-0.1567 (± 0.1802)	+0.0951 (± 0.1496)	+0.0929 (± 0.1453)
W×A	-0.0542 (± 0.0938)	-0.0566 (± 0.0956)	-0.0540 (± 0.0938)	-0.1041 (± 0.1433)	-0.0542 (± 0.0938)	-0.0542 (± 0.0768)	+0.0107 (± 0.0938)	-0.0542 (± 0.0938)	-0.1067 (± 0.1446)	-0.0965 (± 0.1436)	+0.0192 (± 0.1029)	+0.0243 (± 0.1035)
W×B	-0.0515 (± 0.0907)	-0.0515 (± 0.0907)	-0.0514 (± 0.0908)	-0.1018 (± 0.1406)	-0.0515 (± 0.0907)	-0.0515 (± 0.0907)	+0.0105 (± 0.0832)	-0.0515 (± 0.0907)	-0.1021 (± 0.1405)	-0.0936 (± 0.1410)	+0.0105 (± 0.0832)	+0.0176 (± 0.0848)
A×A	+0.0052 (± 0.0447)	+0.0052 (± 0.0447)	+0.0052 (± 0.0447)	+0.0052 (± 0.0447)	+0.0052 (± 0.0447)	+0.0052 (± 0.0447)	+0.0029 (± 0.0430)	+0.0052 (± 0.0447)	+0.0052 (± 0.0447)	+0.0063 (± 0.0450)	+0.0029 (± 0.0441)	+0.0033 (± 0.0441)
A×B	+0.0051 (± 0.0451)	+0.0051 (± 0.0451)	+0.0050 (± 0.0452)	+0.0051 (± 0.0451)	+0.0051 (± 0.0451)	+0.0051 (± 0.0451)	+0.0029 (± 0.0457)	+0.0051 (± 0.0451)	+0.0051 (± 0.0451)	+0.0059 (± 0.0455)	+0.0029 (± 0.0457)	+0.0038 (± 0.0459)
B×B	+0.0060 (± 0.0433)	+0.0060 (± 0.0433)	+0.0060 (± 0.0433)	+0.0060 (± 0.0433)	+0.0060 (± 0.0433)	+0.0060 (± 0.0433)	+0.0041 (± 0.0433)	+0.0060 (± 0.0433)	+0.0060 (± 0.0433)	+0.0060 (± 0.0433)	+0.0041 (± 0.0435)	+0.0041 (± 0.0435)

Tabelle B.4: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die t-Copula mit $\tau = 0.25$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$		<i>t-Copula, $\nu = 4$, mit Kendall'schem Tau $\tau = 0.50$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.</i>											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring	
		-0.0375 (± 0.1108)	-0.0612 (± 0.1227)	-0.0133 (± 0.1038)	-0.0795 (± 0.1583)	-0.0380 (± 0.1122)	-0.0377 (± 0.1108)	+0.0097 (± 0.0544)	-0.0375 (± 0.1108)	-0.0730 (± 0.1444)	-0.0585 (± 0.1500)	-0.0061 (± 0.0601)	+0.0003 (± 0.0538)
<i>Modifiziert</i> [†]		+0.0038 (± 0.0339)	+0.0038 (± 0.0339)	+0.0038 (± 0.0339)	+0.0038 (± 0.0339)	+0.0038 (± 0.0339)	+0.0038 (± 0.0355)	+0.0016 (± 0.0339)	+0.0038 (± 0.0339)	+0.0038 (± 0.0339)	+0.0038 (± 0.0339)	+0.0016 (± 0.0355)	+0.0016 (± 0.0355)
$r = 10^{-6}$		+0.0002 (± 0.0339)	+0.0002 (± 0.0339)	+0.0002 (± 0.0339)	+0.0002 (± 0.0339)	+0.0002 (± 0.0339)	-0.0015 (± 0.0353)	+0.0002 (± 0.0339)	-0.0015 (± 0.0353)	-0.0013 (± 0.0348)			
$r = 10^{-4}$		+0.0023 (± 0.0341)	-0.0036 (± 0.0341)	+0.0019 (± 0.0371)	-0.0001 (± 0.0339)	-0.0005 (± 0.0340)	-0.0007 (± 0.0341)	- (-)	+0.0023 (± 0.0341)	-0.0102 (± 0.0346)	+0.0009 (± 0.0377)	- (-)	- (-)
$r = 10^{-2}$		-0.0338 (± 0.0500)	-0.0338 (± 0.0500)	-0.0077 (± 0.0404)	-0.0349 (± 0.0508)	-0.0338 (± 0.0500)	-0.0338 (± 0.0500)	-0.0438 (± 0.0567)	-0.0338 (± 0.0500)	-0.0351 (± 0.0510)	-0.0186 (± 0.0430)	-0.0438 (± 0.0567)	-0.0259 (± 0.0468)
$p = 0.025$		-0.0721 (± 0.0815)	-0.0725 (± 0.0818)	+0.0024 (± 0.0505)	-0.0760 (± 0.0853)	-0.0721 (± 0.0815)	-0.0721 (± 0.0815)	-0.0821 (± 0.0907)	-0.0721 (± 0.0815)	-0.0767 (± 0.0857)	-0.0252 (± 0.0534)	-0.0821 (± 0.0907)	-0.0318 (± 0.0563)
$p = 0.050$		-0.1437 (± 0.1505)	-0.1495 (± 0.1560)	+0.0540 (± 0.0811)	-0.1555 (± 0.1617)	-0.1437 (± 0.1505)	-0.1437 (± 0.1505)	-0.1549 (± 0.1628)	-0.1437 (± 0.1505)	-0.1629 (± 0.1687)	-0.0034 (± 0.0575)	-0.1549 (± 0.1628)	-0.0088 (± 0.0534)
$p = 0.100$		+0.0179 (± 0.2259)	-0.2513 (± 0.2857)	+0.0088 (± 0.2321)	-0.2985 (± 0.3390)	+0.0140 (± 0.2345)	+0.0179 (± 0.2261)	- (-)	+0.0179 (± 0.2259)	-0.2046 (± 0.2516)	-0.2910 (± 0.3341)	+0.0588 (± 0.1331)	+0.0592 (± 0.1314)
$W \times W$		-0.1140 (± 0.1670)	-0.1172 (± 0.1706)	-0.1132 (± 0.1671)	-0.1972 (± 0.2630)	-0.1140 (± 0.1670)	-0.1140 (± 0.1670)	+0.0116 (± 0.0604)	-0.1140 (± 0.1670)	-0.1942 (± 0.2599)	-0.1874 (± 0.2648)	+0.0235 (± 0.0866)	+0.0301 (± 0.0878)
$W \times A$		-0.1217 (± 0.1727)	-0.1217 (± 0.1727)	-0.1207 (± 0.1728)	-0.2072 (± 0.2720)	-0.1217 (± 0.1727)	-0.1217 (± 0.1727)	+0.0182 (± 0.0749)	-0.1217 (± 0.1727)	-0.2076 (± 0.2723)	-0.1952 (± 0.2706)	+0.0182 (± 0.0749)	+0.0286 (± 0.0773)
$W \times B$		+0.0024 (± 0.0316)	+0.0024 (± 0.0316)	+0.0023 (± 0.0318)	+0.0024 (± 0.0316)	+0.0024 (± 0.0316)	+0.0024 (± 0.0316)	+0.0003 (± 0.0326)	+0.0024 (± 0.0316)	+0.0024 (± 0.0316)	+0.0003 (± 0.0319)	+0.0015 (± 0.0334)	
$A \times A$		+0.0032 (± 0.0336)	+0.0032 (± 0.0336)	+0.0033 (± 0.0337)	+0.0032 (± 0.0336)	+0.0032 (± 0.0336)	+0.0032 (± 0.0336)	+0.0014 (± 0.0351)	+0.0032 (± 0.0336)	+0.0032 (± 0.0336)	+0.0045 (± 0.0338)	+0.0014 (± 0.0351)	+0.0028 (± 0.0351)
$A \times B$		+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0037 (± 0.0346)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0037 (± 0.0346)	+0.0037 (± 0.0346)
$B \times B$		+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0037 (± 0.0346)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0058 (± 0.0331)	+0.0037 (± 0.0346)	+0.0037 (± 0.0346)

Tabelle B.5: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die t-Copula mit $\tau = 0.50$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten ($W = \text{Worst}$, $A = \text{Average}$, $B = \text{Best}$). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	<i>t-Copula, $\nu = 4$, mit Kendall'schem Tau $\tau = 0.75$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.</i>											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	-0.0712 (± 0.1538)	-0.0941 (± 0.1835)	-0.0469 (± 0.1426)	-0.1191 (± 0.2359)	-0.0715 (± 0.1538)	-0.0715 (± 0.1538)	-0.0149 (± 0.0454)	-0.0712 (± 0.1538)	-0.1074 (± 0.2129)	-0.0979 (± 0.2291)	-0.0109 (± 0.0472)	-0.0077 (± 0.0414)
$r = 10^{-6}$	-0.0007 (± 0.0187)	-0.0007 (± 0.0187)	-0.0007 (± 0.0187)	-0.0007 (± 0.0187)	-0.0007 (± 0.0187)	-0.0007 (± 0.0187)	-0.0020 (± 0.0198)	-0.0007 (± 0.0187)	-0.0007 (± 0.0187)	-0.0007 (± 0.0187)	-0.0020 (± 0.0198)	-0.0020 (± 0.0198)
$r = 10^{-4}$	-0.0011 (± 0.0184)	-0.0011 (± 0.0184)	-0.0011 (± 0.0184)	-0.0012 (± 0.0184)	-0.0011 (± 0.0184)	-0.0011 (± 0.0184)	-0.0011 (± 0.0192)	-0.0011 (± 0.0184)	-0.0012 (± 0.0184)	-0.0011 (± 0.0184)	-0.0011 (± 0.0192)	-0.0013 (± 0.0200)
$r = 10^{-2}$	-0.0004 (± 0.0174)	-0.0059 (± 0.0183)	-0.0006 (± 0.0198)	-0.0039 (± 0.0179)	-0.0038 (± 0.0180)	-0.0038 (± 0.0179)	-	-0.0004 (± 0.0174)	-0.0164 (± 0.0234)	-0.0034 (± 0.0217)	-	-
$p = 0.025$	-0.0356 (± 0.0429)	-0.0357 (± 0.0429)	-0.0163 (± 0.0285)	-0.0361 (± 0.0434)	-0.0356 (± 0.0429)	-0.0356 (± 0.0429)	-0.0550 (± 0.0607)	-0.0356 (± 0.0429)	-0.0360 (± 0.0433)	-0.0248 (± 0.0352)	-0.0550 (± 0.0607)	-0.0404 (± 0.0487)
$p = 0.050$	-0.0790 (± 0.0849)	-0.0791 (± 0.0850)	-0.0118 (± 0.0376)	-0.0813 (± 0.0873)	-0.0790 (± 0.0849)	-0.0790 (± 0.0849)	-0.0934 (± 0.0982)	-0.0790 (± 0.0849)	-0.0815 (± 0.0876)	-0.0348 (± 0.0542)	-0.0934 (± 0.0982)	-0.0477 (± 0.0654)
$p = 0.100$	-0.1855 (± 0.1911)	-0.1893 (± 0.1949)	+0.0202 (± 0.0460)	-0.1979 (± 0.2038)	-0.1855 (± 0.1911)	-0.1855 (± 0.1911)	-0.1805 (± 0.1849)	-0.1855 (± 0.1911)	-0.2021 (± 0.2079)	-0.0140 (± 0.0518)	-0.1805 (± 0.1849)	-0.0199 (± 0.0531)
W×W	-0.1195 (± 0.2616)	-0.3812 (± 0.4302)	-0.1219 (± 0.2639)	-0.4723 (± 0.5223)	-0.1194 (± 0.2617)	-0.1194 (± 0.2617)	-	-0.1195 (± 0.2616)	-0.3146 (± 0.3832)	-0.4732 (± 0.5273)	+0.0140 (± 0.0610)	+0.0141 (± 0.0608)
W×A	-0.2229 (± 0.2939)	-0.2262 (± 0.2978)	-0.2214 (± 0.2939)	-0.3282 (± 0.4206)	-0.2229 (± 0.2939)	-0.2229 (± 0.2939)	+0.0050 (± 0.0285)	-0.2229 (± 0.2939)	-0.3259 (± 0.4184)	-0.3208 (± 0.4188)	+0.0131 (± 0.0494)	+0.0150 (± 0.0491)
W×B	-0.2103 (± 0.2854)	-0.2103 (± 0.2854)	-0.2097 (± 0.2854)	-0.3081 (± 0.4060)	-0.2103 (± 0.2854)	-0.2103 (± 0.2854)	+0.0066 (± 0.0486)	-0.2103 (± 0.2854)	-0.3101 (± 0.4084)	-0.3033 (± 0.4090)	+0.0066 (± 0.0486)	+0.0120 (± 0.0490)
A×A	-0.0003 (± 0.0182)	-0.0003 (± 0.0182)	-0.0002 (± 0.0182)	-0.0003 (± 0.0182)	-0.0003 (± 0.0182)	-0.0003 (± 0.0182)	-0.0010 (± 0.0198)	-0.0003 (± 0.0182)	-0.0003 (± 0.0182)	+0.0008 (± 0.0180)	-0.0010 (± 0.0198)	+0.0002 (± 0.0199)
A×B	+0.0004 (± 0.0188)	+0.0004 (± 0.0188)	+0.0004 (± 0.0188)	+0.0004 (± 0.0188)	+0.0004 (± 0.0188)	+0.0004 (± 0.0188)	-0.0010 (± 0.0191)	+0.0004 (± 0.0188)	+0.0004 (± 0.0188)	+0.0011 (± 0.0188)	-0.0010 (± 0.0191)	+0.0000 (± 0.0190)
B×B	-0.0001 (± 0.0178)	-0.0001 (± 0.0178)	-0.0001 (± 0.0178)	-0.0001 (± 0.0178)	-0.0001 (± 0.0178)	-0.0001 (± 0.0178)	-0.0009 (± 0.0196)	-0.0001 (± 0.0178)	-0.0001 (± 0.0178)	-0.0001 (± 0.0178)	-0.0009 (± 0.0196)	-0.0009 (± 0.0196)

