Essays on Risk Management for Agricultural Commodity Futures Market

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Ying Wang, M.A., M.A.S.

Graduate Program in Agricultural, Environmental and Development Economics

The Ohio State University

2016

Dissertation Committee:

Professor Matthew Roberts, Advisor

Professor Stanley Thompson

Professor Mario Miranda
Abstract

Funding risk, which caused the $1.3 billion derivatives-related loss at MG Refining & Marketing, Inc. in 1993, has long been overlooked in the risk management literature. The key to understanding funding risk is that, as futures hedging practice requires substantial infusions of cash to meet variation margin calls, the maximum margin required may occur well before the expiration of the futures contract, but must be met in order to maintain the futures positions. This paper approaches the question of how to properly measure the “funding risk” of commodity futures positions by estimating a CD-Vine copula model for the dependence of corn, soybean and wheat futures at multiple forecast intervals, using Harrison’s method and the Extreme Value Theory to calibrate the distribution of the maximum. This is the first attempt in the literature to model the extreme prices of futures contracts over a given time period in an agricultural commodity portfolio context. The adoption of the recently-developed CD-Vine copula model allows one to model the dependence structure in a more flexible manner than the previous standard multivariate models based on Gaussian or Student’s $t$ distributions.

Witnessing the recent surge in price and volatility of agricultural commodity markets, it cannot be emphasized enough how important it is to assess the probability of rare and extreme price movements in the risk management of agricultural commodity futures. Similar to other financial time series, agricultural commodity futures exhibit the characteristics of time-varying volatility and fat tails. In this chapter, we employ the
McNeil and Frey’s two step approach and conditional Extreme Value Theory to estimate Value-at-Risk (VaR) and Expected Shortfall (ES) for long and short positions in the agricultural commodity futures market at multiple significance levels, and compare this approach to conventional multivariate Normal or Student’s $t$ distribution based models, Historical Simulation, RiskMetrics etc. The backtesting demonstrates that this GARCH-EVT approach provides a significant improvement over the widely used Normal and Student’s $t$ distribution based VaR and ES models, which tend to underestimate the true risk and fail to provide accurate VaR estimates that are statistically no different from the corresponding significance level. To capture the tail dependence and properly estimate portfolio VaR, copula models are introduced to estimate a portfolio measure of risk in a multi-commodity setting. This has broad applications, for instance, for an agricultural commodity end-user that is purchasing corn, wheat and soybeans simultaneously. The conventional approach to this problem is to use a multivariate GARCH (a.k.a. MGARCH) model to estimate the conditional covariance between the futures prices. However, the typical MGARCH model approach inevitably suffers from being unduly restrictive because of the classical joint multivariate Gaussian assumption, despite the empirical evidence against elliptical distributions in commodity price returns. Also, from the perspective of computational efforts, the number of parameters to be estimated in the MGARCH specification often increases rapidly, stemming from the high-dimensional nature of the problem.
Dedication

This is dedicated to my parents Shutang Wang and Yunqiu Feng, my husband Scott Murdoch, and my daughter Chloe Murdoch.
Acknowledgments

I would like to sincerely thank all those people who have helped and inspired me during my doctoral dissertation research. This doctoral dissertation would not have been possible without the support of many kind people, to only some of whom it is possible to include thanks here.

I would like to express my deepest gratitude to my advisor Dr. Matthew Roberts for his excellent guidance, encouragement, patience, and high expectations throughout my doctoral studies. This dissertation would not have been possible without his generous contributions of time, expertise, resources, and references. I am also deeply indebted to Dr. Stanley Thompson and Dr. Mario Miranda for their generous help and insightful guidance as my dissertation committee members. My research has benefitted greatly from their vast knowledge base and unique perspectives.

I am also appreciative of my fellow graduate students in the department, particularly the members of my first year cohort. Special thanks go to Dr. Carlianne Patrick, Dr. Greg Howard, Dr. Asif Ali, Dr. Isai Guizar-Mateos, and Dr. Nicholas Dadzie for all the thought-provoking discussions, warm encouragement and support. I would also like to thank Dr. Qiping Xu, Dr. Haoxuan Tong, Dr. Shanshan He, Dr. Ruohan Wu, Dr. Ying An,
Dr. Stacie Wang, Dr. Li Ma, and Dr. Zhenzhen Liu for sharing the ups and downs, laughs and tears in our own pursuit of Ph.D. in diverse fields.

Being a Ph.D. candidate while pursuing a full-time career in finance would not have been possible without the help of many colleagues and mentors. I would like to thank Dr. SuBang Choe and Dr. Suzhou Huang at Ford Motor Credit Company for showing me the way to apply quantitative methods in solving the real world problems. I would especially like to thank Paul Sassieni and James Kane at Northern Trust for affording me the opportunity to work on important projects applying my Ph.D. training, while creating a friendly work environment with their witty words and constant encouragement.

I would also like to acknowledge the academic, financial, and technical support of the Department of Agricultural, Environmental and Development Economics at The Ohio State University. In particular, I would like to thank Dr. Stanley Thompson, Dr. Mario Miranda, and Susan Miller for all their efforts on my behalf.

Last, but certainly not least, I would like to extend my deepest gratitude to my family. None of this would have been possible without the unwavering love and support from my parents Shutang Wang and Yunqiu Feng, my husband Scott Murdoch, and my daughter Chloe Murdoch. To my husband Scott Murdoch, I am eternally grateful for your unceasing encouragement of my Ph.D. pursuit, your emotional support under time of stress, and your kindness in bearing the burden of household chores without complaint. To my daughter Chloe Murdoch, I am grateful for the “superpower” bestowed on me when you came into this world. It is with your love that I become a stronger and better self.
Vita

2008..................................................B.S. Quantitative Finance,
Zhejiang University

2010..................................................M.A. Economics,
The Ohio State University

2011..................................................Master of Applied Statistics,
The Ohio State University

2008 to 2012 .....................................Graduate Teaching/Research Associate,
The Ohio State University

2012 to 2014 .....................................Intermediate Quantitative Analyst,
AEGON USA Investment Management

2014 to present ....................................Second Vice President,
Northern Trust

Fields of Study

Major Field: Agricultural, Environmental and Development Economics
Table of Contents

Abstract.................................................................................................................................................. ii

Dedication................................................................................................................................................ iv

Acknowledgments.................................................................................................................................... v

Vita........................................................................................................................................................... vii

List of Tables ............................................................................................................................................ xi

List of Figures ............................................................................................................................................ xii

Chapter 1: Calibrating the probability Distribution of the Maximum Futures Movements in a Multi-Commodity Setting .............................................................................................................. 1

1.1 Introduction ....................................................................................................................................... 1

1.2 The Model ......................................................................................................................................... 5

1.2.1 Distribution of the Maximum ........................................................................................................ 5

1.2.2 Futures Market Margins ................................................................................................................. 13

1.2.3 Copula and CD-Vine ..................................................................................................................... 15

1.3 Empirical Evaluation and Results ..................................................................................................... 26

1.3.1 Data Description and Preliminary Tests ...................................................................................... 28

1.3.2 Comparison of Harrison's Method and EVT ................................................................................ 32
List of Tables

Table 1. Summary Characteristics of Bivariate Copula Families ........................................ 20
Table 2. Summary Descriptive Statistics ............................................................................. 31
Table 3. Proportion of Exceedance using Weekly Data ..................................................... 33
Table 4. Proportion of Exceedance using Monthly Data ..................................................... 35
Table 5. Out-of-Sample Log-Likelihood Function (OOSLLF) Values ................................. 36
Table 6. Proportion of Exceedance using Weekly Data with CD-Vine Copula ................. 43
Table 7. Proportion of Exceedance using Monthly Data with CD-Vine Copula ............... 43
Table 8. Contingency Table for Markov Independence Test ............................................. 78
Table 9. Summary Descriptive Statistics ............................................................................. 87
Table 10. (G)ARCH Effect Test ...................................................................................... 88
Table 11. Summary Results for the Marginal Distribution ............................................... 90
Table 12. Conditional Correlation Measures Matrix ....................................................... 93
Table 13. Summary of Observations in Each Tail ............................................................ 95
Table 14. VaR Violation Ratio of Competing Risk Models ............................................... 101
Table 15. Bootstrap Unconditional/Conditional Coverage Test ...................................... 102
Table 16. Bootstrap Backtesting Results of Expected Shortfall ...................................... 103
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>The Fréchet, Gumbel, and Weibull Extreme Value Distributions</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Contour Plots Illustration of Common Copula Families</td>
<td>21</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Illustration of Five-dimensional Canonical Vine Trees with Edge Indices</td>
<td>23</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Illustration of Five-dimensional D-Vine Trees with Edge Indices</td>
<td>24</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Daily Prices, Returns, and Squared Returns</td>
<td>29</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Correlation Plot with Weekly Harrison’s Marginal</td>
<td>40</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Correlation Plot with Monthly Harrison’s Marginal</td>
<td>40</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Correlation Plot with Weekly EVT Marginal</td>
<td>41</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Correlation Plot with Monthly EVT Marginal</td>
<td>41</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Illustration of C-Vine Tree Plot of Weekly Harrison’s Marginal</td>
<td>42</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Illustration of D-Vine Tree Plot of Weekly Harrison’s Marginal</td>
<td>42</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Illustration of C-Vine Tree Plot of Weekly EVT Marginal</td>
<td>42</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Illustration of D-Vine Tree Plot of Weekly EVT Marginal</td>
<td>42</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Generalized Pareto Distributions- Cumulative Probability Functions</td>
<td>61</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Daily Prices, Returns, and Squared Returns</td>
<td>85</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Normal QQ-Plot Before vs. After GARCH Filter</td>
<td>92</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Correlation Plot of GARCH-filtered Residuals</td>
<td>94</td>
</tr>
</tbody>
</table>
Figure 18. C-Vine Tree Plot of Uniformed GARCH-filtered Residuals .......................... 97
Figure 19. D-Vine Tree Plot of Uniformed GARCH-filtered Residuals .......................... 98
Chapter 1: Calibrating the probability Distribution of the Maximum Futures Movements in a Multi-Commodity Setting

1.1 Introduction

At year-end 1993, a derivatives-related loss of $1.3 billion - a world-record according to press accounts at that time – incurred by MG Refining & Marketing, Inc. (MGRM) shook the financial world. MGRM was a U.S. subsidiary of the German industrial conglomerate Metallgesellschaft AG (MG AG), whose derivatives were part of a marketing program, under which it offered clients fixed prices for up to ten years on gasoline, heating oil, and diesel fuel purchased from MGRM. The firm hedged its resulting exposure to the spot price fluctuations with futures contracts. Because futures contracts are marked to market each day, cash payments must be made at the end of the trading day to meet variation margin payments when futures prices fluctuate. After several consecutive months of falling prices in the autumn of 1993, MGRM’s German parent reacted to margin calls in excess of $900 million (Hanley(1994)) by liquidating the hedge. The supervisory board’s decision to unwind the futures leg proved unfortunate on several counts, turning “paper losses” into realized losses, sending a distress signal to MGRM’s over-the-counter (OTC) derivatives counterparties, and leaving MGRM
exposed to rising prices on its remaining fixed-price contracts (Falloon (1994); Culp and Miller (1995)).

Funding risk is a natural suspect for MGRM’s problems because futures hedging practice requires substantial infusions of cash to meet variation margin calls amid increases and decreases in commodity prices and volatility of the broader financial market. In 1993, the prices of oil and oil products plunged after OPEC failed to reach agreement on production quotas, resulting in substantial margin payments due from MGRM to the NYMEX\(^1\). By December 1993, rumors of MGRM’s possible cash flow problem surfaced. As that information spread, the NYMEX demanded a “super-margin” from MGRM to protect its clearing house, which further exacerbated the liquidity crisis. Had MGRM not liquidated its futures positions prematurely because of the transitory cash drain, the transaction would have realized a profit of more than $1 billion\(^2\) (Bollen and Whaley (1998)).

MGRM’s large derivative loss is not just an isolated case, but part of a larger funding issue common to the broader commodity futures market. On April 22\(^{nd}\), 2008, the Commodity Futures Trading Commission (CFTC) held hearings on the agricultural futures market\(^3\). Tom Buis of the National Farmers Union stated that they started raising concerns in March 2008 after feedback from members that they weren’t able to roll over forward contracts due to their elevators having credit issues. Also, evidence began to

---

\(^1\)On March 17, 2008, the New York Mercantile Exchange (NYMEX) accepted an acquisition offer from CME Group, the parent of the Chicago Mercantile Exchange. The acquisition was formally completed on August 22, 2008, and the NYMEX systems were fully integrated by September 30, 2009.

\(^2\)This comes from a statistical simulation of MGRM’s future positions at that time by Bollen and Whaley (1998). However, the authors also noted that MGRM might have been required to suffer a negative cash flow of as much as $1 billion to sustain the position.