Tabelle B.6: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die t-Copula mit $\tau = 0.75$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	Clayton-Copula mit Kendall'schem Tau $\tau = 0.25$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	+0.0051 (± 0.0788)	+0.0008 (± 0.0721)	+0.0189 (± 0.0898)	-0.0018 (± 0.0865)	+0.0174 (± 0.0865)	+0.0176 (± 0.0860)	+0.0604 (± 0.1339)	+0.0176 (± 0.0860)	-0.0004 (± 0.0804)	-0.0122 (± 0.0734)	+0.0512 (± 0.1196)	+0.0620 (± 0.1360)
$r = 10^{-6}$	+0.0134 (± 0.0394)	+0.0134 (± 0.0394)	+0.0134 (± 0.0394)	+0.0134 (± 0.0394)	+0.0134 (± 0.0394)	+0.0134 (± 0.0394)	- (-)	+0.0134 (± 0.0394)	+0.0134 (± 0.0394)	+0.0134 (± 0.0394)	- (-)	- (-)
$r = 10^{-4}$	+0.0113 (± 0.0412)	+0.0113 (± 0.0412)	+0.0114 (± 0.0411)	+0.0114 (± 0.0411)	+0.0114 (± 0.0410)	+0.0115 (± 0.0410)	- (-)	+0.0114 (± 0.0409)	+0.0114 (± 0.0410)	+0.0115 (± 0.0410)	- (-)	- (-)
$r = 10^{-2}$	+0.0137 (± 0.0409)	+0.0105 (± 0.0396)	+0.0127 (± 0.0436)	+0.0103 (± 0.0398)	+0.0119 (± 0.0402)	+0.0138 (± 0.0407)	+0.0015 (± 0.0392)	+0.0131 (± 0.0398)	+0.0062 (± 0.0364)	+0.0125 (± 0.0435)	-0.0041 (± 0.0379)	-0.0013 (± 0.0417)
$p = 0.025$	-0.0143 (± 0.0418)	-0.0144 (± 0.0419)	-0.0009 (± 0.0439)	-0.0075 (± 0.0434)	-0.0091 (± 0.0430)	-0.0092 (± 0.0430)	- (-)	-0.0093 (± 0.0431)	-0.0102 (± 0.0437)	-0.0102 (± 0.0394)	- (-)	- (-)
$p = 0.050$	-0.0021 (± 0.0501)	-0.0024 (± 0.0504)	+0.0337 (± 0.0781)	+0.0291 (± 0.0758)	+0.0240 (± 0.0753)	+0.0232 (± 0.0753)	- (-)	+0.0230 (± 0.0753)	+0.0224 (± 0.0755)	+0.0019 (± 0.0495)	- (-)	- (-)
$p = 0.100$	+0.0279 (± 0.0727)	+0.0274 (± 0.0733)	+0.1385 (± 0.1556)	+0.1379 (± 0.1552)	+0.1222 (± 0.1539)	+0.1207 (± 0.1533)	- (-)	+0.1207 (± 0.1533)	+0.1205 (± 0.1534)	+0.0342 (± 0.0716)	- (-)	- (-)
W×W	+0.0421 (± 0.2087)	-0.0053 (± 0.1763)	+0.0423 (± 0.2085)	-0.1143 (± 0.1539)	+0.0550 (± 0.1991)	+0.0558 (± 0.1981)	+0.1014 (± 0.1866)	+0.0559 (± 0.1983)	-0.0698 (± 0.1172)	-0.1202 (± 0.1452)	+0.0869 (± 0.1635)	+0.1023 (± 0.1850)
W×A	-0.0341 (± 0.0782)	-0.0350 (± 0.0787)	-0.0307 (± 0.0759)	-0.0698 (± 0.1092)	-0.0281 (± 0.0657)	-0.0271 (± 0.0637)	+0.1702 (± 0.1744)	-0.0271 (± 0.0637)	-0.0660 (± 0.1037)	-0.0617 (± 0.1071)	+0.1710 (± 0.1749)	+0.1713 (± 0.1751)
W×B	-0.0397 (± 0.0821)	-0.0397 (± 0.0821)	-0.0381 (± 0.0808)	-0.0774 (± 0.1145)	-0.0350 (± 0.0720)	-0.0339 (± 0.0699)	+0.1886 (± 0.2088)	-0.0334 (± 0.0694)	-0.0767 (± 0.1115)	-0.0724 (± 0.1146)	+0.1886 (± 0.2088)	+0.1886 (± 0.2088)
A×A	+0.0157 (± 0.0396)	+0.0157 (± 0.0396)	+0.0157 (± 0.0397)	+0.0157 (± 0.0397)	+0.0157 (± 0.0397)	+0.0157 (± 0.0397)	- (-)	+0.0157 (± 0.0397)	+0.0157 (± 0.0397)	+0.0164 (± 0.0400)	- (-)	- (-)
A×B	+0.0139 (± 0.0412)	+0.0139 (± 0.0412)	+0.0139 (± 0.0412)	+0.0139 (± 0.0412)	+0.0139 (± 0.0413)	+0.0139 (± 0.0413)	- (-)	+0.0139 (± 0.0412)	+0.0139 (± 0.0413)	+0.0145 (± 0.0415)	- (-)	- (-)
B×B	+0.014 (± 0.0361)	+0.014 (± 0.0361)	- (-)	+0.014 (± 0.0361)	+0.014 (± 0.0361)	+0.014 (± 0.0361)	- (-)	- (-)				

Tabelle B.7: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Clayton-Copula mit $\tau = 0.25$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	Clayton-Copula mit Kendall'schem Tau $\tau = 0.50$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	-0.0630 (± 0.1258)	-0.0740 (± 0.1314)	-0.0196 (± 0.0966)	-0.0991 (± 0.1716)	-0.0636 (± 0.1278)	-0.0636 (± 0.1280)	+0.0243 (± 0.0513)	-0.0621 (± 0.1259)	-0.0954 (± 0.1634)	-0.0649 (± 0.1431)	+0.0200 (± 0.0465)	+0.0249 (± 0.0498)
$r = 10^{-6}$	+0.0007 (± 0.0309)	+0.0007 (± 0.0309)	+0.0007 (± 0.0309)	+0.0007 (± 0.0309)	+0.0007 (± 0.0309)	+0.0007 (± 0.0309)	+0.0191 (± 0.0272)	+0.0007 (± 0.0309)	+0.0007 (± 0.0309)	+0.0007 (± 0.0309)	+0.0191 (± 0.0272)	+0.0191 (± 0.0272)
$r = 10^{-4}$	+0.0011 (± 0.0320)	+0.0011 (± 0.0320)	+0.0011 (± 0.0320)	+0.0011 (± 0.0320)	+0.0012 (± 0.0319)	+0.0012 (± 0.0319)	+0.0182 (± 0.0277)	+0.0011 (± 0.0320)	+0.0011 (± 0.0319)	+0.0011 (± 0.0320)	+0.0182 (± 0.0277)	+0.0181 (± 0.0277)
$r = 10^{-2}$	+0.0004 (± 0.0309)	-0.0044 (± 0.0307)	+0.0001 (± 0.0333)	-0.0034 (± 0.0307)	-0.0003 (± 0.0311)	+0.0017 (± 0.0313)	+0.0148 (± 0.0279)	+0.0004 (± 0.0309)	-0.0171 (± 0.0332)	-0.0034 (± 0.0346)	+0.0040 (± 0.0227)	+0.0175 (± 0.0336)
$p = 0.025$	-0.1084 (± 0.1148)	-0.1084 (± 0.1148)	-0.0511 (± 0.0684)	-0.1178 (± 0.1245)	-0.1084 (± 0.1148)	-0.1084 (± 0.1148)	+0.0032 (± 0.0032)	-0.1084 (± 0.1148)	-0.1178 (± 0.1245)	-0.0754 (± 0.0884)	+0.0032 (± 0.0032)	+0.0132 (± 0.0231)
$p = 0.050$	-0.1762 (± 0.1802)	-0.1769 (± 0.1809)	-0.0178 (± 0.0874)	-0.2006 (± 0.2049)	-0.1762 (± 0.1802)	-0.1762 (± 0.1802)	- (-)	-0.1762 (± 0.1802)	-0.2024 (± 0.2067)	-0.0738 (± 0.1109)	- (-)	+0.0175 (± 0.0266)
$p = 0.100$	-0.2417 (± 0.2461)	-0.2486 (± 0.2528)	+0.0779 (± 0.1086)	-0.2370 (± 0.2773)	-0.2130 (± 0.2371)	-0.2154 (± 0.2381)	- (-)	-0.2169 (± 0.2388)	-0.2695 (± 0.2933)	-0.0228 (± 0.0754)	- (-)	+0.0148 (± 0.0254)
W×W	-0.0367 (± 0.1807)	-0.1566 (± 0.2150)	-0.0429 (± 0.1852)	-0.2763 (± 0.3087)	-0.0435 (± 0.1849)	-0.0435 (± 0.1849)	+0.0404 (± 0.0875)	-0.0384 (± 0.1818)	-0.1844 (± 0.2281)	-0.2695 (± 0.3053)	+0.0125 (± 0.0651)	+0.0411 (± 0.0876)
W×A	-0.1037 (± 0.1531)	-0.1054 (± 0.1550)	-0.1030 (± 0.1532)	-0.1742 (± 0.2368)	-0.1096 (± 0.1598)	-0.1096 (± 0.1598)	+0.0352 (± 0.0718)	-0.1039 (± 0.1535)	-0.1733 (± 0.2360)	-0.1646 (± 0.2399)	+0.0349 (± 0.0714)	+0.0416 (± 0.0735)
W×B	-0.1140 (± 0.1600)	-0.1140 (± 0.1600)	-0.1141 (± 0.1602)	-0.1903 (± 0.2490)	-0.1221 (± 0.1687)	-0.1221 (± 0.1687)	+0.0262 (± 0.0645)	-0.1149 (± 0.1615)	-0.1918 (± 0.2509)	-0.1827 (± 0.2517)	+0.0262 (± 0.0645)	+0.0313 (± 0.0657)
A×A	+0.0033 (± 0.0303)	+0.0033 (± 0.0303)	+0.0033 (± 0.0304)	+0.0033 (± 0.0303)	+0.0033 (± 0.0303)	+0.0033 (± 0.0303)	+0.0186 (± 0.0279)	+0.0033 (± 0.0303)	+0.0033 (± 0.0303)	+0.0046 (± 0.0305)	+0.0186 (± 0.0279)	+0.0190 (± 0.0285)
A×B	+0.0019 (± 0.0319)	+0.0019 (± 0.0319)	+0.0019 (± 0.0319)	+0.0019 (± 0.0319)	+0.0019 (± 0.0319)	+0.0019 (± 0.0319)	+0.0200 (± 0.0289)	+0.0019 (± 0.0319)	+0.0019 (± 0.0319)	+0.0029 (± 0.0319)	+0.0200 (± 0.0289)	+0.0203 (± 0.0293)
B×B	+0.0027 (± 0.0324)	+0.0027 (± 0.0324)	+0.0027 (± 0.0324)	+0.0027 (± 0.0324)	+0.0027 (± 0.0324)	+0.0027 (± 0.0324)	+0.0201 (± 0.0286)	+0.0027 (± 0.0324)	+0.0027 (± 0.0324)	+0.0027 (± 0.0324)	+0.0201 (± 0.0286)	+0.0201 (± 0.0286)

Tabelle B.8: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Clayton-Copula mit $\tau = 0.50$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$		<i>Clayton-Copula mit Kendall'schem Tau</i> $\tau = 0.75$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring	
Modifiziert†	-0.1283 (± 0.2131)	-0.1429 (± 0.2286)	-0.0702 (± 0.1533)	-0.1790 (± 0.2845)	-0.1344 (± 0.2206)	-0.1342 (± 0.2206)	-0.0321 (± 0.0856)	-0.1283 (± 0.2131)	-0.1705 (± 0.2704)	-0.1244 (± 0.2365)	-0.0408 (± 0.0892)	-0.0239 (± 0.0700)	
$r = 10^{-6}$	-0.0057 (± 0.0204)	-0.0057 (± 0.0204)	-0.0057 (± 0.0204)	-0.0058 (± 0.0205)	-0.0057 (± 0.0204)	-0.0057 (± 0.0204)	-0.0067 (± 0.0210)	-0.0057 (± 0.0204)	-0.0057 (± 0.0204)	-0.0057 (± 0.0204)	-0.0067 (± 0.0210)	-0.0067 (± 0.0210)	
$r = 10^{-4}$	-0.0045 (± 0.0182)	-0.0045 (± 0.0182)	-0.0045 (± 0.0182)	-0.0046 (± 0.0182)	-0.0044 (± 0.0182)	-0.0044 (± 0.0182)	-0.0054 (± 0.0186)	-0.0045 (± 0.0182)	-0.0046 (± 0.0182)	-0.0046 (± 0.0182)	-0.0054 (± 0.0186)	-0.0053 (± 0.0186)	
$r = 10^{-2}$	-0.0055 (± 0.0197)	-0.0102 (± 0.0213)	-0.0062 (± 0.0213)	-0.0128 (± 0.0225)	-0.0106 (± 0.0217)	-0.0092 (± 0.0214)	-0.0043 (± 0.0205)	-0.0055 (± 0.0197)	-0.0286 (± 0.0337)	-0.0142 (± 0.0258)	-0.0238 (± 0.0298)	-0.0035 (± 0.0237)	
$p = 0.025$	-0.1793 (± 0.1849)	-0.1793 (± 0.1849)	-0.1042 (± 0.1147)	-0.1901 (± 0.1963)	-0.1793 (± 0.1849)	-0.1793 (± 0.1849)	-0.1615 (± 0.1664)	-0.1793 (± 0.1849)	-0.1900 (± 0.1962)	-0.1282 (± 0.1412)	-0.1616 (± 0.1666)	-0.1150 (± 0.1264)	
$p = 0.050$	-0.2952 (± 0.2988)	-0.2955 (± 0.2991)	-0.1086 (± 0.1450)	-0.3229 (± 0.3271)	-0.2952 (± 0.2988)	-0.2952 (± 0.2988)	-0.2294 (± 0.2310)	-0.2952 (± 0.2988)	-0.3238 (± 0.3278)	-0.1660 (± 0.2031)	-0.2307 (± 0.2322)	-0.1297 (± 0.1609)	
$p = 0.100$	-0.4242 (± 0.4263)	-0.4308 (± 0.4328)	+0.0062 (± 0.0914)	-0.4869 (± 0.4893)	-0.4242 (± 0.4263)	-0.4242 (± 0.4263)	-0.2086 (± 0.2148)	-0.4242 (± 0.4263)	-0.4968 (± 0.4990)	-0.0514 (± 0.1152)	-0.1837 (± 0.1837)	-0.0224 (± 0.0343)	
W×W	-0.1672 (± 0.2580)	-0.3317 (± 0.3784)	-0.1714 (± 0.2629)	-0.4558 (± 0.4969)	-0.1846 (± 0.2774)	-0.1845 (± 0.2771)	+0.0033 (± 0.0903)	-0.1672 (± 0.2580)	-0.3249 (± 0.3758)	-0.4595 (± 0.5019)	-0.0707 (± 0.1223)	+0.0019 (± 0.0893)	
W×A	-0.2357 (± 0.3044)	-0.2386 (± 0.3073)	-0.2357 (± 0.3054)	-0.3357 (± 0.4183)	-0.2520 (± 0.3203)	-0.2520 (± 0.3203)	-0.0086 (± 0.0586)	-0.2357 (± 0.3044)	-0.3336 (± 0.4153)	-0.3306 (± 0.4185)	-0.0093 (± 0.0587)	-0.0048 (± 0.0571)	
W×B	-0.2249 (± 0.2911)	-0.2249 (± 0.2911)	-0.2239 (± 0.2911)	-0.3260 (± 0.4064)	-0.2442 (± 0.3114)	-0.2442 (± 0.3114)	+0.0014 (± 0.0561)	-0.2249 (± 0.2911)	-0.3290 (± 0.4092)	-0.3240 (± 0.4101)	+0.0014 (± 0.0561)	+0.0047 (± 0.0541)	
A×A	-0.0046 (± 0.0190)	-0.0046 (± 0.0190)	-0.0045 (± 0.0191)	-0.0046 (± 0.0190)	-0.0046 (± 0.0190)	-0.0046 (± 0.0190)	-0.0055 (± 0.0194)	-0.0046 (± 0.0190)	-0.0046 (± 0.0190)	-0.0036 (± 0.0188)	-0.0055 (± 0.0194)	-0.0039 (± 0.0190)	
A×B	-0.0024 (± 0.0189)	-0.0024 (± 0.0189)	-0.0024 (± 0.0190)	-0.0024 (± 0.0189)	-0.0024 (± 0.0189)	-0.0024 (± 0.0189)	-0.0031 (± 0.0193)	-0.0024 (± 0.0189)	-0.0024 (± 0.0189)	-0.0015 (± 0.0187)	-0.0031 (± 0.0193)	-0.0019 (± 0.0190)	
B×B	-0.0052 (± 0.0186)	-0.0052 (± 0.0186)	-0.0052 (± 0.0186)	-0.0052 (± 0.0186)	-0.0052 (± 0.0186)	-0.0052 (± 0.0186)	-0.0061 (± 0.0190)	-0.0052 (± 0.0186)	-0.0052 (± 0.0186)	-0.0052 (± 0.0186)	-0.0061 (± 0.0190)	-0.0061 (± 0.0190)	