\(^3\) The agenda of the hearing is available at [http://www.cftc.gov/PressRoom/PressReleases/pr5484-08](http://www.cftc.gov/PressRoom/PressReleases/pr5484-08).
mount that the confidence of the physical participants in the commodity market had been
shaken in the previous few years. Additionally, representatives of the Federal Reserve
Bank of Kansas City began a discussion on bank elevator financing in their region. The
banks had not stepped up their credit standards, but upped their monitoring and scrutiny.
The credit crunch had little direct effect on the elevators, but borrowing money would
become much more tiresome.

However, the attempt to accurately measure the “funding risk”, i.e. a form of
liquidity risk, defined as “the inability to meet cash flow obligations at an acceptable
price as they become due”\(^4\), has long been overlooked in the risk management literature.
Traditional risk management tools, such as Value-at-Risk (VaR) (Linsmeier and Pearson
(2000)), or price distributions derived from options prices, are of little use in estimating
funding risk. Previous risk estimates are based upon the terminal distribution of prices,
i.e., the realized portfolio value in the terminal time period. The key to understanding the
risk of financing is that the maximum margin required may occur well before the
expiration of the futures contract, sometimes months or even years earlier, but must be
met in order to maintain the futures positions. Therefore, simply using the cumulative
distribution function (CDF) recovered from options prices is not sufficient. Instead, a
properly calibrated distribution for the extreme price reached between any moment of
time \(t\) after the inception of the futures contract and the terminal period \(T\) is necessary.
Given this distribution, a distribution of maximum margin requirements can then be
constructed, from which an inverse CDF can be used to create a VaR type measure of

Handbook (Jan., 1997)
margin exposure that can later be used by participants, by exchanges and clearing houses, or by regulators in the futures market to estimate the margin finance that may be needed under various scenarios.

In this paper, we approach the question of how to properly measure the “funding risk” of commodity futures positions by estimating a CD-Vine copula model (Bedford and Cooke (2001); Bedford and Cooke (2002); Kurowicka and Cooke (2006); Aas et al. (2009); Guégan and Maugis (2010)) for the dependence of corn, soybean and wheat futures at multiple forecast intervals. The contribution is three-fold: First, to the best of my knowledge, this is the first attempt in the literature to model the extreme prices of futures contracts over a given time period in a multi-commodity context, a leap toward making this approach feasible for more realistic applications. Second, the adoption of the copula model allows one to model the dependence structure in a more flexible manner than the previous standard multivariate models based on Gaussian or Student's t distributions (Nelson 2006). Also, the use of copula modeling makes it possible to separate the dependence model from the marginal distributions (Pelletier 2006). Third, the CD-Vine copula implemented in this paper is a novel type of copula structure only recently introduced into the finance literature that allows for general types of dependence. While modeling a bivariate dependence structure doesn’t raise concerns in regards to a lack of flexibility (Joe, Li, and Nikoloulopoulos (2010)), it does present problems when generalizing to an arbitrary number of dimensions, given that the choice of copulas is usually limited to the standard multivariate copula functions such as the multivariate Gaussian or the Student's t, as well as Archimedean copulas. These standard multivariate
copulas lack the flexibility to accurately model the dependence structure in higher dimensions, even under reasonable generalizations. They typically become rather intricate in structure, giving rise to other limitations such as parameter restrictions. CD-Vine copulas allow one to overcome these limitations (Rodriguez 2007).

The organization of the paper is as follows. Section II formally demonstrates the theoretical model which consists of the following: Harrison’s method and Extreme Value Theory in calibrating the distribution of the maximum, as well as that of the future’s margin level, the copula method and the CD-Vine dependence modeling structure. Section III empirically estimates the probability distribution of the extreme price movements of different agricultural commodities, controlling for the dependence structure with the CD-Vine copula, with the univariate version as benchmark. Daily price data for corn, soybean, and wheat futures traded on futures exchanges during the period are analyzed. Section IV provides summarizing comments and concluding remarks.

1.2 The Model

1.2.1 Distribution of the Maximum

Given a process $X_t$ with observations over the time period $0, \ldots, t$, define a process $M_t \equiv \sup \{X_s, 0 \leq s \leq t\}$ that denotes the maximum level of $X_t$ over the time interval $0, \ldots, t$. Two approaches to estimate $M_t$ are utilized in this study: Harrison (1985) derives a closed-form distribution for $M_t$ under the assumption that $X_t$ is a Brownian motion, requiring a volatility estimate for the Brownian motion; alternatively, one may use the Extreme Value Theory (EVT) as in de Oliveira (1983) to empirically estimate the
distribution of the maximum observation of the process $X_t$ over previous intervals of length $0, ..., t$. The neat closed-form solution offered by Harrison’s method facilitates the dependence structure modeling in the CD-Vine copula stage. Additionally, the empirical evaluation by Roberts (2008) concludes that Harrison’s method, by virtue of simplicity, performs at least as well if not better than the extreme value method. Therefore, in this paper, we propose to expand on the analysis in Roberts (2008) to a multi-commodity portfolio with the CD-Vine copula method, using both Harrison’s method and EVT to estimate the distribution of the maximum/minimum at weekly and monthly frequency as the marginal distribution for the copula.

1.2.1.1 Harrison’s Method

According to Harrison, given a Wiener process with drift $\mu$ and infinitesimal variance $\sigma^2$, $X_t = \mu t + \sigma W_t$, where the standard Wiener process $W_t$ is characterized by three properties (Durrett (1996)):

1. $W_0 = 0$;
2. The function $t \rightarrow W_t$ is almost surely continuous;
3. $W_t$ has independent increments with $W_t - W_s \sim N(0, t - s)$ for $0 \leq s < t$, where $N(\mu, \sigma^2)$ denotes the normal distribution with expected value $\mu$ and variance $\sigma^2$.

The common practice in mathematical finance to model prices of financial instruments, such as stocks or commodity futures, are using the stochastic process Geometric Brownian motion (Black and Scholes (1973); Black (1976)). Geometric Brownian
motion, a.k.a exponential Brownian motion, is a continuous-time stochastic process in which the logarithm of the random variable follows a Brownian motion, or Wiener process. Without loss of generality, we hereby derive the distribution of the maximum under the assumption of Brownian motion instead of Geometric Brownian motion, because the price data can be simply transformed into its log form before the implementation of the Harrison’s method.

Define a process $M_t \equiv \sup \{X_s, 0 \leq s \leq t\}$ that denotes the maximum level of $X_t$ over the time interval $0, ..., t$. Without loss of generality, the following mathematical derivation is kept in its 2-dimensional case. The joint distribution function can then be defined as $F_t(x, M_t) \equiv P[X_t \leq x, M_t \leq y]$. In order to ease solution, $X_0 = 0$ is assumed. Therefore, the joint distribution is well-defined only when both conditions are satisfied, namely $y \geq 0$ and $x \leq y$.

For the sake of simplicity, assume that $\mu = 0$ and $\sigma = 1$, i.e. $X_t$ is reduced to the standard Wiener process $W_t$. Then

$$F_t(x, y) = P[X_t \leq x] - P[X_t \leq x, M_t > y] = \Phi \left( \frac{x}{\sqrt{t}} \right) - P[X_t \leq x, M_t > y]$$ (1)

where $\Phi(.)$ is the standard normal cumulative distribution function. Harrison (1985) provides a heuristic\(^5\) explanation of the decomposition of $P[X_t \leq x, M_t > y]$ using the Reflection Principle:

Therefore, when $X_0 = 0$, $\mu = 0$, and $\sigma = 1$,

$$F_t(X, M) \equiv P[X_t \leq x, M_t \leq y] = \Phi \left( \frac{x}{\sqrt{t}} \right) - \Phi \left( \frac{x-2y}{\sqrt{t}} \right)$$ (3)

\(^5\) Along with a rigorous proof in Harrison (1985), see page 8-9.
The generalization to arbitrary \((\mu, \sigma)\) combination is more complex; hence the net result is presented as following, with the mathematical derivation subsumed\(^6\).

\[
F_t(x, y) = \Phi\left(\frac{x-\mu t}{\sigma \sqrt{t}}\right) - e^{2\mu y/\sqrt{t}} \Phi\left(\frac{x-2y-\mu t}{\sigma \sqrt{t}}\right) \quad (4)
\]

Finally, if \(x = y\), i.e. \(X_t\) is permitted to take any value less than \(M_t\) at \(t\), then:

\[
F_t(x, y) \equiv P\{X_t \leq x, M_t \leq y\} = P\{X_t \leq y, M_t \leq y\} = P\{M_t \leq y\} = \Phi\left(\frac{y-\mu t}{\sigma \sqrt{t}}\right) - e^{2\mu y/\sqrt{t}} \Phi\left(\frac{-y-\mu t}{\sigma \sqrt{t}}\right) \quad (5)
\]

which gives a cumulative distribution function for \(M_t\)\(^7\), the maximum value of \(X_t\) over the time span \(0, ..., t\).

1.2.1.2 Extreme Value Theory (EVT)

Previous research on modeling extreme values of commodity futures prices sought to estimate the maximum clearing house exposure to commodity positions. Longin (1999) and Cotter (2001) use Extreme Value Theory (EVT) to estimate the maximum futures price change over one trading day or a few trading days and, therefore, derive the margin requirements accordingly. Extreme Value Theory originates from the book *Statistics of Extremes* by Gumbel (1958). Extremes are defined as the highest observation (the maximum) and the lowest observation (the minimum) over a given time period. Extreme

\(^6\)For further reading please refer to Harrison (1985).

\(^7\)Note that regarding the setting of margins, the calculated thresholds are actually too large, as margins are only charged against end of day prices, not the highest intraday prices which is calculated according to Harrison’s method. See Wilmott, P. (2000) page 278-280 for a discussion in the context of look-back option pricing.
Value Theory provides the theoretical ground for the estimation of the exact and asymptotic distribution of the extremes of a random process within a given time period. One significant feature of this theory is that the form of the distribution of the extremes is largely independent of the distribution of the underlying process. That is, different underlying processes lead to the same form of the distribution of extremes; the distributions of extremes derived from different underlying processes are differentiated by the value of the parameters of the distributions of extremes only. Hence the extreme value approach for the funding risk problem exploits the generality of this result.

Futures prices, \( p_t \), are assumed to follow a random walk. Over a pre-defined time period \( t \), such as a week or a month, the maximum and minimum prices attained are denoted as \( p_t^+ \) and \( p_t^- \), respectively. As shown in Gumbel (1958), if the variables \( p_0, p_1, p_2, \ldots, p_n \) are independent and identically distributed (i.i.d), then the exact distribution of \( p_t^- \) is given by eq. (6):

\[
F(p_t^-) = 1 - [1 - F(p_t)]^n
\]  

(6)

Similarly, the exact distribution of \( p_t^+ \) is given by eq. (7):

\[
F(p_t^+) = [F(p_t)]^n
\]  

(7)

where \( F(p_t) \) is the cumulative distribution function of the underlying random variable, futures prices \( p_t \).

The distributions \( F(p_t^-) \) and \( F(p_t^+) \) depend mainly on the properties of \( F(p_t) \) for large negative and large positive values of \( p_t \). Indeed, for small absolute values of \( p_t \), the influence of \( F(p_t) \) decreases rapidly with \( n \). Hence, the most important information about the extremes is contained in the tails of the distribution. Therefore, the limiting
distributions of the extremes $F(p_t^-)$ and $F(p_t^+)$ obtained by letting $n$ tend to infinity are degenerate.

The exact results for the distribution of extremes are not, however, especially interesting. In practice, the distribution of the underlying variable is not precisely known and, therefore, if this distribution is unknown, neither is the exact distribution of the extremes. For this reason, the asymptotic behavior of the minimum $p_t^-$ and of the maximum $p_t^+$ is studied. As shown by de Oliveira (1983)\(^8\), if $p_t^+$ and $p_t^-$ are normalized by a location parameter $\beta_n^+/\beta_n^-$ and a scale parameter $\alpha_n^+/\alpha_n^-$ (assumed to be positive), and $n$, the number of time periods, is sufficiently large, then there exists a non-degenerate asymptotic distribution of the extreme values, known as the Generalized Extreme Value (GEV) distribution. The Generalized Extreme Value distribution is a generalization of the Fréchet, Gumbel, and Weibull distributions\(^9\), determined by a tail index parameter $\tau$ of the limiting distribution of the standardized variable. The tail index reflects the fatness of the distribution (that is, the weight of the tails), whereas the parameters of scale $\alpha_n^+/\alpha_n^-$ and of location $\beta_n^+/\beta_n^-$ represent the dispersion and the average of the extremes, respectively. The Fréchet distribution is the limiting distribution of the extremes of fat-tailed distributions (such as Student's $t$, Cauchy, or other stable distributions other than the Normal distribution), whereas Gumbel and Weibull are limiting distributions for the extremes of Normal and bounded distributions, respectively. Figure 1 illustrates the three

\(^8\)de Oliveira (1983) argues: “As, in general, we deal with sufficiently large samples, it is natural and in general sufficient for practical uses to find limiting distributions for the maximum or the minimum conveniently reduced and use them.”

\(^9\)The Gumbel distribution can be regarded as a transitional limiting form between the Fréchet and the Weibull distributions.
types of extreme value distribution. Figure 1 gives the detail of the distribution tails: The Fréchet distribution is fat-tailed as its tail is slowly decreasing; the Gumbel distribution is thin-tailed as its tail is rapidly decreasing; and the Weibull has no tail—after a certain point there are no extremes.

The top figure in Figure 1 represents the extreme value distributions: the Fréchet distribution, the Gumbel distribution, and the Weibull distribution. The bottom figure in Figure 1 represents in detail the tails of the three types of extreme value distribution. The distributions represented in both figures are standardized extreme value distributions $(\alpha_n = 1, \beta_n = 0)$ with tail index values equal to -0.5 for the Fréchet case, 0 for the Gumbel case and 0.5 for the Weibull case.
Figure 1. The Fréchet, Gumbel, and Weibull Extreme Value Distributions
1.2.2 Futures Market Margins

To minimize credit risk to the exchange, futures clearinghouses require market participants to post a performance bond, commonly referred to as a margin\textsuperscript{10}, typically 5%-15% of the contract's notional value. Due to the extreme leverage available in the futures market, margins serve as a cushion for clearinghouses amid market turbulence. However, for those with a considerable amount of outstanding futures contracts, variation margin requirements present potential funding risks. For instance, after entering into a long-term contract to purchase or sell grain, an offsetting futures position can be established to mitigate the risk of price movements between contract establishment and closure. Profits and losses of the physical and futures positions should largely offset when the physical contract is executed. However, in the interim, the hedger may face large liquidity needs in order to finance margin calls, as in the case of MGRM.

As implemented by the Chicago Board of Trade (CBOT)\textsuperscript{11}, Chicago Mercantile Exchange (CME), and most other exchanges, margins have two components: initial margin, and variation margin. Initial margin is required at the time when the futures position is established. Variation margin is required if the futures position loses money, in order to maintain an adequate guarantee against non-performance. The level of variation margin required is that which keeps the sum of the initial margin contribution

\textsuperscript{10}Margin requirements are waived or reduced in some cases for hedgers who have physical ownership of the covered commodity or spread traders who have offsetting contracts balancing the position.

\textsuperscript{11}On July 12\textsuperscript{th}, 2007, the CBOT merged with the Chicago Mercantile Exchange (CME) to form the CME Group, a CME/Chicago Board of Trade Company. CBOT and three other exchanges (CME, NYMEX, and COMEX) now operate as designated contract markets (DCM) of the CME Group. See also http://www.chicagotribune.com/news/nationworld/chi-0703160138mar16_0,5630269.story. However, since CBOT still operates its exchange activity under its own regulations, in this paper I hereby still refer to CBOT as the data source rather than the CME group.
(IM), marked-to-market gains or losses (MtM), and variation margin contributions (VM) above a required maintenance margin level (MM), i.e. to be in good standing a margin account must satisfy $IM + MtM + VM \geq MM$. At the CBOT, initial and maintenance margins are equal for hedgers; therefore, for these positions, $VM \geq -MtM$.

That is, to maintain the open position, hedgers are required to post a variation margin daily to cancel out the marked-to-market movement at the minimum. Also, $MtM$ can further be decomposed into $nt(p_t - p_0)$, where $n$ is the size of the net position in futures contracts, $\tau$ is the value of each futures tick, $p_t$ is the time $t$ price level and $p_0$ is the initial price level. Normalizing $p_0 = 0$ and $\tau = 1$ yields $VM \geq -np_t$.

Combining the above results with equation (5), the distributions of maximum margin can be derived. Again, futures prices are martingales, and are thus driftless, i.e. $\mu = 0$. The net result is that for a long future position ($n = 1$),

$$P\{\max VM \geq A\} = 1 - \Phi\left(\frac{A}{\sigma \sqrt{t}}\right) + \Phi\left(-\frac{A}{\sigma \sqrt{t}}\right) \quad (8)$$

and for a short future position ($n = -1$),

$$P\{\max VM \geq A\} = \Phi\left(\frac{A}{\sigma \sqrt{t}}\right) - \Phi\left(-\frac{A}{\sigma \sqrt{t}}\right) \quad (9)$$

where $\sigma$ is the estimate of implied volatility calculated from the options data.

One issue worth mentioning is that in setting margin levels, it is commonly assumed that the price movements of different commodities (such as gold and silver) are independent of one another. For instance, the margin imposed on a participant who trades gold is the same whether or not he is simultaneously also trading silver or any other commodity. However, this independence assumption is far from reality, if not completely
incorrect. That is, by admitting certain positive/negative relationships between the price movements of a spectrum of different commodity prices, it can be shown that a futures clearinghouse ignoring this dependence will set margins that are systematically either too high or too low (Edwards and Neftci (1988)). As a consequence, depending upon the direction and the degree of correlation among the price extremes, from a regulator’s perspective, it will obtain either more or less safety than it would consider optimal; likewise, from a market participant’s perspective, he will have to put up more or less margin than is necessary. Further, the greater the correlation between the price movements of different commodity prices, the greater the discrepancy between the margin level set based on the independent assumption and that taking into account of the dependence structure in the co-movement of the price extremes.

1.2.3 Copula and CD-Vine

The multivariate Gaussian distribution has been well-established in the statistical dependence modeling literature. However, there’s a growing demand for non-Gaussian modeling in finance. With the availability of large multivariate data samples, it is possible to investigate non-Gaussian dependency models and to estimate parameters efficiently. Recently, Copula modeling has become increasingly popular in many fields of application. Standard references on copula theory include the books by Joe (1997) and Nelson (2006).

According to Nelson (2006), Copulas are “functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions.” The
The word ‘copula’ derives from the Latin verb *copulare* and means to bond or to tie. The backbone for the Copula modeling theory is the famous Sklar’s Theorem (Nelson (2006)).

The converse of Sklar’s Theorem is especially important, because it allows one to construct general multivariate distributions from arbitrarily chosen copulas and marginal distributions. Also, if \( \mathcal{F} \) is a continuous multivariate distribution function, it is possible to separate the univariate margins from the dependence structure, because the dependence structure is captured entirely by the copula \( \mathcal{C} \). To see this more clearly, if we assume the \( F_i \)'s are differentiable, and \( \mathcal{C} \) and \( \mathcal{F} \) are \( n \) times differentiable. Take the derivative of both sides of (10) to get the density of \( \mathcal{F} \), we get

\[
\frac{\partial^n \mathcal{F}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} = \frac{\partial^n \mathcal{C}(F_1(x_1), \ldots, F_n(x_n))}{\partial x_1 \cdots \partial x_n} \times f_1(x_1) \times \ldots \times f_n(x_n) \quad (11)
\]

That is, the density of \( \mathcal{F} \) has been decomposed as the product of the copula density and a sequence of the univariate marginal densities. Hence in this sense the copula has all the information about the dependence structure.

Copula methods have several neat properties that are very useful in the study of dependence. Copulas are invariant to strictly increasing transformations, and the widely-used measures of concordance between random variables, like Kendall's tau and Spearman's rho, are also well-defined. While these motivate the adoption of the copula method in dependence modeling, the problem in practical applications is how to identify the copula function properly, especially in an arbitrary higher-dimensional space. For the bivariate case, a rich variety of copula families is available and well-investigated; however, the choice of copula families in arbitrary dimension is rather limited.
1.2.3.1 Elliptical Copula

The elliptical copula family consists of two mostly common copulas: Normal (Gaussian) copula and Student’s t copula, both of which feature symmetric tail dependency only. Gaussian copula has the dependence function associated with multivariate normality. Without the loss of generality, its bivariate case can be written as:

$$C_N(u, v; \rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi \sqrt{1-\rho^2}} e^{-\frac{r^2 - 2\rho rs + s^2}{2(1-\rho^2)}} dr ds$$  \hspace{1cm} (12)

where $$\Phi^{-1}(\cdot)$$ is the inverse of the standard Normal CDF, and $$\rho$$ is the linear correlation coefficient of the corresponding bivariate normal distribution. Gaussian copula has neither upper tail dependence nor lower tail dependence for $$\rho < 1$$.

The bivariate Student’s t copula, a.k.a. t-copula, with $$n$$ degrees of freedom and correlation $$\rho = Corr(t_n^{-1}(u), t_n^{-1}(v))$$ can be defined as:

$$C_t(u, v; n, \rho) = \int_{-\infty}^{t_n^{-1}(u)} \int_{-\infty}^{t_n^{-1}(v)} \frac{1}{2\pi n \sqrt{1-\rho^2}} \left(1 + \frac{r^2 - 2\rho rs + s^2}{2(1-\rho^2)}\right)^{-\frac{n+2}{2}} dr ds$$  \hspace{1cm} (13)

where $$t_n^{-1}(\cdot)$$ denotes the inverse of the standard univariate Student's t CDF with $$n$$ degrees of freedom. The parameter $$n$$ controls the heaviness of the tails and as $$n \to \infty$$, $$C_t(u, v; n, \rho) \to \Phi(u, v; \rho)$$, where $$\Phi(u, v; \rho)$$ is the standard bivariate normal distribution.

1.2.3.2 Archimedean Copula

As alluded to earlier, it’s not uncommon to spot asymmetric tail dependence and volatility clustering in financial returns data. Archimedean copulas can model such asymmetries. Again, for the sake of simplicity, all functional form of copulas are limited
to bivariate versions. It should be a straightforward exercise to extend to the multivariate version with arbitrary dimensions. For the Archimedean family of copulas,
\[ C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \]  \hspace{1cm} (14)
where the generator function \( \varphi: (0,1] \rightarrow [0, +\infty) \) is a strictly decreasing convex function with \( \varphi^{-1}(1) = 0 \), assuming \( \varphi^{-1}(t) \neq 0 \) for all \( t \geq \lim_{u \to 0} \varphi(u) \). The independence copula \( C(u, v) = uv \) is an Archimedean copula with generator \( \varphi(t) = -\log t \). We’ll discuss four one-parameter Archimedean copulas: Gumbel copula (1960a), Clayton copula (1978), Frank copula (1979), and Joe copula (1997) in this chapter.

The Gumbel copula corresponds to \( \varphi(t) = (-\log t)^\theta \) and it is defined as:
\[ C_G(u, v; \theta) = \exp\left(-\left[\left(\log u\right)^\theta + \left(\log v\right)^\theta\right]\frac{1}{\theta}\right) \]  \hspace{1cm} (15)
with the dependence parameter \( \theta \) restricted to the interval \([1, +\infty)\), in which 1 and \(+\infty\) correspond to independence and the Fréchet upper bound. However, the Gumbel copula does not attain the Fréchet-Hoeffding upper bound for any value of parameter \( \theta \). In addition, the Gumbel copula does not allow negative dependence and exhibits strong right tail dependence and relatively weak left tail dependence. The Gumbel copula is appropriate for data known to be strongly correlated at higher values (gain) but less correlated at lower values (loss).

The Clayton copula corresponds to \( \varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1) \) and it is defined as:
\[ C_C(u, v; \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}} \]  \hspace{1cm} (16)
with the dependence parameter \( \theta \) restricted to the interval \((0, +\infty)\). As \( \theta \) goes to zero, the marginal distributions become independent; as \( \theta \) approaches infinity, the copula attains
the Fréchet-Hoeffding upper bound. Similar to the Gumbel copula, the Clayton copula cannot account for negative dependence. It is widely used in the field of insurance risk management as it exhibits strong left tail (loss) dependence and relatively weak right tail (gain) dependence (Embrechts, Kluppelberg, and Mikosch (1997)).

The Frank copula corresponds to \( \varphi(t) = -\log \left( \frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right) \) and it is defined as:

\[
C_F(u, v; \theta) = -\frac{1}{\theta} \log \left[ 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right]
\] (17)

The dependence parameter \( \theta \) can assume any real value in \( \mathbb{R}: (-\infty, +\infty) \). Values of \(-\infty, 0, +\infty\) correspond to the Fréchet-Hoeffding lower bound, independence, and Fréchet-Hoeffding upper bound, respectively. These characteristics make the Frank copula a convenient choice for several reasons: first, unlike Gumbel or Clayton copulas, the Frank copula permits negative dependence between the marginal distributions; secondly, similar to the elliptical copulas, dependence is symmetric in both tails for the Frank copula; lastly, the Frank copula is most encompassing in the sense that both Fréchet-Hoeffding lower bound and Fréchet-Hoeffding upper bound are included in the range of permissible dependence. The Frank copula is most appropriate for data that exhibits weak tail dependence.