Tabelle B.9: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Clayton-Copula mit $\tau = 0.75$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	Gumbel-Copula mit Kendall'schem Tau $\tau = 0.25$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	-0.0037 (± 0.0768)	-0.0225 (± 0.0650)	+0.0088 (± 0.0768)	-0.0302 (± 0.0743)	-0.0041 (± 0.0768)	-0.0043 (± 0.0768)	+0.0550 (± 0.1541)	-0.0037 (± 0.0768)	-0.0283 (± 0.0716)	-0.0191 (± 0.0727)	+0.0113 (± 0.1011)	+0.0573 (± 0.1580)
$r = 10^{-6}$	+0.0046 (± 0.0408)	+0.0046 (± 0.0408)	+0.0046 (± 0.0408)	+0.0046 (± 0.0408)	+0.0046 (± 0.0408)	+0.0046 (± 0.0408)	- (-)	+0.0046 (± 0.0408)	+0.0046 (± 0.0408)	+0.0046 (± 0.0408)	- (-)	- (-)
$r = 10^{-4}$	+0.0073 (± 0.0422)	+0.0073 (± 0.0422)	+0.0074 (± 0.0422)	+0.0073 (± 0.0421)	+0.0072 (± 0.0421)	+0.0072 (± 0.0421)	- (-)	+0.0073 (± 0.0422)	+0.0073 (± 0.0422)	+0.0073 (± 0.0421)	- (-)	- (-)
$r = 10^{-2}$	+0.0062 (± 0.0400)	+0.0023 (± 0.0389)	+0.0062 (± 0.0444)	+0.0044 (± 0.0398)	+0.0018 (± 0.0400)	-0.0001 (± 0.0374)	+0.0018 (± 0.0400)	+0.0062 (± 0.0400)	-0.0011 (± 0.0380)	+0.0065 (± 0.0445)	-0.0016 (± 0.0362)	+0.0015 (± 0.0404)
$p = 0.025$	-0.0177 (± 0.0446)	-0.0178 (± 0.0446)	-0.0051 (± 0.0433)	-0.0184 (± 0.0448)	-0.0177 (± 0.0446)	-0.0177 (± 0.0446)	- (-)	-0.0177 (± 0.0446)	-0.0185 (± 0.0448)	-0.0089 (± 0.0429)	- (-)	- (-)
$p = 0.050$	-0.0347 (± 0.0540)	-0.0353 (± 0.0543)	+0.0034 (± 0.0499)	-0.0365 (± 0.0550)	-0.0347 (± 0.0540)	-0.0347 (± 0.0540)	- (-)	-0.0347 (± 0.0540)	-0.0374 (± 0.0556)	-0.0069 (± 0.0463)	- (-)	- (-)
$p = 0.100$	-0.0616 (± 0.0767)	-0.0671 (± 0.0809)	+0.0452 (± 0.0697)	-0.0654 (± 0.0795)	-0.0616 (± 0.0767)	-0.0616 (± 0.0767)	- (-)	-0.0616 (± 0.0767)	-0.0723 (± 0.0848)	+0.0109 (± 0.0482)	- (-)	- (-)
W×W	+0.0880 (± 0.2061)	-0.1250 (± 0.1464)	+0.0809 (± 0.2085)	-0.1251 (± 0.1447)	+0.0880 (± 0.2061)	+0.0880 (± 0.2849)	+0.1787 (± 0.2061)	+0.0880 (± 0.1209)	-0.0870 (± 0.1443)	-0.1234 (± 0.2378)	+0.0168 (± 0.2851)	+0.1789 (± 0.2851)
W×A	-0.0275 (± 0.0621)	-0.0301 (± 0.0643)	-0.0272 (± 0.0623)	-0.0787 (± 0.1123)	-0.0275 (± 0.0621)	-0.0275 (± 0.0621)	+0.2795 (± 0.2795)	-0.0275 (± 0.0621)	-0.0785 (± 0.1121)	-0.0727 (± 0.1164)	+0.2794 (± 0.2794)	+0.2795 (± 0.2795)
W×B	-0.0273 (± 0.0622)	-0.0273 (± 0.0622)	-0.0273 (± 0.0624)	-0.0731 (± 0.1072)	-0.0273 (± 0.0622)	-0.0273 (± 0.0622)	+0.2502 (± 0.2508)	-0.0273 (± 0.0622)	-0.0753 (± 0.1108)	-0.0662 (± 0.1115)	+0.2502 (± 0.2508)	+0.2502 (± 0.2508)
A×A	+0.0070 (± 0.0413)	+0.0070 (± 0.0413)	+0.0071 (± 0.0414)	+0.0070 (± 0.0413)	+0.0070 (± 0.0413)	+0.0070 (± 0.0413)	- (-)	+0.0070 (± 0.0413)	+0.0070 (± 0.0413)	+0.0081 (± 0.0417)	- (-)	- (-)
A×B	+0.0060 (± 0.0401)	+0.0060 (± 0.0401)	+0.0059 (± 0.0400)	+0.0060 (± 0.0401)	+0.0060 (± 0.0401)	+0.0060 (± 0.0401)	- (-)	+0.0060 (± 0.0401)	+0.0059 (± 0.0401)	+0.0068 (± 0.0404)	- (-)	- (-)
B×B	+0.0051 (± 0.0384)	+0.0051 (± 0.0384)	+0.0051 (± 0.0384)	+0.0051 (± 0.0384)	+0.0051 (± 0.0384)	+0.0051 (± 0.0384)	- (-)	+0.0051 (± 0.0384)	+0.0051 (± 0.0384)	+0.0051 (± 0.0384)	- (-)	- (-)

Tabelle B.10: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gumbel-Copula mit $\tau = 0.25$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	Gumbel-Copula mit Kendall'schem Tau $\tau = 0.50$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π -Jitter	mix $\hat{\tau}_n$ -Jitter	Censoring	Kojadinovic	V-Rd. Rnd-Ranks	G-Rd. Rnd-Ranks	V-Rd. Censoring	G-Rd. Censoring
Modifiziert†	-0.0357 (± 0.0915)	-0.0542 (± 0.1109)	-0.0150 (± 0.0823)	-0.0701 (± 0.1403)	-0.0361 (± 0.0915)	-0.0362 (± 0.0915)	+0.0324 (± 0.0505)	-0.0357 (± 0.0915)	-0.0652 (± 0.1296)	-0.0501 (± 0.1326)	+0.0293 (± 0.0459)	+0.0336 (± 0.0499)
$r = 10^{-6}$	+0.0022 (± 0.0323)	+0.0022 (± 0.0323)	+0.0022 (± 0.0323)	+0.0022 (± 0.0323)	+0.0022 (± 0.0323)	+0.0022 (± 0.0323)	+0.0248 (± 0.0318)	+0.0022 (± 0.0323)	+0.0022 (± 0.0323)	+0.0022 (± 0.0323)	+0.0248 (± 0.0318)	+0.0248 (± 0.0318)
$r = 10^{-4}$	+0.0052 (± 0.0318)	+0.0052 (± 0.0318)	+0.0052 (± 0.0317)	+0.0052 (± 0.0318)	+0.0051 (± 0.0318)	+0.0051 (± 0.0318)	+0.0242 (± 0.0311)	+0.0052 (± 0.0318)	+0.0052 (± 0.0318)	+0.0052 (± 0.0317)	+0.0242 (± 0.0311)	+0.0242 (± 0.0311)
$r = 10^{-2}$	+0.0074 (± 0.0312)	+0.0022 (± 0.0300)	+0.0075 (± 0.0334)	+0.0049 (± 0.0303)	+0.0028 (± 0.0306)	+0.0017 (± 0.0308)	+0.0219 (± 0.0316)	+0.0074 (± 0.0312)	-0.0065 (± 0.0295)	+0.0067 (± 0.0361)	+0.0140 (± 0.0254)	+0.0255 (± 0.0368)
$p = 0.025$	-0.0363 (± 0.0501)	-0.0363 (± 0.0501)	-0.0148 (± 0.0398)	-0.0375 (± 0.0510)	-0.0363 (± 0.0501)	-0.0363 (± 0.0237)	+0.0167 (± 0.0501)	-0.0363 (± 0.0501)	-0.0375 (± 0.0418)	-0.0207 (± 0.0511)	+0.0169 (± 0.0238)	+0.0199 (± 0.0273)
$p = 0.050$	-0.0727 (± 0.0819)	-0.0731 (± 0.0823)	-0.0094 (± 0.0486)	-0.0760 (± 0.0849)	-0.0727 (± 0.0819)	-0.0727 (± 0.0138)	+0.0089 (± 0.0819)	-0.0727 (± 0.0819)	-0.0766 (± 0.0854)	-0.0250 (± 0.0508)	+0.0086 (± 0.0138)	+0.0248 (± 0.0315)
$p = 0.100$	-0.1348 (± 0.1417)	-0.1391 (± 0.1459)	+0.0350 (± 0.0601)	-0.1425 (± 0.1491)	-0.1348 (± 0.1418)	-0.1348 (± 0.1418)	- (-)	-0.1348 (± 0.1417)	-0.1486 (± 0.1550)	-0.0020 (± 0.0453)	- (-)	+0.0302 (± 0.0374)
W×W	-0.0239 (± 0.1631)	-0.2335 (± 0.2677)	-0.0324 (± 0.1704)	-0.2688 (± 0.3026)	-0.0239 (± 0.1631)	-0.0239 (± 0.0928)	+0.0572 (± 0.1631)	-0.0239 (± 0.2315)	-0.1899 (± 0.2965)	-0.2606 (± 0.0777)	+0.0386 (± 0.0931)	+0.0583 (± 0.0931)
W×A	-0.0925 (± 0.1385)	-0.0952 (± 0.1416)	-0.0913 (± 0.1387)	-0.1685 (± 0.2308)	-0.0925 (± 0.1385)	-0.0925 (± 0.0666)	+0.0469 (± 0.1385)	-0.0925 (± 0.2327)	-0.1696 (± 0.2325)	-0.1570 (± 0.0665)	+0.0470 (± 0.0672)	+0.0507 (± 0.0672)
W×B	-0.0925 (± 0.1393)	-0.0925 (± 0.1393)	-0.0919 (± 0.1394)	-0.1700 (± 0.2316)	-0.0925 (± 0.1393)	-0.0925 (± 0.0585)	+0.0425 (± 0.1393)	-0.0925 (± 0.2325)	-0.1706 (± 0.2371)	-0.1621 (± 0.0585)	+0.0425 (± 0.0608)	+0.0473 (± 0.0608)
A×A	+0.0043 (± 0.0320)	+0.0043 (± 0.0320)	+0.0045 (± 0.0323)	+0.0043 (± 0.0320)	+0.0043 (± 0.0320)	+0.0043 (± 0.0310)	+0.0243 (± 0.0320)	+0.0043 (± 0.0320)	+0.0043 (± 0.0325)	+0.0060 (± 0.0319)	+0.0243 (± 0.0319)	+0.0253 (± 0.0319)
A×B	+0.0037 (± 0.0310)	+0.0037 (± 0.0310)	+0.0038 (± 0.0312)	+0.0037 (± 0.0310)	+0.0037 (± 0.0310)	+0.0037 (± 0.0314)	+0.0241 (± 0.0310)	+0.0037 (± 0.0310)	+0.0037 (± 0.0313)	+0.0050 (± 0.0314)	+0.0241 (± 0.0321)	+0.0247 (± 0.0321)
B×B	+0.0017 (± 0.0306)	+0.0017 (± 0.0306)	+0.0017 (± 0.0306)	+0.0017 (± 0.0306)	+0.0017 (± 0.0306)	+0.0017 (± 0.0304)	+0.0241 (± 0.0306)	+0.0017 (± 0.0306)	+0.0017 (± 0.0306)	+0.0017 (± 0.0304)	+0.0241 (± 0.0304)	+0.0241 (± 0.0304)

Tabelle B.11: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gumbel-Copula mit $\tau = 0.50$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