The Joe copula is \( \varphi(t) = -\log \left[ 1 - (1 - t)^{\theta} \right] \) and it is defined as:

\[
C_J(u, v; \theta) = 1 - [(1 - u)^{\theta} + (1 - v)^{\theta} - (1 - u)^{\theta}(1 - v)^{\theta}]^{\frac{1}{\theta}}, 1 \leq \theta < \infty
\] (18)

The Joe copula is similar to the Gumbel and Clayton copula in that it cannot account for negative dependence. It attains the Fréchet-Hoeffding upper bound as \( \theta \to \infty \), but cannot achieve the Fréchet-Hoeffding lower bound. The range of Kendall’s \( \tau \) is between 0 and
1, with independence \((\tau = 0)\) corresponding to \(\theta = 1\). As illustrated in Figure 2, the Joe copula is similar to the Gumbel, but the right tail positive dependence is stronger (as can be observed from the tighter clustering of points in the right tail). The Joe copula is closer to being the reverse of the Clayton copula than is the Gumbel.

Table 1. Summary Characteristics of Bivariate Copula Families

<table>
<thead>
<tr>
<th>Copula</th>
<th>Parameter Range</th>
<th>Kendall’s (\tau)</th>
<th>Tail Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elliptical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>(\rho \in (-1,1))</td>
<td>(\frac{2}{\pi} \arcsin (\rho))</td>
<td>0</td>
</tr>
<tr>
<td>Student’s (t)</td>
<td>(\rho \in (-1,1), n &gt; 2)</td>
<td>(\frac{2}{\pi} \arcsin (\rho))</td>
<td>(2t_{n+1}(-\sqrt{n} + 1, 1 + \rho))</td>
</tr>
<tr>
<td>Archimedean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gumbel</td>
<td>(\theta \geq 1)</td>
<td>(1 - \frac{1}{\theta})</td>
<td>((0, 2 - 2\theta))</td>
</tr>
<tr>
<td>Clayton</td>
<td>(\theta &gt; 0)</td>
<td>(\frac{\theta + 2}{\theta + 4} \frac{D_1(\theta)}{\theta})</td>
<td>((2 - \frac{1}{\theta}, 0))</td>
</tr>
<tr>
<td>Frank</td>
<td>(\theta \in \mathbb{R}\setminus{0})</td>
<td>(1 - \frac{4}{\theta} + \frac{D_1(\theta)}{\theta})</td>
<td>((0, 0))</td>
</tr>
<tr>
<td>Joe</td>
<td>(\theta &gt; 1)</td>
<td>(1 + \frac{4}{\theta^2} \int_0^1 \log(t)(1 - t)^{2(1-\theta)/\theta} dt)</td>
<td>((0, 2 - 2\theta))</td>
</tr>
</tbody>
</table>

The contour plots presented in Figure 2 below show that different copulas can account for many kinds of dependence structure. The upper left and upper middle panel displays the Gaussian copula and the Student’s \(t\) copula, which have the familiar symmetric shape. However, knowing only the marginal distributions and linear correlation is not sufficient to describe the joint distribution. The contour of the Student’s \(t\) copula in the upper right panel behaves quite differently in the first quadrant (“positive-
positive”) versus the third quadrant (“negative-negative”), where the contours are more tightly clustered along the diagonal even though it’s symmetric and elliptically-shaped like the Gaussian copula. The Gumbel copula is fit to account for returns more highly correlated in bull markets than bear markets; while on the contrary, the Clayton copula is fit to account for returns more highly correlated in bear markets than bull markets, which is the case for many financial time-series. Frank copula can generate symmetric dependence, complementing the Gumbel copula and the Clayton copula. The main advantages of Archimedean copulas are two-fold: these copulas do have a closed-form solution and various dependence structures can be modeled with them. A drawback is their complex structure for higher dimensions.

![Contour Plots Illustration of Common Copula Families](image)

Figure 2. Contour Plots Illustration of Common Copula Families
1.2.3.3 Pair-Copula Decomposition of a Multivariate Distribution

While other multivariate copulas suffer from criticisms of having a rather inflexible structure, vine copulas not only overcome these limitations but also are able to model rather complex dependency patterns by benefiting from leveraging the rich variety of bivariate copula families. Initially proposed by Joe (1996) and developed in more detail in Bedford and Cooke (2001; 2002), a vine copula is “a flexible graphical model for describing multivariate copulas built up using a cascade of bivariate copulas”, so called “pair-copula construction (PCC)” (Aas et al. (2009); Brechmann and Schepsmeier (2013)). These copulas were first introduced in finance by Aas et al. (2009), in which they derived statistical inference techniques for the two classes of Canonical (C-) Vine and D-Vine copulas.

To formally define the general notion of regular vine (R-vine), according to Bedford and Cooke (2001; 2002), $\mathcal{V} = (T_1, \ldots, T_{n-1})$ is an R-vine on $n$ elements if

i. $T_1$ is a tree with nodes $N_1 = \{1, \ldots, n\}$ and a set of edges denoted $E_1$

ii. For $i = 2, \ldots, n-1$, $T_i$ is a tree with nodes $N_i = E_{i-1}$ and edge set $E_i$

iii. For $i = 2, \ldots, n-1$, and $\{a, b\} \in E_i$ with $a = \{a_1, a_2\}$, and $b = \{b_1, b_2\}$ it must hold that $\#(a \cap b) = 1$ (proximity condition)

That is, an R-vine on $n$ elements is a nested set of $n - 1$ trees such that the edges of tree $j$ become the nodes of tree $j + 1$. The proximity condition insures that two nodes in
tree $j + 1$ are only connected by an edge if these nodes share a common node in tree $j + 1$.

Figure 3. Illustration of Five-dimensional Canonical Vine Trees with Edge Indices
As illustrated in Figure 3 and Figure 4, C-Vine tree structures have a “star” property, while D-Vine tree structures have a path dependency. That is, one constructs a C-Vine by choosing the most influential variable, while constructing D-Vines by choosing a specific order of the variables. In the first D-Vine tree, the dependence of the 1\textsuperscript{st} and 2\textsuperscript{nd} variable, of the 2\textsuperscript{nd} and 3\textsuperscript{rd}, of the 3\textsuperscript{rd} and 4\textsuperscript{th}, and so on, is modeled using pair-copula, i.e., if one assumes the order 1,...,n, one models the pairs (1,2), (2,3), (3,4), etc. In the second tree, the conditional dependence of the 1\textsuperscript{st} and 3\textsuperscript{rd} given the 2\textsuperscript{nd} variable (the pair (1,3|2)), the 2\textsuperscript{nd} and 4\textsuperscript{th} given the 3\textsuperscript{rd} variable (the pair (2,4|3)), and so on, is modeled. In a similar way, pairwise dependencies of variables a and b are modeled in subsequent trees conditioned on those variables which lie between the variables a and b in the first tree, e.g., the pair (1,5|2,3,4). Hence, this leads to the decomposition of a multivariate density, the CD-Vine density (Brechmann and Schepsmeier (2013)): 

![Figure 4. Illustration of Five-dimensional D-Vine Trees with Edge Indices](image)
\[ f(x) = \]
\[ \prod_{k=1}^{n} f_k(x_k) \times \]
\[ \prod_{i=1}^{n-1} \prod_{j=1}^{n-i} c_{j, j+i | (j+1), (j+i-1)} F(x_j | x_{j+1}, ..., x_{j+i-1}), F(x_{j+i} | x_{j+1}, ..., x_{j+i-1}) | \theta_{j, j+i | (j+1), (j+i-1)} \] (19)

where \( c \) is the corresponding density of the \( n \)-dimension Copula function \( C \).

By allowing arbitrary bivariate copulas for each pair-copula term in the decomposition (eq. (19)), the multivariate copulas obtained from CD-Vine structure constitute very flexible models, since bivariate copulas can easily accommodate complex dependence structures such as asymmetric dependence or strong joint tail behavior (Joe, Li, and Nikoloulopoulos (2010)). The last question is how should one specify the CD-Vine structure in the dependency context; in particular, how should one determine the order of the nodes since C-Vine structure has a "star" variable, while D-Vine structure is path dependent. In this paper, I choose a sequential method to select a CD-Vine copula specification based on Kendall’s \( \tau \) (Dissmann et al. (2011)).

**Algorithm**

Sequential method to select an R-Vine model based on Kendall’s \( \tau \).

1. Calculate the empirical Kendall’s \( \tau \) for all possible variable pairs \( \{i, j\}, 1 \leq i < j \leq n \).

2. Select the spanning tree\(^{12}\) that maximizes the sum of absolute empirical Kendall’s \( \tau \), i.e.

\(^{12}\) A spanning tree is a graph in which each two nodes are connected by a unique sequence of edges.
Argmax \sum_{e=(i,j) \text{ in spanning tree}} |\hat{\tau}_{i,j}|

3. For all these pairs select a copula and fit the corresponding parameter.

4. Transform the observations using the copulas and parameters from Step 3 to obtain the transformed values.

5. Use these transformed observations to calculate empirical Kendall’s \( \tau \) for all possible pairs. By “possible” it is meant that the proximity condition must be fulfilled.

6. Proceed with Step 2 until the R-Vine is fully specified.

Note that the sequence method presented above does not guarantee finding a global optimum\(^{13}\) (Vuong (1989)). However, this sequential approach is reasonable because the specification ensures that the first tree of the R-vine captures the strongest dependencies among the marginal distributions, which is crucial because in most cases the first tree of the R-vine has the greatest influence on the model fit.

1.3 Empirical Evaluation and Results

The empirical analysis is conducted on CBOT corn, soybeans, and wheat futures at weekly and monthly time intervals, from February 1989 to January 2010. As a benchmark, as well as the building block for the CD-Vine Copula modeling stage, Harrison’s method and the Extreme Value Theory are used on the panel of agricultural products.

\(^{13}\)Here, global optimum is meant in terms of model fit, e.g. higher likelihood, smaller AIC/BIC or superior in terms of the likelihood-ratio based test for comparing non-nested models proposed by Vuong (1989).
commodity futures for the univariate case to calibrate the distribution of the maximum. Harrison's method and EVT are compared via Kupiec's method and the out-of-sample log-likelihood (OOSLLF) method of Norwood, Roberts, and Lusk (2004) on CBOT corn, soybeans, and wheat futures at weekly and monthly time intervals at various significance levels. One input in Harrison’s method is the estimate of the implied volatility. It is calculated by taking the average of the implied volatilities of the two closest at-the-money put options and call options. There are four options in total, with maturities longer than the forecast interval on the final trading day prior to the start of the forecast interval. Implied volatilities are calculated following the methodology proposed by Barone-Adesi and Whaley (1987). For EVT, the distribution of the maximum and minimum are estimated using the previous 1 and 5 years of data for weekly and monthly distributions, respectively.

Before implementing the CD-Vine Copula dependence structure to estimate the joint distribution of extreme prices movement of corn, soybeans, and wheat futures, Harrison’s method and EVT are used to estimate the distribution of extrema over the same period for a single commodity. The findings from this can also serve as a benchmark as the univariate version. However, it should be noted that the use of this single-commodity method may lead to systematic under- or over-estimations of liquidity needs. That is because most firms participating in commodity futures market are simultaneously hedging across multiple commodities. Failing to take into account of the direction and magnitude of correlations among the portfolio of commodity futures will adversely affect the effectiveness of hedging and liquidity needs.
To expand the analysis into the multivariate case, the distribution of extrema for the single commodity, i.e., corn, soybeans, and wheat, obtained separately using Harrison’s method and EVT can be transformed into the corresponding uniform margins using either an empirical or parametric probability integral transformation, because a copula is a multivariate distribution function with standard uniform marginal distributions. Then, the CD-Vine model of the joint distribution of extreme futures movement of corn, soybeans, and wheat futures is estimated, by executing the algorithm of the sequential method based on Kendall’s τ. Further statistical inferences can be made by generating Monte Carlo simulations of the estimated dependence structure.

1.3.1 Data Description and Preliminary Tests

To study the tail behavior of agricultural commodity futures, we select all futures contracts that are most liquid and widely held. End-of-day prices of the front month futures of corn, soybean, and wheat traded on Chicago Board of Trade (a.k.a. CME Group) have been obtained from Commodity Research Bureau (CRB) Futures Price Database. The daily returns are defined as $r_t = \log \frac{P_t}{P_{t-1}} \times 100$. Figure 5 plots daily prices, returns and squared returns for each series analyzed. Each plot of daily returns confirms the typical empirical time series properties. All series of returns exhibit mean-reverting and volatility clustering, i.e. periods of low volatility were followed by periods of high volatility. The squared daily returns also confirm the observation of volatility clustering with large changes typically followed by large changes. For risk management
purposes, any measure of risk exposure should be conditional on the current volatility regime.

![Image](image.png)

Figure 5. Daily Prices, Returns, and Squared Returns

Basic statistics for the front month futures are summarized in Table 2. As is the case with most financial time series, the front month agricultural commodity futures exhibit
evidence of fat tails, with the kurtosis of soybean and wheat far larger than 3. The Jarque-Bera normality test also confirms that all returns are not normally distributed. The Augmented Dickey-Fuller test rejects the null hypothesis of a unit root for all series. Among all three agricultural commodity futures, wheat is more volatile than others. The standard deviation of wheat front month futures returns is 50% higher than the standard deviation of soybean front month futures returns. Interestingly, all front month agricultural commodity futures returns exhibit positive skewness. The squared daily returns shown in Figure 5 exhibit volatility clustering, i.e. large changes tend to be followed by large changes in bursts. The Ljung-Box Q-statistics reported in Table 2 also reject the null hypothesis of no autocorrelation through 20-lags at a 1% significance level for all series.

---

14 The Jarque-Bera test statistic JB is defined as \( JB = \frac{n}{6} S^2 + \frac{1}{4} K^2 \), where \( n \) is the number of observations. \( S \) is the sample skewness and \( K \) is the sample kurtosis. The statistic \( JB \) has an asymptotic chi-square distribution with two degrees of freedom and can be used to test the null hypothesis that the data are from a Normal distribution. The null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being 0.

15 The Ljung-Box test tests whether any of a group of autocorrelations of a time series are different from zero. The test statistics can be calculated as \( Q = n(n + 2) \sum_{k=1}^{h} \hat{\beta}_k^2 \frac{1}{n-k} \), where \( n \) is the sample size, \( \hat{\beta}_k \) is the sample autocorrelation at lag \( k \), and \( h \) is the number of lags being tested. For significance level \( \alpha \), the critical region for rejection of the hypothesis of randomness is \( Q > \chi^2_{1-a,h} \), where \( \chi^2_{1-a,h} \) is the \( \alpha \)-quantile of the chi-square distribution with \( h \) degrees of freedom.
Table 2. Summary Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Corn</th>
<th>Soybean</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start Date</td>
<td>24-Feb-89</td>
<td>24-Feb-89</td>
<td>24-Feb-89</td>
</tr>
<tr>
<td>End Date</td>
<td>15-Jan-10</td>
<td>15-Jan-10</td>
<td>15-Jan-10</td>
</tr>
<tr>
<td>Sample Size</td>
<td>5206</td>
<td>5196</td>
<td>5201</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0061</td>
<td>0.0047</td>
<td>0.0036</td>
</tr>
<tr>
<td>Median</td>
<td>-2.0013</td>
<td>-0.5020</td>
<td>-3.7469</td>
</tr>
<tr>
<td>Max</td>
<td>16.1557</td>
<td>23.2138</td>
<td>37.7102</td>
</tr>
<tr>
<td>Min</td>
<td>-16.3992</td>
<td>-26.6629</td>
<td>-31.4843</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>3.0627</td>
<td>2.4963</td>
<td>3.6590</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.1004</td>
<td>0.0119</td>
<td>0.2466</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.3307</td>
<td>19.3834</td>
<td>6.9454</td>
</tr>
<tr>
<td>Jarque-Bera Test</td>
<td>1189.1816</td>
<td>81571.5028</td>
<td>10528.0869</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>ADF Unit-Root Test$^{16}$</td>
<td>-17.061</td>
<td>-17.374</td>
<td>-18.814</td>
</tr>
<tr>
<td>p-value</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Auto Corr- $r$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>-0.343</td>
<td>-0.328</td>
<td>-0.369</td>
</tr>
<tr>
<td>Lag 5</td>
<td>-0.003</td>
<td>-0.067</td>
<td>-0.043</td>
</tr>
<tr>
<td>Lag 10</td>
<td>-0.031</td>
<td>0.034</td>
<td>0.024</td>
</tr>
<tr>
<td>Lag 20</td>
<td>0.012</td>
<td>0.016</td>
<td>0.006</td>
</tr>
<tr>
<td>Ljung-Box (20)</td>
<td>645.88</td>
<td>831.65</td>
<td>748.49</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Auto Corr- $r^2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>0.259</td>
<td>0.473</td>
<td>0.338</td>
</tr>
<tr>
<td>Lag 5</td>
<td>0.228</td>
<td>0.267</td>
<td>0.143</td>
</tr>
<tr>
<td>Lag 10</td>
<td>0.187</td>
<td>0.263</td>
<td>0.198</td>
</tr>
<tr>
<td>Lag 20</td>
<td>0.171</td>
<td>0.277</td>
<td>0.057</td>
</tr>
<tr>
<td>Ljung-Box (20)</td>
<td>4046.5</td>
<td>10164</td>
<td>2992.8</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

$^{16}$ An intercept term (no trend) is included in the integrating regression for the ADF Unit-Root test.
1.3.2 Comparison of Harrison's Method and EVT

Price levels at the 90th, 95th, and 99th percentiles for positive and negative extrema realizations are computed for Harrison's method and EVT on CBOT corn, soybean, and wheat futures at weekly, and monthly time intervals. Years from 1989 through 2010 are used to evaluate the models, resulting in 796 weekly forecasts for corn, 827 weekly forecasts for soybean, 759 weekly forecasts for wheat, and 227 monthly forecasts for all three agricultural commodity futures. The maximum positive and negative extremes in the forecast periods are compared to the corresponding percentiles to evaluate the forecasting ability of each method in terms of the distribution of the extreme price movements. The EVT distributions are estimated prior to each out-of-sample period. Weekly estimates are calendar weeks, using the closing price on the final trading day of the previous week, and calculating the extrema based upon the maximum and minimum closing prices. Monthly estimates are calendar months, using the closing price on the final trading day of the previous month, and calculating the extrema based upon the maximum and minimum closing prices. As OOSLLF statistics have no distribution, they cannot be used to make interval estimates of the methods’ predictive performance. Therefore, bootstrap distributions of the OOSLLF statistics were generated to test whether one model out-performs the other. For each commodity/frequency combination, 1000 samples were drawn with replacement for each time horizons. For each replication, the mean of the OOSLLF for Harrison’s method and EVT were calculated. If a particular

\[17\] Certain weekly and monthly observations were omitted due to the lack of useful options prices. In these cases, the corresponding EVT forecasts were also omitted.
method had a higher mean in 95%, 97.5%, or 99.5% of the replications, it was noted as being significant at the 1% level in Table 5.

Table 3. Proportion of Exceedance using Weekly Data

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Distribution</th>
<th>1% LB</th>
<th>5% LB</th>
<th>10% LB</th>
<th>10% UB</th>
<th>5% UB</th>
<th>1% UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corn</td>
<td>Harrison</td>
<td>0.0050</td>
<td>0.0239*</td>
<td>0.0440*</td>
<td>0.0302*</td>
<td>0.0101*</td>
<td>0.0013*</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0239*</td>
<td>0.0616</td>
<td>0.1030</td>
<td>0.1018</td>
<td>0.0666*</td>
<td>0.0289*</td>
</tr>
<tr>
<td>Soybean</td>
<td>Harrison</td>
<td>0.0097</td>
<td>0.0266*</td>
<td>0.0496*</td>
<td>0.0278*</td>
<td>0.0109*</td>
<td>0.0000*</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0230*</td>
<td>0.0641*</td>
<td>0.0992</td>
<td>0.0883</td>
<td>0.0520</td>
<td>0.0181*</td>
</tr>
<tr>
<td>Wheat</td>
<td>Harrison</td>
<td>0.0040*</td>
<td>0.0145*</td>
<td>0.0356*</td>
<td>0.0395*</td>
<td>0.0184*</td>
<td>0.0040*</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0184*</td>
<td>0.0487</td>
<td>0.1041</td>
<td>0.1146</td>
<td>0.0685*</td>
<td>0.0224*</td>
</tr>
<tr>
<td>All</td>
<td>Harrison</td>
<td>0.0063*</td>
<td>0.0218*</td>
<td>0.0432*</td>
<td>0.0323*</td>
<td>0.0130*</td>
<td>0.0017*</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0218*</td>
<td>0.0584*</td>
<td>0.1020</td>
<td>0.1012</td>
<td>0.0621*</td>
<td>0.0231*</td>
</tr>
</tbody>
</table>

^ The body of the table reports the proportion of the observations that were less than the LB or greater than the UB level
* Indicates significance at the 1%: the performance is significantly different from the expected proportion

Table 3 and Table 4 present the proportion of actual extreme prices which were outside the intervals predicted by the Harrison’s method and EVT. Table 3 reports the percentage of the actual extrema that fell outside of the forecasted intervals with a weekly horizon. For example, for corn, 0.5% and 2.39% of the out-of-sample extrema fell below the 5% lower bound forecasted by the Harrison’s method and EVT, respectively. Using a simple test of proportions, the number of observations falling below the 5% lower bound
is significantly different- at the 10% level- for EVT, but not for the Harrison’s method. For the weekly data, EVT is rejected in 9 out of 18 individual commodity tests, but 4 out of 6 of the combined tests; while Harrison’s method is rejected in 16 out of 18 tests, and all 6 of the combined tests. Certain patterns of prediction are indicated by each method. EVT consistently has too many observations falling beyond the predicted boundaries, indicating that its estimates are too conservative; while Harrison’s method consistently has too few observations falling beyond the boundaries, indicating that the predicted intervals are too wide. These differences hold true across all agricultural commodity return time series and time periods. Table 4 reports the monthly results. Among these results, 8 out of 18 individual commodity tests are rejected for Harrison’s method, compared to 9 for EVT. Five are rejected for each method in the combined tests.

Table 5 represents the mean OOSLLF scores for each model and the bootstrap test results. At the weekly horizon, Harrison’s method and EVT has very comparable OOSLLFs; however, at the monthly horizon, Harrison’s method has meaningfully higher OOSLLFs for both upper and lower tails, across all agricultural commodities return time series.
Table 4. Proportion of Exceedance using Monthly Data

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Distribution</th>
<th>1% LB</th>
<th>5% LB</th>
<th>10% LB</th>
<th>10% UB</th>
<th>5% UB</th>
<th>1% UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corn</td>
<td>Harrison</td>
<td>0.0088</td>
<td>0.0352</td>
<td>0.0617*</td>
<td>0.0396*</td>
<td>0.0220*</td>
<td>0.0044</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0176</td>
<td>0.0925*</td>
<td>0.1322</td>
<td>0.1454*</td>
<td>0.0661</td>
<td>0.0088</td>
</tr>
<tr>
<td>Soybean</td>
<td>Harrison</td>
<td>0.0088</td>
<td>0.0352</td>
<td>0.0705</td>
<td>0.0529*</td>
<td>0.0088*</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0264*</td>
<td>0.0749*</td>
<td>0.1189</td>
<td>0.1454*</td>
<td>0.0705</td>
<td>0.0044</td>
</tr>
<tr>
<td>Wheat</td>
<td>Harrison</td>
<td>0.0088</td>
<td>0.0308</td>
<td>0.0617*</td>
<td>0.0529*</td>
<td>0.0132*</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0308*</td>
<td>0.0661</td>
<td>0.1454*</td>
<td>0.1233</td>
<td>0.0881*</td>
<td>0.0220*</td>
</tr>
<tr>
<td>All</td>
<td>Harrison</td>
<td>0.0088</td>
<td>0.0338*</td>
<td>0.0646*</td>
<td>0.0485*</td>
<td>0.0147*</td>
<td>0.0015*</td>
</tr>
<tr>
<td></td>
<td>EVT</td>
<td>0.0250*</td>
<td>0.0778*</td>
<td>0.1322*</td>
<td>0.1380*</td>
<td>0.0749*</td>
<td>0.0117</td>
</tr>
</tbody>
</table>

^ The body of the table reports the proportion of the observations that were less than the LB or greater than the UB level

* Indicates significance at the 1%: the performance is significantly different from the expected proportion
Table 5. Out-of-Sample Log-Likelihood Function (OOSLLF) Values

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Tail</th>
<th>Weekly Harrison</th>
<th>Weekly EVT</th>
<th>Monthly Harrison</th>
<th>Monthly EVT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2.9555***</td>
<td>2.3364***</td>
<td>2.1714***</td>
<td>1.8797</td>
</tr>
<tr>
<td>Corn</td>
<td>Upper</td>
<td>2.9202***</td>
<td>2.3532***</td>
<td>2.0950***</td>
<td>1.7617</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>2.9887***</td>
<td>2.4026***</td>
<td>2.2192***</td>
<td>1.8761</td>
</tr>
<tr>
<td>Soybean</td>
<td>Upper</td>
<td>2.9428***</td>
<td>2.3760***</td>
<td>2.2037***</td>
<td>1.8717</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>2.8571***</td>
<td>2.3191***</td>
<td>2.1105***</td>
<td>1.6333</td>
</tr>
<tr>
<td>Wheat</td>
<td>Upper</td>
<td>2.8542***</td>
<td>2.3414***</td>
<td>2.0388***</td>
<td>1.6317</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>2.9357***</td>
<td>2.3539***</td>
<td>2.1670***</td>
<td>1.7964</td>
</tr>
<tr>
<td>All</td>
<td>Upper</td>
<td>2.9070***</td>
<td>2.3574***</td>
<td>2.1125***</td>
<td>1.7550</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>2.9202***</td>
<td>2.3532***</td>
<td>2.0950***</td>
<td>1.7617</td>
</tr>
</tbody>
</table>

*** Indicates bootstrap test rejection of identical means at the 1% level.

1.3.3 Results of CD-Vine Copula

One question remained to be answered is: why when we simply combine corn, soybean, and wheat together to calculate Kupiec’s test for weekly and monthly returns across all agricultural commodities, both Harrison’s method and EVT are largely rejected at 10% significance level across the board? In other words, the naïve way of portfolio construction, assuming a perfect correlation of one, is not a good way to calibrate hedge finance, because it fails to consider the interconnections of price movements in an agricultural commodities portfolio. In this section, we generalize the Harrison’s method.
and EVT to the multivariate setting using the CD-Vine copula defined in the earlier section.