MPLE $\hat{\tau}_n = \tau(\hat{\vartheta}_n)$	Gumbel-Copula mit Kendall'schem Tau $\tau = 0.75$ Empirischer Bias und RMSE ($\pm \dots$) des Schätzers $\hat{\tau}_n$.											
Modifikation der Daten*	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	Π-Jitter	mix $_{\hat{\tau}_n}$ -Jitter	Censoring	Kojadinovic	V-Red. Rnd-Ranks	G-Red. Rnd-Ranks	V-Red. Censoring	G-Red. Censoring
Modifiziert†	-0.0790 (± 0.1511)	-0.0988 (± 0.1801)	-0.0518 (± 0.1340)	-0.1230 (± 0.2281)	-0.0795 (± 0.1511)	-0.0795 (± 0.1511)	-0.0287 (± 0.0722)	-0.0790 (± 0.1511)	-0.1112 (± 0.2057)	-0.0963 (± 0.2166)	-0.0327 (± 0.0736)	-0.0068 (± 0.0377)
$r = 10^{-6}$	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)	-0.0011 (± 0.0193)
$r = 10^{-4}$	-0.0021 (± 0.0187)	-0.0021 (± 0.0187)	-0.0021 (± 0.0187)	-0.0022 (± 0.0187)	-0.0022 (± 0.0187)	-0.0022 (± 0.0187)	-0.0021 (± 0.0187)	-0.0021 (± 0.0187)	-0.0022 (± 0.0187)	-0.0022 (± 0.0187)	-0.0021 (± 0.0187)	-0.0021 (± 0.0187)
$r = 10^{-2}$	-0.0012 (± 0.0178)	-0.0058 (± 0.0185)	-0.0014 (± 0.0181)	-0.0052 (± 0.0188)	-0.0064 (± 0.0197)	-0.0072 (± 0.0180)	+0.0001 (± 0.0178)	-0.0012 (± 0.0178)	-0.0173 (± 0.0242)	-0.0051 (± 0.0213)	-0.0141 (± 0.0221)	+0.0007 (± 0.0220)
$p = 0.025$	-0.0616 (± 0.0680)	-0.0616 (± 0.0680)	-0.0345 (± 0.0439)	-0.0629 (± 0.0692)	-0.0616 (± 0.0680)	-0.0616 (± 0.0677)	-0.0613 (± 0.0680)	-0.0616 (± 0.0693)	-0.0395 (± 0.0482)	-0.0613 (± 0.0677)	-0.0390 (± 0.0475)	
$p = 0.050$	-0.1141 (± 0.1193)	-0.1143 (± 0.1196)	-0.0330 (± 0.0547)	-0.1179 (± 0.1232)	-0.1141 (± 0.1193)	-0.1141 (± 0.1193)	-0.1133 (± 0.1187)	-0.1141 (± 0.1193)	-0.1182 (± 0.1235)	-0.0454 (± 0.0636)	-0.1138 (± 0.1192)	-0.0438 (± 0.0616)
$p = 0.100$	-0.2149 (± 0.2190)	-0.2175 (± 0.2215)	+0.0028 (± 0.0513)	-0.2252 (± 0.2292)	-0.2149 (± 0.2190)	-0.2149 (± 0.2190)	-0.2005 (± 0.2030)	-0.2149 (± 0.2190)	-0.2294 (± 0.2334)	-0.0179 (± 0.0555)	-0.2034 (± 0.2056)	-0.0152 (± 0.0453)
W×W	-0.1366 (± 0.2348)	-0.3640 (± 0.4091)	-0.1379 (± 0.2360)	-0.4451 (± 0.4890)	-0.1366 (± 0.2348)	-0.1366 (± 0.2348)	+0.0074 (± 0.0556)	-0.1366 (± 0.2348)	-0.2897 (± 0.3516)	-0.4434 (± 0.4900)	-0.0311 (± 0.0767)	+0.0073 (± 0.0555)
W×A	-0.2048 (± 0.2733)	-0.2074 (± 0.2764)	-0.2035 (± 0.2734)	-0.3038 (± 0.3935)	-0.2048 (± 0.2733)	-0.2048 (± 0.2733)	+0.0046 (± 0.0460)	-0.2048 (± 0.2733)	-0.3016 (± 0.3902)	-0.2962 (± 0.3927)	+0.0042 (± 0.0460)	+0.0101 (± 0.0467)
W×B	-0.2077 (± 0.2744)	-0.2077 (± 0.2744)	-0.2068 (± 0.2744)	-0.3082 (± 0.3947)	-0.2077 (± 0.2744)	-0.2077 (± 0.2744)	+0.0005 (± 0.0398)	-0.2077 (± 0.2744)	-0.3073 (± 0.3938)	-0.3027 (± 0.3962)	+0.0005 (± 0.0398)	+0.0051 (± 0.0407)
A×A	-0.0012 (± 0.0168)	-0.0012 (± 0.0168)	-0.0012 (± 0.0168)	-0.0012 (± 0.0168)	-0.0012 (± 0.0168)	-0.0012 (± 0.0168)	-0.0012 (± 0.0168)	-0.0012 (± 0.0168)	-0.0001 (± 0.0168)	-0.0012 (± 0.0168)	-0.0001 (± 0.0168)	-0.0001 (± 0.0167)
A×B	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0008 (± 0.0187)	-0.0018 (± 0.0187)	-0.0018 (± 0.0187)	-0.0008 (± 0.0187)
B×B	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)	-0.0014 (± 0.0179)

Tabelle B.12: Resultate der Modellanpassung bzw. Schätzung des Kendall'schen Taus für die Gumbel-Copula mit $\tau = 0.75$. Es wurden $N = 500$ Iterationen mit Stichproben der Größe $n = 250$ berücksichtigt. Die MPLE $\hat{\tau}_n(x_k)$ des Parameters ϑ_0 auf Grundlage der Stichprobe x_k wurden über die Funktionen aus Beispiel 2.4.4 in die jeweiligen Werte $\hat{\tau}_n$ des Kendall'schen Taus übersetzt. Die niedrigsten Werte für Bias und RMSE je Copula und Modifikation sind blau und, falls Bias und RMSE gleichzeitig minimal sind, rot hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). Der Stichprobenumfang beträgt hier $N = 6\,000$. Quelle: Eigene Erstellung.

Anhang C

Detaillierte Testresultate zum Test der Anpassungsgüte

Die Tabellen C.1 bis C.4 umfassen die Simulationsergebnisse bzgl. der Tests der Anpassungsgüte für die Copulas aus den Gauß'schen, t-, Clayton- und Gumbel-Copulafamilien untereinander.

In den Tabellen sind die \mathcal{H}_0 -Ablehnungsraten für die jeweiligen \mathcal{H}_1 - und \mathcal{H}_0 -Copulas dargestellt. Rote Einträge geben Werte unterhalb 5.0 % an. Blaue Einträge umfassen die Methoden mit den höchsten Ablehnungsraten je Copula, Kendall'schem Tau und Datenmodifikation.

Es wurden für diese Ergebnisse keine zusätzlichen Abbildungen bereitgestellt.

Tests der Anpassungsgüte		<i>Gauß'sche Copulafamilie (\mathcal{H}_0), $\tau \in \{0.25, 0.50, 0.75\}$</i> \mathcal{H}_0 -Ablehnungsraten in % für \mathcal{H}_0 - und \mathcal{H}_1 -Copulas											
Modifikation der Daten*	Copula-familien	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix _{$\hat{\tau}_n$} -Jitter	Censoring	Kojadinovic	V-Reduct. Rnd-Ranks	G-Reduct. Rnd-Ranks	V-Reduct. Censoring	V-Reduct. Censoring
Modifiziert [†] ($N = 2325$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	30.2 31.6 98.1 37.8	30.3 32.4 99.2 38.6	26.3 25.8 97.1 35.5	28.3 24.7 99.6 33.6	28.5 25.7 99.3 33.6	30.8 27.6 99.5 36.6	54.7 100.0 99.6 100.0	23.3 23.5 98.0 29.6	28.0 25.6 99.1 33.5	22.7 19.2 99.0 31.6	55.0 100.0 99.3 100.0	46.5 100.0 97.3 100.0
$r = 10^{-6}$ ($N = 150$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	4.7 14.0 98.4 33.3	4.0 12.0 99.1 33.3	4.9 13.9 99.2 32.6	4.0 11.3 98.4 36.0	5.3 14.7 98.4 30.0	4.7 14.0 98.4 34.0	12.0 100.0 100.0 —	3.3 14.0 97.6 32.7	4.7 14.0 99.2 31.3	4.2 13.2 99.1 33.3	11.3 100.0 100.0 —	13.2 100.0 100.0 —
$r = 10^{-4}$ ($N = 200$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	3.5 15.5 98.2 35.0	3.0 15.0 100.0 35.5	4.8 14.3 98.1 33.3	5.5 12.0 98.8 34.0	5.0 14.0 98.2 34.5	3.5 14.5 99.5 36.0	12.0 100.0 99.5 —	5.0 14.5 98.8 33.5	5.0 13.5 98.2 34.0	4.8 19.0 100.0 33.3	14.5 100.0 100.0 —	23.8 100.0 100.0 —
$r = 10^{-2}$ ($N = 150$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	4.7 23.3 97.7 43.3	6.7 20.7 98.8 40.0	6.0 16.0 97.5 31.3	5.3 19.3 100.0 35.3	5.3 20.7 98.8 32.7	5.3 20.0 98.8 34.7	51.3 — 95.0 —	4.7 18.7 98.8 32.7	7.3 18.0 100.0 33.3	4.0 14.7 100.0 28.7	55.3 — 93.3 —	47.3 — 87.0 —
$p = 0.025$ ($N = 250$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	42.8 18.8 100.0 24.0	41.6 19.2 100.0 23.6	38.4 15.3 98.7 26.0	49.6 18.0 100.0 27.2	50.8 19.2 100.0 26.0	51.2 18.0 100.0 —	69.2 100.0 100.0 —	48.8 19.6 100.0 26.4	49.2 16.4 100.0 26.8	40.1 16.4 100.0 29.4	70.0 100.0 100.0 —	62.7 100.0 99.1 —
$p = 0.050$ ($N = 225$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	56.4 47.1 100.0 29.8	58.2 49.3 100.0 28.9	39.9 25.7 97.9 23.9	67.6 48.0 100.0 33.3	68.0 44.9 100.0 33.8	67.6 46.2 100.0 35.1	90.7 100.0 100.0 —	66.7 51.1 100.0 40.0	68.0 47.1 100.0 35.1	48.2 27.1 100.0 28.0	90.2 100.0 100.0 —	65.1 100.0 96.7 —
$p = 0.100$ ($N = 225$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	78.2 76.0 100.0 58.7	79.1 76.9 100.0 62.2	27.1 23.1 100.0 32.9	81.8 73.8 100.0 57.3	80.4 72.9 100.0 58.2	81.8 72.9 100.0 58.7	100.0 100.0 100.0 —	81.8 78.2 100.0 63.6	84.9 76.4 100.0 60.0	29.3 20.9 96.2 32.9	100.0 100.0 100.0 —	41.3 100.0 92.5 —
W×W ($N = 150$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	47.0 46.3 76.6 47.0	53.0 56.4 100.0 56.8	49.0 47.7 73.5 48.3	28.7 20.7 100.0 28.7	30.0 26.7 100.0 31.3	61.3 56.7 100.0 66.0	100.0 — 100.0 100.0	7.4 12.1 75.0 8.1	22.7 19.3 85.7 25.5	28.7 26.0 100.0 31.3	100.0 100.0 100.0 100.0	99.3 100.0 100.0 100.0
W×A ($N = 150$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	54.7 57.3 100.0 58.0	54.7 56.0 100.0 56.0	54.0 58.0 100.0 60.0	32.0 21.3 100.0 32.7	33.3 23.3 100.0 36.0	35.3 26.7 100.0 36.0	100.0 100.0 100.0 —	6.0 6.0 100.0 8.7	32.0 26.7 100.0 34.0	38.7 28.7 100.0 40.7	100.0 100.0 100.0 —	73.8 100.0 99.0 —
W×B ($N = 150$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	52.7 56.7 100.0 60.0	52.0 58.7 100.0 63.0	55.3 59.3 100.0 63.3	30.7 23.3 100.0 34.0	30.7 26.7 100.0 38.0	34.7 26.0 100.0 35.3	100.0 100.0 100.0 —	4.7 8.7 100.0 8.0	28.0 24.7 100.0 37.3	32.0 28.7 97.4 41.3	100.0 100.0 100.0 —	73.2 100.0 100.0 —
A×A ($N = 275$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	4.4 13.1 100.0 30.2	4.0 12.7 100.0 28.7	4.9 10.5 100.0 29.4	4.7 13.1 100.0 29.1	4.0 13.8 100.0 29.1	3.6 12.4 100.0 29.1	17.5 99.5 100.0 —	5.5 12.0 100.0 27.3	4.4 12.4 100.0 28.0	4.2 10.5 100.0 28.7	16.4 100.0 99.5 —	14.0 100.0 99.0 —
A×B ($N = 175$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	5.7 12.0 100.0 24.0	2.9 10.9 100.0 24.0	1.6 12.7 100.0 22.2	5.1 12.6 100.0 24.6	5.1 12.0 100.0 24.0	4.6 10.9 100.0 25.1	10.9 100.0 100.0 —	4.0 13.1 100.0 26.3	4.0 13.7 100.0 24.6	3.2 12.7 100.0 24.6	12.6 100.0 97.6 —	11.1 100.0 96.6 —
B×B ($N = 225$)	Gauß (\mathcal{H}_0) t (\mathcal{H}_1) Clayton (\mathcal{H}_1) Gumbel (\mathcal{H}_1)	7.6 11.6 98.4 28.0	6.7 14.2 98.5 32.4	6.2 11.1 98.5 29.8	6.2 12.9 99.2 30.2	6.2 12.9 99.2 29.8	6.2 14.2 100.0 31.1	16.9 100.0 100.0 —	6.7 13.8 99.2 29.3	5.3 12.9 99.2 30.2	5.8 10.7 99.4 29.8	15.1 100.0 99.4 —	14.7 100.0 99.4 —