The major advantage of vine-based pair-copula models is their flexibility in modeling multivariate dependence. They allow for flexible specification of the dependence between different pairs of marginal distributions using bivariate copula functions, while specifying an overall dependence between all of the marginal distributions. We utilize both the C-Vine and the D-Vine structures in this chapter to account for the dependence structure of Harrison’s and EVT marginal at different confidence levels for both upper and lower tails. For the three-dimensional case, the full decomposition is:

\[
f(x_1, x_2, x_3) = f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|1,2}(x_3|x_1, x_2)
\]

\[
= f_1(x_1) \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot f_{2}(x_2) \cdot c_{23|1}(F(x_2|x_1), F(x_3|x_1)) \cdot c_{13}(F_1(x_1), F(x_3)) \cdot f_{3}(x_3)
\]

That is to say, in a three-dimensional vine copula, there are three different decompositions and each of the three decompositions is both a C-Vine and a D-Vine.

In order to further investigate the dependence relationship between agricultural commodity futures, we map the standardized residuals into the unit square applying the probability-integral transformation to estimate the pair-copulas via maximum likelihood (MLE). The parameters of the marginal distribution and the pair-copula can be estimated simultaneously using MLE; however, the computation becomes very complex very

---

18 In general, for the \(d\)-dimensional case, there are \(n! / 2\) different C-Vines and \(n! / 2\) distinct D-Vines.

19 If \(X\) is a continuous random variable with CDF \(F_X(x)\) and if \(Y = F_X(X)\), then \(Y\) is a uniform random variable on the interval [0,1].
quickly as the dimension of the problem increases. Therefore, we use the Inference Function for Margins (IFM) method, where the parameter estimation is performed in two steps: the parameters in the marginal distributions are estimated first; and the copula parameters are estimated conditioned on the marginal distributions estimates in the previous step.

For C-Vine decomposition, the log-likelihood is given by:

\[
l(x; \Theta) = \sum_{j=1}^{d-1} \sum_{i=1}^{d-j} \sum_{t=1}^{T} \log \left( c_{j,i+1,i,...,i+1,t}(F(x_{j,t} \mid x_{1,t}, ..., x_{j-1,t}), F(x_{j+1,t} \mid x_{1,t}, ..., x_{j-1,t})) \right)
\]

For D-Vine decomposition, the log-likelihood is given by:

\[
l(x; \Theta) = \sum_{j=1}^{d-1} \sum_{i=1}^{d-j} \sum_{t=1}^{T} \log \left( c_{i,j+1,i+1,...,i+1,t}(F(x_{i,t} \mid x_{i+1,t}, ..., x_{i+j-1,t}), F(x_{i+j,t} \mid x_{i+1,t}, ..., x_{i+j-1,t})) \right)
\]

We consider a total of six different copulas for the selection and estimation of pair-copula, namely Gaussian copula, Student’s t copula, Clayton copula, Gumbel copula, Frank copula, and Joe copula. In addition, we need to determine which key variable to use that governs interactions in the dataset at the root of the C-Vine. The key variable is identified by selecting the series that has the strongest dependence relationship with other series, using the C-Vine structure selection criterion described by Czado et al. (2012). As shown in Figure 6 and Figure 7, corn futures exhibit the strongest dependence among the weekly and monthly Harrison’s and EVT marginal series. Therefore, we have decided to locate corn futures at the root of the canonical vine. By selecting all further C-Vine root nodes in the similar fashion, the root node order in the data set is determined as corn,
soybean, and wheat. The chosen pair-copula decomposition, along with best copula fits and parameter estimates using the full sample for the C-Vine is shown in Figure 10 as an illustration. In the case of the Student’s t copula, the parameter is the degrees of freedom. In the case of the Archimedean copula, the parameter is a scalar value of the rank correlation. The chosen D-Vine pair-copula decomposition, along with best copula fits and parameter estimates using the full sample is shown in Figure 11 as an illustration.
Figure 6. Correlation Plot with Weekly Harrison’s Marginal

Figure 7. Correlation Plot with Monthly Harrison’s Marginal
Figure 8. Correlation Plot with Weekly EVT Marginal

Figure 9. Correlation Plot with Monthly EVT Marginal
Figure 10. Illustration of C-Vine Tree Plot of Weekly Harrison’s Marginal

Figure 11. Illustration of D-Vine Tree Plot of Weekly Harrison’s Marginal

Figure 12. Illustration of C-Vine Tree Plot of Weekly EVT Marginal

Figure 13. Illustration of D-Vine Tree Plot of Weekly EVT Marginal
Table 6. Proportion of Exceedance using Weekly Data with CD-Vine Copula

<table>
<thead>
<tr>
<th>Marginal Distribution</th>
<th>Pair-Copula</th>
<th>1% LB</th>
<th>5% LB</th>
<th>10% LB</th>
<th>10% UB</th>
<th>5% UB</th>
<th>1% UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harrison C-Vine</td>
<td>0.0124</td>
<td>0.0124*</td>
<td>0.0186*</td>
<td>0.0062*</td>
<td>0.0165*</td>
<td>0.0124</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0083</td>
<td>0.0103*</td>
<td>0.0269*</td>
<td>0.0083*</td>
<td>0.0124*</td>
<td>0.0041</td>
<td></td>
</tr>
<tr>
<td>EVT C-Vine</td>
<td>0.0000</td>
<td>0.0103*</td>
<td>0.0310*</td>
<td>0.0186*</td>
<td>0.0062*</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0000</td>
<td>0.0021*</td>
<td>0.0393*</td>
<td>0.0434*</td>
<td>0.0021*</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>

* Indicates significance at the 10%: the performance is significantly different from the expected proportion

Table 7. Proportion of Exceedance using Monthly Data with CD-Vine Copula

<table>
<thead>
<tr>
<th>Marginal Distribution</th>
<th>Pair-Copula</th>
<th>1% LB</th>
<th>5% LB</th>
<th>10% LB</th>
<th>10% UB</th>
<th>5% UB</th>
<th>1% UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harrison C-Vine</td>
<td>0.0000</td>
<td>0.0114</td>
<td>0.0114*</td>
<td>0.0227</td>
<td>0.0455</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0114</td>
<td>0.0000</td>
<td>0.0341</td>
<td>0.0114*</td>
<td>0.0341</td>
<td>0.0114</td>
<td></td>
</tr>
<tr>
<td>EVT C-Vine</td>
<td>0.0000</td>
<td>0.0114</td>
<td>0.0568</td>
<td>0.0227</td>
<td>0.0114</td>
<td>0.0114</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0000</td>
<td>0.0114</td>
<td>0.0455</td>
<td>0.0455</td>
<td>0.0114</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>

* Indicates significance at the 10%: the performance is significantly different from the expected proportion

^ The body of the table reports the proportion of the observations that were less than the LB or greater than the UB level
Based on the fitted C-Vine and D-Vine copula illustration in Figure 10 through Figure 13, it is clear even though the two pair-copulas look structurally similar, the copula chosen and the corresponding parameters estimated show some resemblance and yet is still different. This is because the C-Vine copula has a “star” structure, while the D-Vine copula exhibits path-dependency. For instance, if we compare Tree 1 from Figure 10 to that from Figure 11, the dependence between Corn and Soybean is best fitted with Clayton copula with rank correlation of 0.84 under both C-Vine and D-Vine; however, the dependence between Corn and Wheat is best fitted with a Gumbel copula with rank correlation of 0.95 under C-Vine (Figure 10), while the dependence between Soybean and Wheat is best fitted with Gaussian copula with rank correlation of 0.91 (Figure 11) under D-Vine. Consequently, Tree 2 in C-Vine shows the dependence between the pair Corn/Wheat and Corn/Soybean is best fitted by the Clayton copula with rank correlation of 0.07 (Figure 10); whereas the Tree 2 in D-Vine exhibits the dependence between the pair Corn/Soybean and Soybean/Wheat is best fitted by the Gaussian copula with rank correlation of 0.36 (Figure 11).

Finally, Table 6 and Table 7 summarizes the proportion of actual extreme prices which were outside the intervals predicted by the CD-Vine copula with Harrison’s and EVT marginal with weekly and monthly returns series at 1%, 5%, and 10% significance levels. Compared to the “All” section in Table 3 and Table 4, these results are significantly better across significance levels, especially in the monthly returns series and higher significance levels. Specifically, for the weekly data, Harrison’s method and EVT are rejected in 4 out of 6 tests for both C-Vine and D-Vine; while for the monthly data,
Harrison’s method is rejected in 1 out of 6 tests for both C-Vine and D-Vine; EVT is not rejected in any tests for C-Vine or D-Vine. Therefore, properly accounting for the dependence structure within the agricultural commodity futures portfolio by utilizing the CD-Vine copula on Harrison’s and EVT marginal leads to better estimation of extreme price swings and effective hedge finance.

1.4 Conclusion

Financing risk in hedging can become a binding constraint to hedgers, especially in times of high volatility. Typically, firms have not attempted to estimate the amount of financing that a hedger might need over the course of a hedge due to the perceived complexity. These results show that a relatively simple method for a single commodity case, requiring only time to maturity of the hedge and the implied volatility from options, both readily available from commercial sources, and calculations that can be performed in common statistical software can permit even non-specialists to evaluate the levels of financing that may be required to maintain a long-term hedge.

While the use of Harrison’s method appears to be well-suited to estimating the distribution of extrema over long periods for a single commodity, it should be noted that in practice, without explicitly accounting for the dependence structure across multiple commodities, one may wind up with a systematic under- or over-estimations of liquidity needs, depending on the direction and the magnitude of the correlation across commodities. This is because in the real world, most firms participating in commodity futures market are simultaneously hedging across multiple commodities. The most
important consideration is whether the firm should analyze commodities separately or simultaneously for purposes of hedging finance, since most firms are simultaneously hedging multiple commodities whose correlation of price changes is less than one.

One interesting conclusion of this paper is, at the single commodity level, Harrison’s method performs at least as well as EVT across various significance levels for both upper and lower tails. Extending the single commodity analysis to a multivariate setting with CD-Vine copula methods on Harrison’s and EVT marginal, though forsaking the tractability of the univariate version, proves to be worthwhile in terms of estimating price extrema in agricultural commodity futures portfolio across various significance levels for both upper and lower tails. However, as is often observed, in highly volatile markets, correlation between assets tends towards one, i.e. during large moves such as in 2008, when the far tails of the distribution were reached, assuming a correlation of one might be the most appropriate approximation. For future research, one meaningful extension is to quantify the impact or cost-saving for hedge finance using multi-variate copula method vs. the single commodity method. Another useful extension to the present method would be to explore alternative CD-Vine tree specifications, in addition to Gaussian copula, Student’s $t$ copula, Clayton copula, Gumbel copula, Frank copula, and Joe copula assessed in this paper.
Chapter 2: Estimating Value-at-Risk and Expected Shortfall with a GARCH-Copula Approach in a Multi-Commodity Portfolio

2.1 Introduction

Witnessing the recent surge in price and volatility of agricultural commodity markets, it becomes increasingly important to assess the probability of rare and extreme price movement in the risk management of agricultural commodity futures. Similar to other financial time series, agricultural commodity futures exhibit the characteristics of time-varying volatility and fat tails. An appropriate risk measurement of agricultural commodity futures should be able to capture these two features of the returns. The preferred approach for measuring market risk is Value-at-Risk (VaR), which measures the worst expected losses in the market value over a specific time interval at a given significance level. VaR has gained wide popularity because it is conceptually simple and it summarizes the risk into a single number, which can be easily communicated to relevant stakeholders. A key element in VaR calculation is the distribution assumed for the financial returns under study. The common practice in estimating VaR is to assume that asset returns are normally distributed. However, this fails to capture the observed skewness and kurtosis exhibited in most financial time series. The Normal distribution based VaR models tend to underestimate risk and results in excessive observations of
VaR violations. As Extreme Value Theory focuses on the modeling of the tail behavior of a distribution using only extreme values rather than the whole dataset, it can potentially provide a more accurate estimate of tail risk.