Tabelle C.1: \mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der Gauß'schen Copula. Die Stichprobengröße je Iteration (N) ist $n = 250$. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen höchsten Ablehnungsraten sind blau hervorgehoben.
*: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Fall *Modifiziert* umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Tests der Anpassungsgüte		<i>t-Copulafamilie, $\nu = 4$ (\mathcal{H}_0), $\tau \in \{0.25, 0.50, 0.75\}$</i> \mathcal{H}_0 -Ablehnungsraten in % für \mathcal{H}_0 - und \mathcal{H}_1 -Copulas											
Modifikation der Daten*	Copula-familien	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix $_{\hat{\tau}_n}$ -Jitter	Censoring	Kojadinovic	V-Reduct. Rnd-Ranks	G-Reduct. Rnd-Ranks	V-Reduct. Censoring	V-Reduct. Censoring
Modifiziert [†] ($N = 2525$)	t (\mathcal{H}_0)	19.4	20.3	17.4	15.6	16.1	17.3	100.0	13.0	15.2	12.3	100.0	100.0
	Gauß (\mathcal{H}_1)	30.7	31.0	29.4	30.2	31.3	32.7	59.0	25.9	30.4	28.2	59.8	55.5
	Clayton (\mathcal{H}_1)	96.6	98.1	96.5	98.0	98.5	98.5	97.3	97.1	97.3	97.5	96.9	95.3
	Gumbel (\mathcal{H}_1)	28.7	29.0	29.0	26.4	26.7	28.5	—	21.9	26.3	25.8	—	—
$r = 10^{-6}$ ($N = 225$)	t (\mathcal{H}_0)	3.1	3.6	4.1	4.4	4.0	5.3	100.0	4.9	4.0	3.7	100.0	100.0
	Gauß (\mathcal{H}_1)	9.3	9.8	9.6	10.7	10.2	9.8	27.6	10.2	11.1	10.1	29.3	28.0
	Clayton (\mathcal{H}_1)	97.7	97.0	97.0	97.0	97.8	98.5	97.6	97.8	96.9	97.6	97.0	97.0
	Gumbel (\mathcal{H}_1)	18.7	22.2	19.7	20.4	21.8	21.8	—	20.4	21.3	20.2	—	—
$r = 10^{-4}$ ($N = 325$)	t (\mathcal{H}_0)	4.3	4.9	2.3	4.6	4.9	4.9	100.0	5.2	5.8	2.3	100.0	100.0
	Gauß (\mathcal{H}_1)	11.1	10.5	14.0	11.1	12.6	11.7	31.7	10.8	11.1	11.6	35.7	34.9
	Clayton (\mathcal{H}_1)	96.7	97.3	95.8	97.3	97.8	97.8	96.5	97.3	96.7	96.2	95.1	87.9
	Gumbel (\mathcal{H}_1)	20.6	21.2	39.5	22.5	22.5	22.5	—	22.8	22.2	34.9	—	—
$r = 10^{-2}$ ($N = 150$)	t (\mathcal{H}_0)	6.0	6.0	4.7	7.3	6.7	6.7	—	5.3	5.3	4.7	—	—
	Gauß (\mathcal{H}_1)	14.0	12.7	10.0	14.7	13.3	14.0	66.7	14.7	14.0	13.3	58.7	57.3
	Clayton (\mathcal{H}_1)	95.8	96.7	94.7	97.7	98.9	98.9	93.3	95.7	97.6	92.9	91.5	93.0
	Gumbel (\mathcal{H}_1)	26.7	24.7	17.3	22.7	19.3	21.3	—	18.7	21.3	15.3	—	—
$p = 0.025$ ($N = 250$)	t (\mathcal{H}_0)	17.2	16.0	12.2	15.2	14.8	15.2	100.0	15.6	14.8	10.1	100.0	100.0
	Gauß (\mathcal{H}_1)	53.2	52.0	49.2	58.0	59.6	59.2	76.8	60.8	59.6	52.9	77.2	70.4
	Clayton (\mathcal{H}_1)	100.0	100.0	98.9	100.0	100.0	100.0	99.4	100.0	100.0	98.9	98.7	97.6
	Gumbel (\mathcal{H}_1)	16.8	17.6	18.0	18.8	18.0	19.2	—	20.8	18.4	19.6	—	—
$p = 0.050$ ($N = 225$)	t (\mathcal{H}_0)	36.0	38.2	14.5	32.9	35.6	33.3	100.0	39.6	32.0	13.2	100.0	100.0
	Gauß (\mathcal{H}_1)	64.0	64.0	41.4	77.3	77.3	76.9	95.1	74.2	76.9	45.0	96.9	68.2
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.7	100.0	96.8
	Gumbel (\mathcal{H}_1)	28.4	24.9	28.2	29.8	31.6	31.1	—	34.7	32.0	29.1	—	—
$p = 0.100$ ($N = 150$)	t (\mathcal{H}_0)	66.0	70.0	20.0	62.7	60.7	60.7	100.0	66.0	65.3	20.7	100.0	100.0
	Gauß (\mathcal{H}_1)	72.0	74.7	34.7	73.3	77.3	75.3	100.0	73.3	80.7	40.0	100.0	55.3
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	85.9
	Gumbel (\mathcal{H}_1)	54.0	55.3	26.0	48.0	50.7	49.3	—	54.7	50.7	22.0	—	—
W×W ($N = 150$)	t (\mathcal{H}_0)	42.0	52.0	42.0	28.7	30.0	60.7	—	10.0	22.0	23.3	100.0	100.0
	Gauß (\mathcal{H}_1)	48.0	56.0	48.7	40.7	47.3	72.7	100.0	9.3	26.7	41.3	100.0	100.0
	Clayton (\mathcal{H}_1)	75.8	86.7	78.3	100.0	100.0	100.0	100.0	83.1	81.8	100.0	100.0	100.0
	Gumbel (\mathcal{H}_1)	36.7	50.7	39.3	34.0	37.3	70.0	—	4.0	21.5	37.3	—	—
W×A ($N = 150$)	t (\mathcal{H}_0)	46.7	47.3	48.0	22.7	23.3	18.0	100.0	8.7	23.3	26.0	100.0	100.0
	Gauß (\mathcal{H}_1)	48.0	49.3	49.3	32.0	30.0	28.7	100.0	12.7	33.3	32.0	100.0	77.2
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	98.0
	Gumbel (\mathcal{H}_1)	48.7	47.3	48.7	28.0	30.7	26.7	—	4.7	32.0	32.0	—	—
W×B ($N = 150$)	t (\mathcal{H}_0)	47.3	47.3	51.3	28.7	30.7	29.3	100.0	4.7	27.3	32.0	100.0	100.0
	Gauß (\mathcal{H}_1)	52.0	51.3	59.3	39.3	40.7	45.3	100.0	13.3	45.3	48.0	100.0	85.2
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	97.3	100.0	95.0
	Gumbel (\mathcal{H}_1)	52.7	51.3	56.7	41.3	39.3	42.0	—	8.0	45.3	44.0	—	—
A×A ($N = 275$)	t (\mathcal{H}_0)	3.6	4.0	4.4	3.6	4.7	4.0	100.0	4.0	4.7	5.1	100.0	100.0
	Gauß (\mathcal{H}_1)	13.1	12.0	13.2	12.0	13.8	13.5	32.7	13.8	12.7	11.8	33.8	35.3
	Clayton (\mathcal{H}_1)	96.6	97.3	96.0	96.1	98.1	97.4	93.6	96.6	96.0	96.1	94.6	94.8
	Gumbel (\mathcal{H}_1)	20.0	20.0	19.1	19.3	20.7	18.5	—	18.5	18.2	17.6	—	—
A×B ($N = 250$)	t (\mathcal{H}_0)	5.2	3.6	5.4	4.8	4.8	4.4	100.0	4.0	4.4	4.9	100.0	100.0
	Gauß (\mathcal{H}_1)	10.0	9.6	10.9	10.0	8.8	9.6	27.2	10.0	8.8	10.3	28.4	27.7
	Clayton (\mathcal{H}_1)	97.6	99.2	100.0	99.2	98.4	99.2	96.6	99.2	98.4	98.9	96.0	96.1
	Gumbel (\mathcal{H}_1)	28.4	24.8	26.6	26.0	23.2	24.8	—	24.8	25.2	23.4	—	—
B×B ($N = 225$)	t (\mathcal{H}_0)	4.9	4.0	4.9	4.4	5.3	4.9	100.0	4.0	4.0	4.0	100.0	100.0
	Gauß (\mathcal{H}_1)	12.4	13.3	11.6	11.6	13.8	12.9	29.3	13.3	12.4	13.8	30.7	33.3
	Clayton (\mathcal{H}_1)	97.7	97.7	98.5	97.6	97.7	96.9	95.5	96.9	98.5	97.7	94.9	94.4
	Gumbel (\mathcal{H}_1)	24.9	23.1	24.9	24.0	24.4	23.6	—	24.4	25.3	24.0	—	—

Tabelle C.2: \mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der t-Copula. Die Stichprobengröße je Iteration (N) ist $n = 250$. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen höchsten Ablehnungsraten sind blau hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best).

†: Der Fall *Modifiziert* umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Tests der Anpassungsgüte		<i>Clayton-Copulafamilie (\mathcal{H}_0), $\tau \in \{0.25, 0.50, 0.75\}$</i> \mathcal{H}_0 -Ablehnungsraten in % für \mathcal{H}_0 - und \mathcal{H}_1 -Copulas											
Modifikation der Daten*	Copula-familien	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix _{$\hat{\tau}_n$} -Jitter	Censoring	Kojadinovic	V-Reduct., Rnd-Ranks	G-Reduct., Rnd-Ranks	V-Reduct., Censoring	V-Reduct., Censoring
Modifiziert [†] (N = 9000)	Clayton (\mathcal{H}_0)	25.5	26.2	18.4	21.1	21.3	22.9	50.5	22.0	21.0	21.3	49.8	45.3
	Gauß (\mathcal{H}_1)	65.8	66.5	64.0	68.1	68.1	70.8	—	60.2	67.8	65.1	—	—
	t (\mathcal{H}_1)	64.0	64.5	62.1	62.5	62.3	65.3	—	58.5	62.9	59.6	—	—
	Gumbel (\mathcal{H}_1)	82.8	83.7	82.3	83.1	82.9	86.3	100.0	78.9	84.1	81.8	99.8	100.0
$r = 10^{-6}$ (N = 750)	Clayton (\mathcal{H}_0)	5.6	4.8	3.7	5.0	4.4	5.6	12.3	4.6	4.4	4.5	11.3	11.1
	Gauß (\mathcal{H}_1)	70.9	70.8	71.9	71.3	70.5	71.7	—	71.5	70.9	71.6	—	—
	t (\mathcal{H}_1)	72.0	71.9	72.7	72.0	71.3	72.4	—	71.5	72.3	70.9	—	—
	Gumbel (\mathcal{H}_1)	93.3	93.7	93.7	93.5	93.7	93.7	100.0	93.3	93.7	93.7	100.0	100.0
$r = 10^{-4}$ (N = 750)	Clayton (\mathcal{H}_0)	6.0	4.8	7.1	5.0	5.2	5.8	9.1	4.8	5.4	7.1	8.2	9.8
	Gauß (\mathcal{H}_1)	71.5	70.3	71.1	72.0	70.8	70.5	—	71.7	71.6	71.1	—	—
	t (\mathcal{H}_1)	71.5	70.4	70.0	70.9	72.1	71.3	—	71.3	70.7	71.1	—	—
	Gumbel (\mathcal{H}_1)	93.2	93.5	93.3	93.5	92.8	93.9	100.0	92.8	93.3	93.3	100.0	100.0
$r = 10^{-2}$ (N = 750)	Clayton (\mathcal{H}_0)	3.0	4.0	3.4	5.4	6.0	4.6	33.4	3.2	5.0	4.6	28.8	27.3
	Gauß (\mathcal{H}_1)	72.1	70.4	61.6	70.4	70.9	70.1	—	70.7	69.1	60.5	—	—
	t (\mathcal{H}_1)	72.8	70.3	58.9	71.9	71.6	71.9	—	70.8	69.2	59.3	—	—
	Gumbel (\mathcal{H}_1)	96.0	94.8	91.3	93.7	93.5	93.7	100.0	93.7	92.7	88.4	100.0	99.6
$p = 0.025$ (N = 750)	Clayton (\mathcal{H}_0)	71.6	71.3	58.7	84.2	84.0	85.0	97.0	82.6	84.6	69.0	96.6	84.0
	Gauß (\mathcal{H}_1)	71.3	71.3	71.8	77.1	77.6	76.8	—	76.1	77.3	75.5	—	—
	t (\mathcal{H}_1)	54.4	54.3	57.5	58.3	57.3	56.7	—	59.1	58.0	60.8	—	—
	Gumbel (\mathcal{H}_1)	89.6	89.7	91.2	90.1	90.1	90.1	100.0	89.9	90.8	91.7	100.0	100.0
$p = 0.050$ (N = 750)	Clayton (\mathcal{H}_0)	88.4	91.2	59.3	98.1	99.5	98.6	100.0	98.7	99.1	60.3	100.0	63.3
	Gauß (\mathcal{H}_1)	69.5	69.6	70.6	80.1	79.2	78.8	—	78.7	79.9	75.6	—	—
	t (\mathcal{H}_1)	54.4	54.8	58.6	59.2	58.7	60.0	—	63.1	60.3	61.5	—	—
	Gumbel (\mathcal{H}_1)	87.1	87.2	89.3	88.8	89.1	89.1	100.0	88.8	88.9	90.1	100.0	100.0
$p = 0.100$ (N = 750)	Clayton (\mathcal{H}_0)	98.8	99.3	22.6	100.0	100.0	100.0	100.0	99.4	100.0	14.1	100.0	22.5
	Gauß (\mathcal{H}_1)	72.1	72.3	69.7	84.7	84.7	85.9	—	82.5	85.6	71.9	—	—
	t (\mathcal{H}_1)	68.7	72.0	61.1	72.8	73.2	72.3	—	75.3	75.1	62.3	—	—
	Gumbel (\mathcal{H}_1)	80.7	80.9	90.5	84.8	84.3	84.0	100.0	83.5	84.7	90.8	100.0	100.0
$W \times W$ (N = 750)	Clayton (\mathcal{H}_0)	22.6	42.8	23.1	56.3	60.5	66.1	100.0	11.9	35.3	61.4	99.0	97.3
	Gauß (\mathcal{H}_1)	37.7	46.8	38.8	42.8	42.4	76.9	—	8.0	39.6	41.7	—	—
	t (\mathcal{H}_1)	45.9	50.7	47.1	34.5	34.1	72.9	—	28.9	40.0	35.9	—	—
	Gumbel (\mathcal{H}_1)	46.1	54.0	47.6	46.4	44.0	83.7	100.0	35.7	57.0	44.1	96.9	100.0
$W \times A$ (N = 750)	Clayton (\mathcal{H}_0)	14.9	17.1	11.6	34.8	36.8	35.9	99.8	18.2	35.0	33.3	99.8	71.7
	Gauß (\mathcal{H}_1)	56.8	57.2	58.4	52.5	54.4	54.4	—	25.2	53.6	57.1	—	—
	t (\mathcal{H}_1)	58.4	58.4	60.4	49.3	47.5	47.5	—	23.2	47.7	51.7	—	—
	Gumbel (\mathcal{H}_1)	64.8	64.8	65.9	62.5	64.1	62.4	100.0	42.1	64.4	64.7	100.0	100.0
$W \times B$ (N = 750)	Clayton (\mathcal{H}_0)	11.4	12.3	10.8	25.2	27.6	26.9	99.6	13.4	29.2	32.0	99.8	71.0
	Gauß (\mathcal{H}_1)	56.9	56.5	58.2	54.5	56.0	54.5	—	27.1	57.2	57.5	—	—
	t (\mathcal{H}_1)	59.3	59.8	60.7	52.1	50.4	48.9	—	26.6	50.8	51.2	—	—
	Gumbel (\mathcal{H}_1)	64.5	65.2	65.2	64.0	64.5	64.4	100.0	45.8	63.6	66.3	100.0	100.0
$A \times A$ (N = 750)	Clayton (\mathcal{H}_0)	4.2	4.0	3.7	4.2	3.6	4.0	10.5	4.4	4.6	4.1	10.5	11.5
	Gauß (\mathcal{H}_1)	70.7	71.6	68.9	70.8	70.3	70.7	—	70.4	70.7	70.1	—	—
	t (\mathcal{H}_1)	71.1	70.8	68.4	70.3	70.8	70.3	—	71.5	70.9	68.2	—	—
	Gumbel (\mathcal{H}_1)	92.0	92.9	91.8	92.8	93.1	92.8	100.0	93.5	92.9	93.0	100.0	100.0
$A \times B$ (N = 750)	Clayton (\mathcal{H}_0)	3.8	3.8	3.4	4.0	4.4	4.0	9.3	4.2	4.0	3.4	10.3	8.0
	Gauß (\mathcal{H}_1)	70.0	70.7	69.7	70.0	70.1	69.9	—	70.0	69.1	69.9	—	—
	t (\mathcal{H}_1)	70.5	71.2	71.1	69.9	70.8	70.8	—	70.3	70.4	70.9	—	—
	Gumbel (\mathcal{H}_1)	94.1	93.9	93.9	94.5	93.3	94.3	100.0	94.1	93.6	93.9	100.0	100.0
$B \times B$ (N = 750)	Clayton (\mathcal{H}_0)	3.4	4.0	4.0	3.0	3.2	3.4	8.9	4.2	3.2	3.8	9.4	9.1
	Gauß (\mathcal{H}_1)	69.5	70.4	69.5	71.1	70.0	69.1	—	70.0	69.5	70.5	—	—
	t (\mathcal{H}_1)	69.5	69.9	70.0	69.3	70.1	69.1	—	70.3	69.6	68.3	—	—
	Gumbel (\mathcal{H}_1)	92.0	93.5	93.2	92.9	92.8	93.2	100.0	92.9	92.9	92.5	100.0	100.0

Tabelle C.3: \mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der Clayton-Copula. Die Stichprobengröße je Iteration (N) ist $n = 250$. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen höchsten Ablehnungsraten sind blau hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten ($W = \text{Worst}$, $A = \text{Average}$, $B = \text{Best}$). †: Der Fall *Modifiziert* umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Tests der Anpassungsgüte		<i>Gumbel-Copulafamilie (\mathcal{H}_0), $\tau \in \{0.25, 0.50, 0.75\}$</i> \mathcal{H}_0 -Ablehnungsraten in % für \mathcal{H}_0 - und \mathcal{H}_1 -Copulas											
Modifikation der Daten*	Copula-familien	Reference	V-Reduct.	G-Reduct.	Rnd-Ranks	II-Jitter	mix _{$\hat{\tau}_n$} -Jitter	Censoring	Kojadinovic	V-Reduct. Rnd-Ranks	G-Reduct. Rnd-Ranks	V-Reduct. Censoring	V-Reduct. Censoring
Modifiziert [†] ($N = 8700$)	Gumbel (\mathcal{H}_0)	23.0	24.4	19.7	17.7	17.5	18.8	48.0	12.6	17.3	14.7	46.8	41.5
	Gauß (\mathcal{H}_1)	53.2	53.6	49.5	49.9	49.7	51.1	—	47.1	49.4	44.3	—	—
	t (\mathcal{H}_1)	49.6	50.5	43.7	43.6	43.4	45.0	—	42.7	43.9	37.3	—	—
	Clayton (\mathcal{H}_1)	99.5	100.0	99.3	100.0	100.0	100.0	100.0	99.4	100.0	99.9	100.0	100.0
$r = 10^{-6}$ ($N = 750$)	Gumbel (\mathcal{H}_0)	4.8	4.0	4.7	4.7	4.3	4.3	10.7	4.9	4.4	4.6	9.7	11.6
	Gauß (\mathcal{H}_1)	33.1	33.3	32.9	32.9	33.5	33.3	—	32.7	33.1	33.6	—	—
	t (\mathcal{H}_1)	30.9	31.3	33.3	31.6	32.0	31.3	—	33.2	32.8	31.8	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
$r = 10^{-4}$ ($N = 750$)	Gumbel (\mathcal{H}_0)	4.5	4.4	3.9	4.1	5.1	4.5	11.6	4.4	5.1	4.9	8.6	17.9
	Gauß (\mathcal{H}_1)	36.1	35.9	31.1	36.1	35.7	35.7	—	35.2	35.3	33.0	—	—
	t (\mathcal{H}_1)	34.9	34.3	33.0	33.2	34.4	34.4	—	34.8	35.5	33.0	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
$r = 10^{-2}$ ($N = 750$)	Gumbel (\mathcal{H}_0)	5.3	5.7	4.8	3.7	3.7	4.3	50.9	4.1	3.9	4.1	46.5	53.6
	Gauß (\mathcal{H}_1)	33.9	30.7	27.5	35.9	36.1	35.7	—	35.1	33.2	28.8	—	—
	t (\mathcal{H}_1)	32.7	29.9	25.6	34.0	33.3	33.7	—	32.1	30.4	27.7	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	99.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
$p = 0.025$ ($N = 750$)	Gumbel (\mathcal{H}_0)	10.5	10.5	7.3	11.6	11.5	11.5	38.5	12.1	10.9	8.4	37.6	23.9
	Gauß (\mathcal{H}_1)	69.9	69.6	63.7	72.8	73.1	73.1	—	71.9	72.8	64.9	—	—
	t (\mathcal{H}_1)	54.9	55.2	45.4	56.3	55.6	55.3	—	57.3	55.2	47.8	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
$p = 0.050$ ($N = 750$)	Gumbel (\mathcal{H}_0)	31.7	32.1	15.6	29.7	28.9	27.9	86.5	34.8	30.3	14.6	85.3	44.6
	Gauß (\mathcal{H}_1)	75.6	76.8	62.9	80.8	82.0	81.3	—	81.1	81.9	65.4	—	—
	t (\mathcal{H}_1)	68.1	67.6	49.2	68.3	66.9	67.3	—	70.4	69.3	48.5	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
$p = 0.100$ ($N = 750$)	Gumbel (\mathcal{H}_0)	65.5	67.7	15.3	56.9	57.2	55.6	100.0	61.9	59.5	13.3	100.0	24.4
	Gauß (\mathcal{H}_1)	85.5	88.1	47.7	87.7	87.7	86.8	—	86.7	89.6	48.0	—	—
	t (\mathcal{H}_1)	83.3	85.1	36.7	81.6	81.1	80.8	—	82.8	82.9	35.5	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	99.7	100.0	100.0	100.0	100.0	100.0	100.0	99.7	—	100.0
W×W ($N = 450$)	Gumbel (\mathcal{H}_0)	41.1	65.0	42.0	34.2	31.6	58.4	100.0	6.5	22.2	33.8	99.5	98.1
	Gauß (\mathcal{H}_1)	71.3	74.2	71.6	45.3	44.7	73.3	—	30.9	38.0	45.6	—	—
	t (\mathcal{H}_1)	54.0	74.9	54.0	31.3	34.4	61.6	—	29.6	36.0	35.3	—	—
	Clayton (\mathcal{H}_1)	84.5	100.0	85.7	100.0	100.0	100.0	99.3	87.3	97.2	100.0	100.0	99.3
W×A ($N = 750$)	Gumbel (\mathcal{H}_0)	53.9	53.5	56.5	31.1	30.3	32.5	100.0	2.9	32.3	35.5	100.0	76.4
	Gauß (\mathcal{H}_1)	66.5	66.5	66.0	47.5	47.2	46.4	—	38.1	47.3	47.2	—	—
	t (\mathcal{H}_1)	66.7	65.9	66.3	38.1	35.9	38.3	—	29.7	38.4	40.4	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
W×B ($N = 750$)	Gumbel (\mathcal{H}_0)	53.7	53.9	54.9	29.7	30.9	29.9	99.8	3.3	29.2	34.1	99.8	74.1
	Gauß (\mathcal{H}_1)	66.8	67.2	67.6	47.1	45.3	46.8	—	39.0	47.2	46.8	—	—
	t (\mathcal{H}_1)	66.4	66.1	66.4	41.1	38.4	40.1	—	32.4	39.6	41.2	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.2	100.0	100.0	100.0	100.0
A×A ($N = 750$)	Gumbel (\mathcal{H}_0)	4.5	4.8	4.3	4.9	4.5	4.0	11.5	4.3	4.1	4.8	12.2	12.6
	Gauß (\mathcal{H}_1)	37.5	38.4	37.4	37.9	37.7	37.3	—	36.7	37.1	36.9	—	—
	t (\mathcal{H}_1)	37.2	37.2	36.9	36.4	36.7	36.0	—	37.6	37.3	35.9	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
A×B ($N = 750$)	Gumbel (\mathcal{H}_0)	4.3	4.0	3.9	4.5	4.0	4.4	9.8	5.2	4.1	4.6	11.2	16.1
	Gauß (\mathcal{H}_1)	34.4	34.7	35.4	35.6	35.7	34.8	—	35.7	35.5	35.3	—	—
	t (\mathcal{H}_1)	34.9	34.4	33.5	34.0	34.9	33.1	—	34.3	33.1	35.1	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
B×B ($N = 750$)	Gumbel (\mathcal{H}_0)	4.0	4.0	3.7	3.6	3.3	4.0	11.8	4.3	4.0	3.7	12.7	13.0
	Gauß (\mathcal{H}_1)	35.2	36.0	37.3	36.9	35.7	37.7	—	35.6	37.1	36.1	—	—
	t (\mathcal{H}_1)	33.1	34.1	33.7	32.5	33.3	34.4	—	32.5	32.5	32.0	—	—
	Clayton (\mathcal{H}_1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Tabelle C.4: \mathcal{H}_0 -Ablehnungsraten der Tests der Anpassungsgüte bzgl. der Gumbel-Copula. Die Stichprobengröße je Iteration (N) ist $n = 250$. Die Ablehnungsraten sind für die Testentscheidung bzgl. des Fünfprozentniveaus berechnet worden; Werte kleiner als 5.0% sind rot und die zeilenweisen höchsten Ablehnungsraten sind blau hervorgehoben. *: Die Modifikationen umfassen Anpassungen der Messpräzision r , das Nullsetzen mit Wahrscheinlichkeit p sowie die Transformation bzgl. empirischer Daten (W = Worst, A = Average, B = Best). †: Der Fall *Modifiziert* umfasst alle genannten Modifikationen. Quelle: Eigene Erstellung.

Anhang D

Algorithmen und Implementierungen

Die Implementierungen benötigen jeweils das Paket Copula von Hofert et al. (2020); das Paket wird ad hoc über den Befehl `copula::` aktiviert.

D.1 Allgemeine Funktionen

D.1.1 Implementierung (Generierung synthetischer Datensätze). *Dieser Code entspricht dem Algorithmus 4.2.3 zur Generierung von gebundenen Datensätzen unter Modellierung von Finanzmarktrahmenbedingungen.*

```
1 generate_synthetic_data <- function(copula, n, inverse_marginals = c(NA),
2   rounding_precision = c(NA), omission_probability = c(0), x_default = c(0),
3   x_lower = c(NA), x_upper = c(NA)) {
4
5   # Initial processing of the parameter of the data generating algorithm
6   d <- data.frame(cop=dim(copula))
7   d$im <- length(inverse_marginals)
8   d$rp <- length(rounding_precision)
9   d$op <- length(omission_probability)
10  d$xl <- length(x_lower)
11  d$xu <- length(x_upper)
12  d$xd <- length(x_default)
13
14  # Validate the input of the inverse marginals.
15  if (! d$im %in% c(1, d$cop)) {
16    warning(paste0(
17      "Count of inverse marginals ", d$im, " is not 1 or ", d$cop, ". ",
18      "Missing functions are set to uniform marginals. ",
19      "Additional functions provided will be ignored."
20    ));
21    inverse_marginals <-
22      c(inverse_marginals, rep(NA, max(d$cop - d$im,0)))[1:d$cop]
23  } else {
24    inverse_marginals <- rep(inverse_marginals, d$cop / d$im)
25  }
26}
```

```

27 # Validate the input of the rounding precisions.
28 if (! d$rp %in% c(1, d$cop)) {
29   warning(paste0(
30     "Count of rounding precisions ", d$rp, " is not 1 or ", d$cop, ". ",
31     "Missing parameter will be set to maximum precision. ",
32     "Additional parameter will be ignored."
33  ));
34   rounding_precision <-
35     c(rounding_precision, rep(NA, max(d$cop - d$rp,0)))[1:d$cop]
36 } else {
37   rounding_precision <- rep(rounding_precision, d$cop / d$rp)
38 }
39
40 # Validate the input of the omission probabilities.
41 if (! d$op %in% c(1, d$cop)) {
42   warning(paste0(
43     "Count of omission probabilities ", d$op, " is not 1 or ", d$cop, ". "
44     ,
45     "Missing parameter will be set to zero. ",
46     "Additional parameter will be ignored."
47  ));
48   omission_probability <-
49     c(omission_probability, rep(0, max(d$cop - d$op,0)))[1:d$cop]
50 } else {
51   omission_probability <- rep(omission_probability, d$cop / d$op)
52 }
53
54 # Validate the input of the lower boundary vector.
55 if (! d$x1 %in% c(1, d$cop)) {
56   warning(paste0(
57     "Dimension of the lower boundary ", d$x1, " is not 1 or ", d$cop, ". "
58     ,
59     "Missing parameter will be set to NA / -infinity. ",
60     "Additional parameter will be ignored."
61  ));
62   x_lower <- c(x_lower, rep(NA, max(d$cop - d$x1,0)))[1:d$cop]
63 } else {
64   x_lower <- rep(x_lower, d$cop / d$x1)
65 }
66
67 # Validate the input of the upper boundary vector.
68 if (! d$xu %in% c(1, d$cop)) {
69   warning(paste0(
70     "Dimension of the upper boundary ", d$xu, " is not 1 or ", d$cop, ". "
71     ,
72     "Missing parameter will be set to NA / +infinity. ",
73     "Additional parameter will be ignored."
74  ));
75   x_upper <- c(x_upper, rep(NA, max(d$cop - d$xu,0)))[1:d$cop]
76 } else {
77   x_upper <- rep(x_upper, d$cop / d$xu)
78 }
79
80 # Validate the input of the default value boundary vector.
81 if (! d$xd %in% c(1, d$cop)) {
82   warning(paste0(
83     "Dimension of the default vector ", d$xd, " is not 1 or ", d$cop, ". "
84     ,
85     "Missing parameter will be set to zero. ",
86     "Additional parameter will be ignored."
87  ));