In recent years, an increasing number of research studies have analyzed the extreme events in financial markets resulting from currency crises, stock market crashes and credit crises (Longin 1996; Turan G. Bali 2003; Gençay 2004). It is important to note that Extreme Value Theory (EVT) assumes that the data under study is independently and identically distributed (i.i.d.), which is not necessarily the case for most financial returns. In order to properly account for stochastic volatility, we adopt McNeil and Frey’s (McNeil and Frey (2000)) approach to model agricultural commodity futures return and measure the associated tail risk. McNeil and Frey’s solution to observed volatility clustering in financial returns is to first fit a GARCH-type volatility model to the returns data by pseudo-maximum likelihood (PML), and apply the EVT to the GARCH residuals in the second stage. The advantage of this GARCH-EVT two stage approach lies in its ability to capture conditional heteroscedasticity in the data through the GARCH framework, while at the same time modeling the extreme tail behavior through the EVT method. Applying McNeil and Frey’s approach in various financial markets, Bali and Neftci (2003), Byström(2005), Fernandez (2005), Chan and Gray(2006), Bhattacharyya and Ritolia (2008), and Liu (2011) demonstrate that a risk metric based on the statistics of the extremes can measure the risk exposure more accurately than the Normal-distribution-based approaches.
In this chapter, we employ McNeil and Frey’s two step approach and conditional Extreme Value Theory to estimate VaR and Expected Shortfall (ES) for long and short positions in the agricultural commodity futures market, and compare this approach to conventional unconditional Normal or Student’s t distribution based models. The backtesting demonstrates that this GARCH-EVT approach provides a significant improvement over the widely used Normal and Student’s t distribution based VaR and ES models, which tend to underestimate the true risk and fail to provide statistically significant VaR estimates. To capture the tail dependence and properly estimate portfolio VaR, copula models are introduced to estimate a portfolio measure of risk in a multi-commodity setting. This has broad applications, for instance, for a grain elevator that is purchasing corn, wheat and soybeans simultaneously. The conventional approach to this problem is to use a multivariate GARCH (a.k.a. MGARCH) model to estimate the conditional covariance between the futures prices. However, the typical MGARCH model approach inevitably suffers from being unduly restrictive because of the classical joint multivariate Gaussian assumption, despite the plentiful empirical evidence against elliptical distributions in commodity price returns. Also, from the perspective of computational efforts, the number of parameters to be estimated in the MGARCH specification often increases rapidly, stemming from the high-dimensional nature of the problem.

Copulas, introduced by Sklar in 1959 are statistical functions that join together one-dimensional marginal distributions to form multivariate distributions. Copulas have become a popular multivariate modeling tool in many fields, such as actuarial science,
biomedical science, engineering and empirical finance. During the past decade, there are an increasing number of financial applications of copula theory, mainly attributable to its flexibility in constructing a suitable joint distribution to fit the non-normality in returns data. The key characteristic of copula model is the separation of the joint distribution of returns into two components - the marginal distribution and the dependence structure. The approach is designed to capture stylized facts of non-normal financial returns using proper marginal distributions, leaving all information about the dependence structure to be estimated by copula function separately. Since the pioneering work of Embrechts et al. (1999), copula models have attracted increasing attention due to their ability to model different patterns of dependence while allowing for flexible marginal distributions to capture the skewness and kurtosis in financial returns. The CD-Vine copula (Bedford and Cooke (2001); Bedford and Cooke (2002); Kurowicka and Cooke (2006); Aas et al. (2009); Guégan and Maugis (2010)) implemented in this paper, is a novel type of copula structure recently introduced in finance that allows for very general types of dependence.

The rest of the paper is organized as follows. Section 2 lays out the theoretical model used in this paper: GARCH modeling framework, the Extreme Value Theory, alternative VaR methodology such as RiskMetrics and Historical Simulation, the copula method and vine based pair-copula dependence modeling structure, and VaR and ES backtesting framework. Section 3 empirically investigates the performance of 13 competing VaR and ES estimation approaches for a panel of storable agricultural commodities at 3 different confidence levels (5%, 1%, and 0.1%) for both left tail and right tail. Namely, conventional Normal or Student’s $t$ distribution based models, Historical Simulation,
unconditional EVT, and all their conditional counterparts, multivariate Gaussian or t copula, and GARCH filtered CD-Vine Copula. Section 4 provides the summary comments and concluding remarks.

2.1 Dependence Concepts

Two random variables \((X, Y)\) are dependent if they don’t satisfy the condition of probabilistic independence, i.e. \(F(X, Y) \neq F_1(X)F_2(Y)\). A measure of dependence summarizes the dependence structure of two random variables in a single number. According to Embrechts et al. (2002), a scalar measure of dependence \(\delta(\cdot, \cdot)\) should satisfy the following properties:

1. (Symmetry) \(\delta(X, Y) = \delta(Y, X);\)
2. (Normalization) \(-1 \leq \delta(X, Y) \leq 1;\)
3. (Comonotonic) \(\delta(X, Y) = 1,\)
   (Counter-monotonic) \(\delta(X, Y) = -1;\)
4. For a strictly monotonic transformation \(T: \mathbb{R} \rightarrow \mathbb{R}\)

\[
\delta(T(X), Y) = \begin{cases} \\
\delta(X, Y), & \text{if } T \text{ is increasing} \\
-\delta(X, Y), & \text{if } T \text{ is decreasing} \\
\end{cases}
\]

In general, dependence can be measured using several different concepts: linear correlation, concordance, and tail dependence.
2.1.1 Linear Correlation

The most common measure of dependence between two random variables \((X, Y)\) is the Pearson’s correlation coefficient, a.k.a. the linear correlation coefficient. It is defined as:

\[
\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}
\]

where \(\text{cov}(X, Y)\) is the covariance between \(X\) and \(Y\), \(\text{cov}(X, Y) = E(XY) - E(X)E(Y)\), and \(\sigma_X, \sigma_Y\) is the standard deviation of \(X\) and \(Y\), respectively. The linear correlation is a measure of linear dependence. If two variables are independent, \(\rho(X, Y) = 0\) because \(\text{cov}(X, Y) = 0\); in the case of perfect dependence, i.e. \(Y = aX + b\), \(\rho(X, Y) = \pm 1\); in the case of imperfect dependence, \(-1 < \rho(X, Y) < 1\).

The concept of correlation has gained such wide popularity that people sometimes use the term “correlation” and “dependence” interchangeably. It’s straightforward to calculate the second moments, variance and covariance, to derive the correlation coefficient for most bivariate distributions. Also, linear correlation is invariant with respect to linear transformation of the variables. However, correlation is not an appropriate measure of dependence when returns are not normally distributed. This is because linear correlation is not well-defined for certain heavy-tailed distributions whose second moments do not exist, such as Student’s \(t\) distribution with less than three degrees of freedom. This limitation renders it unsuitable for measuring dependence for financial time series, which often have fat-tails and potentially ill-defined higher moments.

Further, zero linear correlation is a necessary, but not sufficient, condition for independence between two random variables, with the exception of multivariate Normal
distribution. Zero correlation only requires $\text{cov}(X, Y) = 0$, while zero dependence requires $\text{cov}(\Phi_1(X), \Phi_2(Y)) = 0$ for any transformation $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$. Moreover, linear correlation is not invariant to monotonically increasing nonlinear transformation $T: \mathbb{R} \rightarrow \mathbb{R}$, i.e. $\rho(T(X), T(Y)) \neq \rho(X, Y)$. Finally, the linear correlation is not a robust measure in that a single observation can have an arbitrarily high influence on the linear correlation.

2.1.2 Rank Correlation

Recognizing the aforementioned limitations of linear correlation, alternative measures of dependence, such as rank correlation, are considered in the later study (Embrechts, McNeil, and Straumann (1999); (2002)). Rank correlation coefficient measures the correspondence between rankings of two random variables $(X, Y)$ and assesses its significance. Two well-established measures of rank correlation are Spearman’s rho rank correlation (Spearman (1904)) and Kendall’s tau rank correlation (Kendall (1970)).

Let $(X, Y)$ be a random vector with continuous marginal distribution $(F_1, F_2)$. The Spearman’s rho is the linear correlation between $F_1(X)$ and $F_2(Y)$, i.e. $\rho_s = \rho(F_1(X), F_2(Y))$, where $\rho(\cdot, \cdot)$ is the Pearson’s linear correlation. Let $(X_1, Y_1)$ and $(X_2, Y_2)$ be two independent vectors of random variables from the distribution of $(X, Y)$. Kendall’s tau is defined as:

$$\rho_t(X, Y) = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

$$= \Pr(\text{concordance}) - \Pr(\text{discordance})$$
That is, Kendall’s tau is a measure of the difference between the probability that the pairs of \((X, Y)\) are in concordance or in discordance. Both measures are based on the concept of concordance and discordance. The former refers to the property that larger values of one random variable are associated with larger values of another, whereas the latter refers to the property that larger values of one random variable are associated with smaller values of another.

As proved in Embrechts et al. (Embrechts, McNeil, and Straumann (2002)), both \(\rho_s(X, Y)\) and \(\rho_t(X, Y)\) satisfy the property of symmetry, normalization, comonotonicity and counter-monotonicity, and equals zero under independence. Define a copula function \(C = C_{X,Y}\) for the random vector \((X, Y)\). Both \(\rho_s(X, Y)\) and \(\rho_t(X, Y)\) can be expressed in terms of copula as:

\[
\rho_s(X, Y) = 12 \int_0^1 \int_0^1 C(u, v) dudv - 3
\]

\[
\rho_t(X, Y) = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1
\]

The rank correlation measures are superior to the linear correlation because they are invariant under monotonic transformations and can capture a wide variety of dependence. Though rank correlations are not simple functions of moments and hence their computation is more complex, there are some cases when rank correlations are easier to calculate. For multivariate distributions, such as multivariate Normal or multivariate Student’s \(t\) distributions, where first and second moments are well-defined and easily determined, the calculation of linear correlation is straightforward. For other multivariate distributions, which possess a simple closed-form copula such as the Gumbel copula, the
higher moments may be difficult to compute, but the calculation of rank correlation may be easier.

2.1.3 Tail Dependence

Differing from linear correlation and rank correlation, tail dependence measures the level of dependence in the upper right-quadrant tail or the lower left-quadrant tail of a bivariate distribution $(X, Y)$. It measures the conditional probability that one variable exceeds a certain threshold value, given that another variable also exceeds another threshold value. For continuous marginal distributions, tail dependence is a copula property such that it is invariant under monotonically increasing transformation.

Let $X$ and $Y$ be continuous random variables with distribution functions $F(x)$ and $G(y)$, respectively. The upper tail dependence coefficient $\lambda_U$, if exists, is the limit of the conditional probability that $Y$ is greater than the $\alpha$-th percentile of $G(y)$, given that $X$ is greater than the $\alpha$-th percentile of $F(x)$ as $\alpha$ approaches 1, i.e.

$$\lambda_U = \lim_{\alpha \to 1^-} P(Y > G^{-1}(\alpha)|X > F^{-1}(\alpha))$$

If $\lambda_U \in (0,1]$, $X$ and $Y$ are said to be asymptotically dependent in the upper tail; if $\lambda_U = 0$, $X$ and $Y$ are said to be asymptotically independent in the upper tail. Similarly, the lower tail dependence coefficient $\lambda_L$, if exists, is the limit of the conditional probability that $Y$ is smaller than the $\alpha$-th percentile of $G(y)$, given that $X$ is smaller than the $\alpha$-th percentile of $F(x)$ as $\alpha$ approaches 0, i.e.

$$\lambda_L = \lim_{\alpha \to 0^+} P(Y < G^{-1}(\alpha)|X < F^{-1}(\alpha))$$
If $\lambda_L \in [-1,0)$, $X$ and $Y$ are said to be asymptotically dependent in the lower tail; if $\lambda_L = 0$, $X$ and $Y$ are said to be asymptotically independent in the lower tail. When $F(\cdot)$ and $G(\cdot)$ are continuous distributions and the limit exists, then

$$
\lambda_{U} = \lim_{\alpha \to 1^-} \frac{\bar{C}(\alpha, \alpha)}{1 - \alpha}
$$

$$
\lambda_{L} = \lim_{\alpha \to 0^+} \frac{C(\alpha, \alpha)}{1 - \alpha}
$$

where $\bar{C}(u, u) = 1 - 2u + C(u, u)$ denotes the survivor function of the unique copula $C = C_{X,Y}$ associated with $(X, Y)$.