```

```

84   x_default <- c(x_default, rep(0, max(d$cop - d$xd,0)))[1:d$cop]
85 } else {
86   x_default <- rep(x_default, d$cop / d$xd)
87 }
88
89 # 1. Generate a sample  $u \sim$  copula of size  $n$  and initialize the two output
90 # variables  $x$  (modified) and  $y$  (unmodified).
91 u <- copula::rCopula(n, copula)
92 x <- y <- u
93
94 # 2. Quantile Transformation (ITS) w. r. t. the inverse marginals.
95 for (i in 1:d$cop) { if (is.function(inverse_marginals[[i]])) {
96   x[, i] <- inverse_marginals[[i]](u[, i])
97 }
98
99 # 3. Rounding of data.
100 for (i in 1:d$cop) { if (!is.na(rounding_precision[i])) {
101   x[, i] <- round(x[, i] / rounding_precision[i]) * rounding_precision[i]
102 }
103
104 # 4. Random omission of entries.
105 for (i in 1:d$cop) { if (!is.na(omission_probability[i])) {
106   if (0 < omission_probability[i] && omission_probability[i] < 1) {
107     b <- rbinom(n,1,omission_probability[i])
108     x[, i] <- (1 - b) * x[, i] + b * x_default[i]
109   }
110 }
111
112 # 5. Limiting of values w. r. t. the boundaries.
113 for (i in 1:d$cop) { if (!is.na(x_lower[i])) {
114   x[, i][x[, i] <= x_lower[i]] <- x_lower[i]
115 }
116 for (i in 1:d$cop) { if (!is.na(x_upper[i])) {
117   x[, i][x[, i] >= x_upper[i]] <- x_upper[i]
118 }
119
120 # Output of the generated data.
121 out <- list()
122 out$modified <- x
123 out$unmodified <- y
124
125 return(out)
126 }

```

D.2 Grundlegende Tests ohne Modifikation

D.2.1 Implementierung (Rangbasierter Test auf Radial-Symmetrie mit Approximation nach Genest und Nešlehová (2013)). *Die Implementierung in R basiert auf dem Code von Quessy (2021) bereitgestellten MATLAB-Implementierung RadialSym_EmpCopula_BIV.m.*

```

1 radSymTest_genestEtNeslehova <- function(X, ties.method="max", b=1, N=250) {
2   # Calculate the pseudo observations U w.r.t. to the sample X. Measure the
3   # sample size and construct the size-adapted binwidth bn.
4   U <- copula::pobs(X, ties.method = ties.method);
5   n <- nrow(X);
6   bn <- b / sqrt(n);
7
8   # Calculation of the statistic Rn.
9   E <- matrix(nrow = n, ncol = n);
10  for (i.1 in 1:n) { for (i.2 in 1:n) {
11    E[i.1, i.2] <- prod(U[i.1,] <= U[i.2,]) - prod(1 - U[i.1,] <= U[i.2,]);
12  }
13  Rn <- sum(E %*% t(E)) / (n^2);
14
15  # Calculation of the deterministic terms.
16  dCdu1 <- rep(0, n); dCdu2 <- rep(0, n);
17  for (i in 1:n) {
18    C.upr.1 <- mean((U[, 1] <= (U[i, 1] + bn)) * (U[, 2] <= U[i, 2]));
19    C.lwr.1 <- mean((U[, 1] <= (U[i, 1] - bn)) * (U[, 2] <= U[i, 2]));
20    C.upr.2 <- mean((U[, 1] <= U[i, 1]) * (U[, 2] <= (U[i, 2] + bn)));
21    C.lwr.2 <- mean((U[, 1] <= U[i, 1]) * (U[, 2] <= (U[i, 2] - bn)));
22    dCdu1[i] <- (C.upr.1 - C.lwr.1) / (2 * bn);
23    dCdu2[i] <- (C.upr.2 - C.lwr.2) / (2 * bn);
24  }
25
26  E <- matrix(nrow = n, ncol = n);
27  for (i.1 in 1:n) { for (i.2 in 1:n) {
28    Pn <- prod(U[i.1, ] <= U[i.2, ]) - prod((1 - U[i.1, ]) <= U[i.2, ]);
29    Pn.1 <- (U[i.1, 1] <= U[i.2, 1]) - (1 - U[i.1, 1] <= U[i.2, 1]);
30    Pn.2 <- (U[i.1, 2] <= U[i.2, 2]) - (1 - U[i.1, 2] <= U[i.2, 2]);
31    dCdu.Pn.1 <- dCdu1[i.2] * Pn.1;
32    dCdu.Pn.2 <- dCdu2[i.2] * Pn.2;
33    E[i.1, i.2] <- Pn - dCdu.Pn.1 - dCdu.Pn.2;
34  }
35  E2 <- E %*% t(E);
36
37  # Computation of the multiplier bootstrap replicates (random part).
38  Rn_bootstrap = rep(0, N);
39  for (k in 1:N) {
40    xi <- rexp(n);
41    Xi <- xi / mean(xi) - 1;
42    Rn_bootstrap[k] <- t(Xi) %*% E2 %*% Xi / (n^2);
43  }
44
45  # Output of the results
46  structure(class = "htest", list(
47    statistic = c(statistic = Rn),
48    p.value = (sum(Rn_bootstrap > Rn) + 0.5) / (N + 1)
49  ));
50}

```

D.2.2 Implementierung (Rangbasierter Test auf Austauschbarkeit mit Approximation nach **Genest2012**). Die Implementierung in R ist angelehnt an der Implementierung D.2.1 für den Test auf Radial-Symmetrie.

```

1 exchTest_genestNeslehovaEtQuessy <- function(X, ties.method="max", b=1, N
2   =250) {
3   # Calculate the pseudo observations U w.r.t. to the sample X. Measure the
4   # sample size and construct the size-adapted binwidth bn.
5   U <- copula::pobs(X, ties.method = ties.method);
6   n <- nrow(X);
7   bn <- b / sqrt(n);
8
9   # Calculation of the statistic Rn.
10  E <- matrix(nrow = n, ncol = n);
11  for (i.1 in 1:n) { for (i.2 in 1:n) {
12    E[i.1, i.2] <- prod(U[i.1, ] <= U[i.2, 1:2]) - prod(U[i.1, ] <= U[i.2,
13      2:1]);
14  }}
15  An <- sum(E %*% t(E)) / (n^2);
16
17  # Calculation of the deterministic terms.
18  dCdu1 <- rep(0, n);
19  dCdu2 <- rep(0, n);
20
21  for (i in 1:n) {
22    C.upr.1 <- mean((U[, 1] <= (U[i, 1] + bn)) * (U[, 2] <= U[i, 2]));
23    C.lwr.1 <- mean((U[, 1] <= (U[i, 1] - bn)) * (U[, 2] <= U[i, 2]));
24    C.upr.2 <- mean((U[, 1] <= U[i, 1]) * (U[, 2] <= (U[i, 2] + bn)));
25    C.lwr.2 <- mean((U[, 1] <= U[i, 1]) * (U[, 2] <= (U[i, 2] - bn)));
26
27    dCdu1[i] <- (C.upr.1 - C.lwr.1) / (2 * bn);
28    dCdu2[i] <- (C.upr.2 - C.lwr.2) / (2 * bn);
29  }
30
31  E <- matrix(nrow = n, ncol = n);
32  for (i.1 in 1:n) { for (i.2 in 1:n) {
33    Pn <- prod(U[i.1, ] <= U[i.2, 1:2]) - prod(U[i.1, ] <= U[i.2, 2:1]);
34    Pn.1 <- (U[i.1, 1] <= U[i.2, 1]) - (U[i.1, 2] <= U[i.2, 1]);
35    Pn.2 <- (U[i.1, 2] <= U[i.2, 2]) - (U[i.1, 1] <= U[i.2, 2]);
36    dCdu.Pn.1 <- dCdu1[i.2] * Pn.1;
37    dCdu.Pn.2 <- dCdu2[i.2] * Pn.2;
38    E[i.1, i.2] <- Pn - dCdu.Pn.1 - dCdu.Pn.2;
39  }}
40
41  E2 <- E %*% t(E);
42
43  # Computation of the multiplier bootstrap replicates (random part)
44  An_bootstrap = rep(0, N);
45  for (k in 1:N) {
46    xi <- rexp(n);
47    Xi <- xi / mean(xi) - 1;
48    An_bootstrap[k] <- t(Xi) %*% E2 %*% Xi / (n^2);
49  }
50
51  # Output of the results
52  structure(class = "htest", list(
53    statistic = c(statistic = An),
54    p.value = (sum(An_bootstrap > An) + 0.5) / (N + 1)
55  )));
56}

```

D.3 Methoden zur Anpassung der Daten und modifizierte Tests

D.3.1 Implementierung (Stichprobenreduktion bzgl. vollständiger Bindungen). Die Implementierung in R nutzt aus, dass die Funktion `unique()` zeilenweise Abgleiche durchführt, um die einzigartigen Elemente / Zeilen zu identifizieren. Ist der Datensatz so konstruiert, dass die Koordinaten in den Spalten abgebildet werden und die einzelnen Stichprobenelemente die Zeilen sind, so stimmt die Identifikation einzigartiger Zeilen mit der Stichprobenreduktion bzgl. vollständiger Bindungen überein.

```
1 fulltie_samplereduction <- function(x) {
2   return(unique(x))
3 }
```

D.3.2 Implementierung (Stichprobenreduktion über einen Greedy-Algorithmus bzgl. Koordinatenanzahl und Bindungsgrößen, bivariater Fall). Diese Implementierung benutzt den Parameter `N`, um die maximale Anzahl der Elemente, die gelöscht werden sollen, zu steuern. In der Simulationsstudie wurde z. B. `N <- ceiling(nrow(x) * 0.25)` gesetzt.

```
1 greedy_samplereduction_BIV <- function(x, N) {
2   # Initialization
3   n <- nrow(x);
4   b <- g <- c();
5
6   # Calculation of initial weights
7   for (i in 1:n) {
8     b <- c(b, (sum(x[, 1] == x[i, 1]) > 1) + (sum(x[, 2] == x[i, 2]) > 1));
9     g <- c(g, max(sum(x[, 1] == x[i, 1]), sum(x[, 2] == x[i, 2])));
10   }
11   w <- b / g;
12
13   # Order of weights that are greater than zero
14   index <- order(w, decreasing = TRUE);
15   index <- index[w[index] > 0];
16
17   # Select a representative per tie in the greedy selection by removing it
18   # from the index set, decreasing in b. Check for full ties, partial ties.
19   I <- J <- c();
20   for (i in index) { if (!i %in% union(I, J)) {
21
22     B.0 <- B.1 <- B.2 <- c();
23     i.0 <- i.1 <- i.2 <- NA;
24
25     B.0 <- index[x[index, 1] == x[i, 1] & x[index, 2] == x[i, 2]];
26     B.0 <- B.0[!B.0 %in% union(I, J)];
27
28     B.1 <- index[x[index, 1] == x[i, 1]];
29     B.1 <- B.1[!B.1 %in% union(union(I, J), B.0)];
30
31     B.2 <- index[x[index, 2] == x[i, 2]];
32     B.2 <- B.2[!B.2 %in% union(union(I, J), B.0)];
33   }
34 }
```

```

33  if (!any(is.na(B.0)) && !is.na(B.0[1])) {i.0 <- sample(c(B.0, B.0), 1);}
34  if (!any(is.na(B.1)) && !is.na(B.1[1])) {i.1 <- sample(c(B.1, B.1), 1);}
35  if (!any(is.na(B.2)) && !is.na(B.2[1])) {i.2 <- sample(c(B.2, B.2), 1);}
36
37  if (
38    length(J) <= N & N <= length(J) +
39    length(B.0[B.0!=i.0]) + length(B.1[B.1!=i.1]) + length(B.2[B.2!=i.2])
40  ) {
41    if (!is.null(J) & length(J) > 0) {
42      return(x[-J,])
43      stop()
44    } else {
45      return(x)
46      stop()
47    }
48  } else {
49    I <- union(I, c(i.0, i.1, i.2)[!is.na(c(i.0, i.1, i.2))]);
50    J <- union(J, union(B.0[B.0!=i.0],union(B.1[B.1!=i.1],B.2[B.2!=i.2])))
51    J <- J[!is.na(J)]
52  }
53}
54 if (!is.null(J)) {
55  return(x[-J,])
56} else {
57  return(x)
58}
59

```

D.3.3 Implementierung (Jittering mit unabhängigem Rauschen). *Die folgenden zwei Zeilen zeigen das Hinzunehmen eines Zufallsrauschen in die Stichprobe. x sei dabei eine bivariate Stichprobe und rp.jitter die Präzision des Jitterings.*

```

1 noise.PiJitter <- copula::rCopula(nrow(x), copula::indepCopula(2))
2 x.PiJitter <- x + noise.PiJitter * rp.jitter

```

D.3.4 Implementierung (Jittering mit empirisch angepasstem Rauschen). *x sei dabei eine bivariate Stichprobe und rp.jitter die Jittering-Präzision.*

```

1 copulas <- c(copula::normalCopula(0,2), copula::upfhCopula(2))
2 empKTau <- min(max(copula::corKendall(x)[1,2], 0), 1)})
3 mixCopula <- copula::mixCopula(copList, c(empKTau, 1 - empKTau))
4 noise.TauJitter <- copula::rCopula(n, mixCopula) * rp.jitter
5 x.TauJitter <- x + noise.TauJitter * rp.jitter

```

D.3.5 Implementierung (Intervall-zensierte MPLE). Der Code basiert auf dem Quellcode des Pakets *Censorcopula* von Li et al. (2016) unter Anpassung der Notationen und Namensgebung. `censoredMPLE` umfasst den Algorithmus zur Parameterschätzung und entspricht `censorcopula::intervalFitb`. Die Hilfsfunktion `CensorLogLikelihood` entspricht `censorcopula::Newloglik2` und stellt die zensierte Likelihood dar.

```

1 censoredMPLE <- function(
2   copula, method, X, start = NULL, lower = NULL, upper = NULL,
3   optim.control = list(maxit = 1000), estimate.variance = NA,
4   hideWarnings = TRUE, bound.eps = .Machine$double.eps^0.5
5 ) {
6   # Model fitting via maximum pseudo likelihood estimation based on a
7   # censored log pseudo likelihood function CensorLogLikelihood.
8
9   U <- copula::pobs(X);
10
11   # Validation of inputs.
12   stopifnot(is.numeric(d <- ncol(U)), d >= 2);
13   if (copula@dimension != d) {
14     stop("The copula and sample dimensions are not equal.");
15   }
16   if (is.null(start)) {
17     start <- copula:::fitCopStart(copula, U);
18   }
19   if (any(is.na(start))) {
20     stop("The parameter 'start' contains NA values.");
21   }
22
23   if (length(copula@parameters) != length(start)) {
24     stop(paste0(
25       "The vector start (", length(start),") is not feasible. A vector of ",
26       "length ", length(copula@parameters), " for the copula parameter is ",
27       "required."
28     ));
29   }
30
31   # Setup of the optimization control.
32   control <- c(optim.control, fnyscale = -1);
33   control <- control[!vapply(control, is.null, NA)];
34   if (!is.null(optim.control[[1]])) { control <- c(control, optim.control);
35   }
36
37   meth.has.bounds <- method %in% c("Brent", "L-BFGS-B");
38   lower <- ifelse (meth.has.bounds, copula@param.lowbnd + bound.eps, -Inf );
39   upper <- ifelse (meth.has.bounds, copula@param.upbnd - bound.eps, Inf );
40   (if (hideWarnings) { suppressWarnings } else { identity })(  

41     fit <- optim(
42       start, CensorLogLikelihood, lower = lower, upper = upper,
43       method = method, copula = copula, X = X, control = control
44     )
45   )
46   return(fit);
}

```

```

1 CensorLogLikelihood <- function(param, X, copula) {
2   copula@parameters <- param;
3   duplicatesIndex <- function(X) {
4     return <- duplicated(X) | duplicated(X, fromLast=TRUE);
5   }
6   maxRanks <- apply(X, 2, rank, ties.method = "max") / (nrow(X) + 1);
7   minRanks <- apply(X, 2, rank, ties.method = "min") / (nrow(X) + 1);
8   censored <- apply(maxRanks, 2, duplicatesIndex);
9   n <- nrow(X)
10
11 # t = tied => censorship true
12 # f = censorship false
13 tt <- ff <- tf <- ft <- NULL;
14 TT <- FF <- TF <- FT <- 1;
15
16 for (i in 1:n) {
17   c <- censored[i, ];
18   minR <- minRanks[i, ];
19   maxR <- maxRanks[i, ];
20
21   # Check the type of censoring intervals required.
22   if (!c[1] & !c[2]) { ff <- rbind(ff, c(minR, maxR)); }
23   else if (c[1] & c[2]) { tt <- rbind(tt, c(minR, maxR)); }
24   else if (!c[1] & c[2]) { ft <- rbind(ft, c(minR, maxR)); }
25   else if (c[1] & !c[2]) { tf <- rbind(tf, c(minR, maxR)); }
26 }
27
28 # Likelihood terms for the four cases stated above.
29 if (!is.null(tt)) {
30   TT <- copula::pCopula(tt[, 3:4], copula) +
31     copula::pCopula(tt[, 1:2], copula) -
32     copula::pCopula(cbind(tt[, 1], tt[, 4]), copula) -
33     copula::pCopula(cbind(tt[, 3], tt[, 2]), copula);
34 }
35 if (!is.null(ff)) {
36   FF <- copula::dCopula(ff[, 1:2], copula);
37 }
38 if (!is.null(tf)) {
39   TF <- copula:::dCdu(copula, matrix(tf[, 3:4], ncol=2))[, 2] -
40     copula:::dCdu(copula, matrix(tf[, 1:2], ncol=2))[, 2];
41 }
42 if (!is.null(ft)) {
43   FT <- copula:::dCdu(copula, matrix(ft[, 3:4], ncol=2))[, 1] -
44     copula:::dCdu(copula, matrix(ft[, 1:2], ncol=2))[, 1];
45 }
46
47 # Return the censored log pseudo likelihood
48 return((sum(log(TT)) + sum(log(TF)) + sum(log(FT)) + sum(log(FF))))
49 }
```

D.3.6 Implementierung (Intervall-zensierter Test der Anpassungsgüte (am Beispiel für die Nullhypothese bzgl. der Clayton-Copulafamilie)). Für andere Copulafamilien sind die Zeilen 2, 19 und 41 anzupassen.

```

1  gofTest_censored_claytonCopula <- function(X, ties.method="max", N=250,
2      start=NULL) {
3      copula <- copula::claytonCopula();
4      U <- copula::pobs(X);
5      n <- nrow(X);
6
7      F.1.inv <- function(u) {return(quantile(U[, 1], u, type=1, names=FALSE))};
8      F.2.inv <- function(u) {return(quantile(U[, 2], u, type=1, names=FALSE))};
9
10     # First fit for the test statistic calculation
11     theta_n <- copula@parameters;
12     try({theta_n <- censoredMLE(copula=copula, method="BFGS", X, start)$par},
13         silent = TRUE);
14
15     # Stop, if the fit fails.
16     if (any(is.na(theta_n))) {
17         return(structure(class="htest", list(statistic=NA, p.value=NA)));
18         stop();
19     }
20
21     copula.theta_n <- copula::claytonCopula(dim=2, theta_n);
22
23     # Calculate the main test statistic
24     eval.fitCopula <- copula::pCopula(U, copula.theta_n);
25     eval.empCopula <- copula::C.n(U,U);
26
27     Gn <- sum((eval.fitCopula - eval.empCopula) ** 2, na.rm = TRUE);
28
29     # Estimate the p value with bootstrap replicates:
30     Gn_bootstrap = rep(0, N);
31     for (k in 1:N) {
32         U_bootstrap <- copula::rCopula(n, copula.theta_n);
33         X_bootstrap <- U_bootstrap;
34         X_bootstrap[, 1] <- F.1.inv(U_bootstrap[, 1]);
35         X_bootstrap[, 2] <- F.2.inv(U_bootstrap[, 2]);
36
37         theta_n <- copula@parameters;
38         try({ theta_n <- censoredMLE(copula=copula, method="BFGS", X_bootstrap,
39             start)$par }, silent = TRUE);
40
41         if (any(is.na(theta_n))) {
42             Gn_bootstrap[k] <- NA;
43         } else {
44             copula.theta_n <- copula::claytonCopula(dim=2, theta_n);
45             eval.fitCopula <- copula::pCopula(U_bootstrap, copula.theta_n);
46             eval.empCopula <- copula::C.n(U_bootstrap, U_bootstrap);
47             Gn_bootstrap[k] <- sum((eval.fitCopula - eval.empCopula) ** 2, na.rm =
48                 TRUE);
49         }
50
51     # Output of the results
52     structure(class = "htest", list(
53         statistic = c(statistic = Gn),
54         p.value = (sum(Gn_bootstrap > Gn, na.rm = TRUE) + 0.5) / (N + 1)
55     ));
56 }

```

Anhang E

Mathematischer Appendix

E.1 Verteilungsfunktionen

E.1.1 Definition (Multivariate Normalverteilung). *Die Verteilungsfunktion $\Phi_{\mu, \Sigma} : \mathbb{R}^d \rightarrow [0, 1]$ der multivariaten Normalverteilung $\mathcal{N}_d(\mu, \Sigma)$ mit Erwartungsvektor $\mu \in \mathbb{R}^d$ und Kovarianzmatrix $\Sigma \in \mathbb{R}^{d \times d}$ ist definiert als*

$$\Phi_{\mu, \Sigma}(x) := \int_{(-\infty, x]} \phi_{\mu, \Sigma} d\lambda^d = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} \phi_{\mu, \Sigma}(y_1, \dots, y_d) dy_1 \cdots dy_d,$$

$$\phi_{\mu, \Sigma}(x) := \frac{1}{(2\pi)^{d/2} |\det(\Sigma)|^{1/2}} \exp\left(-\frac{(x - \mu)^\top \Sigma^{-1} (x - \mu)}{2}\right)$$

für $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. $\phi_{\mu, \Sigma}$ bezeichne dabei die Dichtefunktion (vgl. McNeil et al. (2015): S. 185, Definition 6.1).

E.1.2 Definition (Multivariate t-Verteilung). *Die Verteilungsfunktion $t_{\mu, \Sigma, \nu} : \mathbb{R}^d \rightarrow [0, 1]$ der multivariaten t-Verteilung $t_d(\mu, \Sigma, \nu)$ mit $\nu > 2$ Freiheitsgraden, dem Erwartungs- bzw. Lagevektor $\mu \in \mathbb{R}^d$ und der Streuungsmatrix Σ bzw. Kovarianzmatrix $(\frac{\nu}{\nu-2}) \Sigma \in \mathbb{R}^{d \times d}$ ist definiert als*

$$t_{\mu, \Sigma, \nu}(x) := \int_{(-\infty, x]} f_{\mu, \Sigma, \nu} d\lambda^d = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f_{\mu, \Sigma, \nu}(y_1, \dots, y_d) dy_1 \cdots dy_d,$$

$$f_{\mu, \Sigma, \nu}(x) := \frac{\Gamma(\frac{d+\nu}{2})}{\Gamma(\frac{\nu}{2}) \cdot (\pi\nu)^{d/2} |\det \Sigma|^{1/2}} \left(1 + \frac{(x - \mu)^\top \Sigma^{-1} (x - \mu)}{\nu}\right)^{-(d+\nu)/2}$$

für $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. $f_{\mu, \Sigma, \nu}$ bezeichne dabei die Dichtefunktion (vgl. McNeil et al. (2015): S. 185, Example 6.7).

E.2 Beweise

Beweis zu Lemma 2.1.4: Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine beliebige Verteilungsfunktion und $F^\leftarrow : (0, 1) \rightarrow \mathbb{R}$ die zugehörige Quantilfunktion.

Es wird zunächst (i.) gezeigt: Seien $\alpha_0, \alpha_1 \in (0, 1)$ mit $\alpha_0 < \alpha_1$ gegeben. Die Isotonie von F impliziert

$$A_0 := \{y \in \mathbb{R} : \alpha_0 \leqq F(y)\} \supseteq \{y \in \mathbb{R} : \alpha_1 \leqq F(y)\} =: A_1.$$

Aus dieser Inklusion wird über die Betrachtung der Infima die Ungleichung

$$F^\leftarrow(\alpha_0) := \inf A_0 \leqq \inf A_1 =: F^\leftarrow(\alpha_1)$$

gefolgert. Somit gilt die Isotonie von F^\leftarrow .

Für die Hinrichtung in (ii.) seien $\alpha \in (0, 1)$ und $x \in \mathbb{R}$ mit $\alpha \leqq F(x)$. Mit Hilfe der Eigenschaft (i.) über die Isotonie von F^\leftarrow gilt

$$F^\leftarrow(\alpha) \leqq F^\leftarrow(F(x)) := \inf \{y \in \mathbb{R} : F(x) \leqq F(y)\} \stackrel{(1)}{\leqq} x,$$

wobei in (1) die Isotonie von F ausgenutzt wurde. Seien nun zur Rückrichtung $\alpha \in (0, 1)$ und $x \in \mathbb{R}$ so, dass $F^\leftarrow(\alpha) \leqq x$. Für alle $\varepsilon > 0$ folgt die Abschätzung

$$\alpha \stackrel{(2)}{\leqq} F(F^\leftarrow(\alpha)) \stackrel{(3)}{\leqq} F(x) \stackrel{(3)}{\leqq} F(x + \varepsilon).$$

In (2) wurde die Definition von $F^\leftarrow(\alpha)$ und in (3) die Isotonie von F bzgl. $x < x + \varepsilon$ genutzt. Da F càdlàg ist und obige Abschätzung für alle $\varepsilon > 0$ gilt, folgt $\alpha \leqq F(x)$. Insgesamt ist damit die Äquivalenz aus (ii.) bewiesen.

Die grundlegenden Abschätzungen in (iii.) und (iv.) folgen direkt aus der Äquivalenz in (ii.). Für die Äquivalenz in (iii.) genügen die folgenden Fälle:

1. Fall: Sei $x \in \mathbb{R}$ mit $x = \inf \{y \in \mathbb{R} : F(x) = F(y)\}$. Dann folgt direkt per definitionem $F^\leftarrow(F(x)) = x$.
2. Fall: Sei $x \in \mathbb{R}$ mit $x > x_0 := \inf \{y \in \mathbb{R} : F(x) = F(y)\}$. Anwenden der Quantilfunktion liefert $F^\leftarrow(F(x)) = F^\leftarrow(F(x_0)) = x_0 < x$.

Es folgt also $x \neq \inf \{y \in \mathbb{R} : F(y) = F(x)\} \Leftrightarrow F^\leftarrow(F(x)) < x$.

Für die Äquivalenzaussage in (iv.) werden zwei weitere Fälle betrachtet:

1. Fall: Sei $\alpha \in \text{Bild}(F)$. Dann existiert ein $x_\alpha \in \mathbb{R}$ mit $F(x_\alpha) = \alpha$ und hierfür gilt die Gleichungskette:

$$\begin{aligned} F(F^\leftarrow(\alpha)) &= F(F^\leftarrow(F(x_\alpha))) \\ &= F(\inf \{y \in \mathbb{R} : F(y) \geqq F(x_\alpha)\}) \\ &= F(x_\alpha) = \alpha. \end{aligned}$$

2. Fall: Sei $\alpha \notin \text{Bild}(F)$. Für $\alpha_1 = \inf(\text{Bild}(F) \cap (\alpha, 1)) = F(F^\leftarrow(\alpha))$ folgt trivial, dass $\alpha_1 > \alpha$ und somit $F(F^\leftarrow(\alpha)) > \alpha$.

Insgesamt folgt die Äquivalenzaussage in (v.). \square

Beweis von Lemma 2.5.1. Die multivariate Dichtefunktion f ist über die bedingten Dichten $f_1, f_{2|1}, \dots, f_{d|1, \dots, d-1}$ der Marginale faktorisierbar als

$$f(x_1, \dots, x_d) = f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdots f_{d|1, \dots, d-1}(x_d|x_1, \dots, x_{d-1}). \quad (5-1)$$

Über die Formel (2-21) ist die bivariate Dichte bzgl. der i -ten und j -ten Koordinaten darstellbar als $f(x_i, x_j) = c_{i,j}(F_i(x_i), F_j(x_j)) \cdot f_i(x_i) \cdot f_j(x_j)$.

Für alle $t \in \{2, \dots, d\}$ gilt dann die Rekursion

$$\begin{aligned} f_{t|1, \dots, t-1}(x_t|x_1, \dots, x_{t-1}) &= c_{t-1, t|1, \dots, t-2} \cdot f_{t|1, \dots, t-2}(x_t|x_1, \dots, x_{t-2}) \\ &= \dots = \left[\prod_{s=1}^{t-1} c_{t-s, t|1, \dots, t-s-1} \right] \cdot f_t(x_t) \end{aligned} \quad (5-2)$$

und die Aussage des Lemmas folgt direkt aus der Gleichung

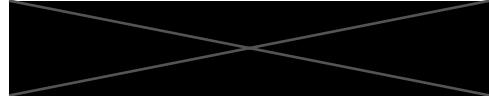
$$\begin{aligned} f(x_1, \dots, x_d) &\stackrel{\textcircled{1}}{=} \left[\prod_{t=2}^d \prod_{s=1}^{t-1} c_{t-s, t|1, \dots, t-s-1} \right] \left[\prod_{k=1}^d f_k(x_k) \right] \\ &\stackrel{\textcircled{2}}{=} \left[\prod_{j=1}^{d-1} \prod_{k=1}^{j-1} c_{j, j+k|1, \dots, j-1} \right] \left[\prod_{j=1}^d f_j(x_j) \right]. \end{aligned}$$

In $\textcircled{1}$ wurde (5-2) in (5-1) eingesetzt und in $\textcircled{2}$ wurden die Indizes $j := t - s$ und $j + k := t$ neu gesetzt (vgl. Czado (2019): S. 89-90). \square

Erklärung

Hiermit versichere ich, dass ich die Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Düsseldorf den 29. November 2021



(Kevin Tuz)