2.2 Extreme Value Theory (EVT) in Risk Management

Extreme Value Theory, similar to the Central Limit Theorem, provides a theoretical framework of analyzing rare events as it specifies what the limiting distributions are as the sample size increases. Broadly speaking, there exist two methods of modeling extrema, namely the block maxima approach and the peaks over threshold (POT) approach. The main difference between these two methods is how the extremes are identified and the principal distribution is used. Under the block maxima approach, the extremes are defined as the maximum data points in the successive periods. Fisher and Tippett (1928) recognized that the limiting distribution of these extremes is the Generalized Extreme Value (GEV) distribution. The POT approach considers the observations that exceed a given threshold. Selecting only the block maxima is not an efficient use of data if there are other extremes available. The threshold methods use data more efficiently and have become a more common method in recent years.
2.2.1 The Block Maxima Approach

The block maxima approach considers the maximum value the variable takes in successive periods. Let $X_1, X_2, \ldots, X_n$ be a sequence of independent, identically distributed (i.i.d.) random variables that come from a distribution function $F(x) = P(X_i \leq x)$, which does not have to be known. The block maxima approach requires grouping the series into non-overlapping successive blocks and identifying the maximum from each block:

$$M_n = \max (X_1, X_2, \ldots, X_n).$$

The limit law of the block maxima is given by the following theorem:

Fisher-Tippett Theorem (a.k.a. The First Theorem in Extreme Value Theory): Let \( \{X_i\}_{i=1,2,\ldots,n} \) be a sequence of i.i.d. random variables. If there exist sequences of constants $c_n > 0, d_n \in \mathbb{R}$ and some non-degenerate distribution function $H$ such that

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H \quad \text{as} \quad n \to \infty$$

Then $H$ belongs to one of the three standard extreme value distributions:

Frechet: $\Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ e^{-x/\alpha}, & x > 0, \alpha > 0 \end{cases}$

Weibull: $\Psi_\alpha(x) = \begin{cases} e^{-x/\alpha}, & x \leq 0 \\ 1, & x > 0, \alpha > 0 \end{cases}$

Gumbel: $\Lambda(x) = e^{-e^{-x}}, x \in \mathbb{R}$

Collectively, these three families of distribution are termed the extreme value distributions. Each family has a location and scale parameter, $d$ and $c$ respectively.
Additionally, the Fréchet and Weibull families have a shape parameter \( \alpha \). The parameter \( \alpha \) is the tail index, which indicates the thickness of the tail of the distribution. The thicker is the tail, the smaller is the tail index. The beauty of this theorem is that these distribution are the only possible limits of the distribution of the extremes \( M_n \), regardless of the distribution \( F \) for the population. In this sense, this theorem provides an extreme value version of the central limit theorem. The shape of the probability density functions for the standard Fréchet, Weibull and Gumbel distributions are shown in Figure 1 in Chapter 1. The density of \( H \) decays polynomially for the Fréchet distribution and therefore the Fréchet distribution suits heavy tailed distribution well, such as Student’s \( t \) distribution.

Jenkinson (1955) and von Mises (1936) suggested that the three of extreme value distribution can be generalized by a one-parameter representation:

\[
H_{\xi}(x) = \begin{cases} 
  e^{-\left(1 + \xi x\right)^{-1/\xi}}, & \text{if } \xi \neq 0 \\
  e^{-e^{-x}}, & \text{if } \xi = 0 
\end{cases} \text{ where } 1 + \xi x > 0
\]

This representation is known as the Generalized Extreme Value (GEV) distribution, where the parameter \( \xi = \alpha^{-1} \). This shape parameter \( \xi \) determines the type of extreme value distribution:

- Fréchet distribution: \( \xi = \alpha^{-1} > 0 \)
- Weibull distribution: \( \xi = \alpha^{-1} < 0 \)
- Gumbel distribution: \( \xi = 0 \)

The biggest criticism of the block maxima approach is that it does not utilize all of the information from the extremes as it considers only the maximum points of the fixed
intervals. Therefore, recent studies on the subject of extreme value theory have concentrated on the behavior of extreme values above a high threshold. This method is the peaks over threshold (POT) approach.

### 2.2.2 The Peaks over Threshold (POT) Approach

The POT approach considers the distribution of the excess over a certain threshold. Let \( \{X_i\}_{i=1,2,...,n} \) be a sequence of i.i.d. random variables with marginal distribution \( F(x) = P(X_i \leq x) \), which does not have to be known. We are interested in estimating the distribution function \( F_u \) of values of \( X_i \) that exceed a certain threshold \( u \). The distribution function \( F_u \) is the conditional excess distribution function defined as:

\[
F_u(y) = P(X - u \leq y | X > u), \quad 0 \leq y \leq x_F - u
\]

Where \( u \) is a given threshold, \( y = X - u \) is termed the excess and \( x_F \leq \infty \) is the right endpoint of the distribution function \( F \). The conditional excess distribution function \( F_u \) represents the probability that the value of \( X \) exceeds the threshold by at most an amount \( y \) given that \( X \) exceeds the threshold \( u \). This conditional probability can be written as:

\[
F_u(y) = \frac{F(u + y) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)}
\]

Due to the limited information available, the estimation of \( F_u \) is not very straightforward. The Fisher-Tippett theorem is the basis for the theorem of peak over threshold. Based on the results of Balkema and de Haan (1974) and Pickands (1975), the distribution of the excess over a high threshold \( u \) can be approximated by the Generalized Pareto Distribution.
Pickands-Balkema and de Haan Theorem (a.k.a. The Second Theorem in Extreme Value Theory): It is possible to find a positive measurable function $\beta$, where $\beta$ is a function of $u$, such that:

$$\lim_{u \to x_F} \sup_{0 \leq x \leq x_F - u} |F_u(x) - G_{\xi, \beta}(x)| = 0$$

If and only if $F \in MDA(H_\xi(x))$. That is, for a large class of underlying distribution $F$, as the threshold $u$ gradually increases, the excess distribution function $F_u$ converges to a Generalized Pareto Distribution.

The Generalized Pareto Distribution (GPD) is the limiting distribution of the peak over threshold approach defined as:

$$G_{\xi, u, \beta}(x) = \begin{cases} 1 - \left(1 + \frac{x - u}{\beta}\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - e^{-\frac{x - u}{\beta}} & \text{if } \xi = 0 \end{cases}$$

with $x \in \begin{cases} [u, \infty) & \text{if } \xi \geq 0 \\ [u, u - \beta/\xi] & \text{if } \xi < 0 \end{cases}$

where $\xi$ is the shape parameter, $\beta$ is the scale parameter, and $u$ is the location parameter.

The Pickands-Balkema and de Haan Theorem implies that, if GEV is the limiting distribution for the block maxima, then the corresponding limiting distribution for threshold excesses is GPD. In addition, the parameters of the GPD of threshold excesses are uniquely determined by those of the associated GEV distribution of block maxima. The shape parameter $\xi$ is equal to that of the corresponding GEV distribution and is
dominant in determining the quantitative behavior of the GPD. Similar to the GEV distribution, the GPD distribution embeds three distributions illustrated by Figure 14:

Figure 14 plots the shape of these distribution functions for illustration purposes. The location parameter \( u \) is set to 0, and the scale parameter \( \beta \) is set to 1. Ordinary Pareto Distribution has the shape parameter \( \xi = \alpha^{-1} > 0 \); Exponential Distribution has the shape parameter \( \xi = 0 \); Pareto type II distribution has the shape parameter \( \xi = \alpha^{-1} < 0 \). In general, financial losses do not have an upper limit. Figure 14 suggests that distributions with shape parameter \( \xi > 0 \) are more suited to model fat-tailed distributions.
An important step in applying POT is to choose an appropriate threshold value $u$. In theory, $u$ should be high enough so that the limiting distribution will converge to the GPD. In practice, the choice of $u$ should allow for enough observations to estimate the parameter. Two approaches are used in this chapter to determine the appropriate threshold value - the mean excess plot, and the maximum likelihood estimation (MLE) of a parametric GPD.

The mean excess plot is a very useful graphical tool for selecting the threshold $u$. The mean excess is the expected value of the excess over a given threshold $u$, given that $u$ is exceeded. The mean excess function $e(.)$ for a random variable $X$ with right endpoint $x_F$ is defined as:

$$e(u) = E(X - u|X > u) \text{ for } u < x_F$$

The mean excess function is better known as the Expected Shortfall in financial risk management. If the underlying distribution $X > u$ has a GPD, then the corresponding mean excess is:

$$e(u) = \frac{\beta + \xi u}{1 - \xi}$$

where $\xi < 1$ so that $e(u)$ is well-defined. That is, the mean excess function is linear in the threshold $u$ when $X > u$ has a GPD. Let $n$ be the number of observations that exceed the threshold $u$. The empirical mean excess function is defined as:

$$e(u) = \frac{\sum_{i=1}^{n} \max(x_i - u, 0)}{\sum_{i=1}^{n} 1_{x_i > u}}$$

---

20 See Embrechts et al. (1997) for a detailed discussion of the properties of this function.
To use the mean excess plot to choose the threshold $u$, one has to look for a threshold $u$ from which the plot presents approximately linear behavior. The mean excess plot will tend towards infinity with a positive slope for a fat-tailed distribution. When the data is exponentially distributed, the plot is a horizontal line. For light-tailed distribution, the plot has a negative slope.

As a preliminary test, using the mean excess plot to select the appropriate threshold is more of an art than a science. A further check on the preliminary conclusion is to estimate the shape parameters using the GPD and look for stability of the parameter estimates as the threshold changes. By the Pickands-Balkema and de Haan Theorem, if the GPD is a reasonable distribution for a threshold, the excesses of a higher level threshold should also follow a GPD with the same shape parameter. Therefore, above a certain level of threshold, the shape parameter should be stable.

Once a threshold level has been selected, the parameters of the GPD can be estimates using several approaches, including maximum likelihood estimation (MLE), method of moments (MOM), biased and unbiased probability weighted moments (PWMB, PWMU), and Bayesian estimation. In this chapter, we use the MLE to estimate the shape parameter $\xi$ and the scale parameter $\beta$. For a high enough threshold $u$ and $n$ excesses of the threshold $(x_1 - u, ..., x_n - u)$, the likelihood function is given by:

$$L(\xi, \beta) = \left\{ \begin{array}{ll} -n\log(\beta) - \frac{1 + \xi}{\xi} \sum_{i=1}^{n} \log \left( 1 + \xi \left( \frac{x_i - u}{\beta} \right) \right), & \text{if } \xi \neq 0 \\ -n\log(\beta) - \sum_{i=1}^{n} (x_i - u), & \text{if } \xi = 0 \end{array} \right.$$
By maximizing the log likelihood function, we can obtain the estimates of the shape parameter $\xi$ and the scale parameter $\beta$.

2.4 Measures of Tail Risk: Value-at-Risk and Expected Shortfall

In the past two decades, Value-at-Risk (VaR) has become the most commonly used tool to measure the downside risk associated with a portfolio. VaR measures the worst expected loss in the market value over a specific time interval at a given significance level. For a given significance level $\alpha$, VaR is defined as the $(1 - \alpha)$-th quantile of the distribution $F$:

$$VaR_{\alpha} = -F^{-1}(1 - \alpha)$$

where $F^{-1}$, the inverse of the distribution function $F$, is the quantile function. The minus sign in front of the quantile function is to transform VaR into a positive number by convention. The existing methods used to calculate VaR can be categorized into three groups: fully parametric methods, non-parametric approaches, and methods based on Extreme Value Theory (EVT). The parametric approach assumes a particular model for the distribution of data, including the variance-covariance method and the assumption of conditional normality (the RiskMetrics method and most models from the ARCH/GARCH family). The non-parametric approach includes historical simulation and Monte Carlo simulation.
2.4.1 Variance-Covariance Method

The classical and the most commonly used approach to estimate VaR is the variance-covariance method. Assume returns series \( \{r_t\}_{t=1,2,...,n} \) follows a martingale process with \( r_t = \mu_t + \varepsilon_t \), where \( \varepsilon_t \) is a white noise with distribution function \( F(\varepsilon_t) \) with zero mean and variance \( \sigma_t^2 \). VaR can be calculated as:

\[
VaR_\alpha = \mu_t + F^{-1}(1 - \alpha)\sigma_t
\]

In practice, the most commonly used distribution function \( F(\varepsilon_t) \) in VaR estimation is the Gaussian distribution for its algebraic simplicity and tractability. However, through the aftermath of the financial crisis, theorists and practitioners have widely criticized as ‘unfit for purpose’ since most financial return time series exhibit the properties of asymmetry and fat-tails. The resulting underestimation of risk has since steered people away from relying solely on the VaR estimates produced by the Gaussian distribution. However, this approach still remains to be the default choice for many VaR-based risk applications since the risk is additive when the return series is based on sample variance assuming normality.

One alternative to account for the fat-tails is to estimate the error term using statistical models such as the family of GARCH models (Engle (1982)). The simplest GARCH(1,1) model is specified as follows:

\[
r_t = \sigma_t z_t, \quad \text{where} \quad z_t \sim i.i.d. N(0,1)
\]

\[
\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2
\]

Although the conditional distribution of the GARCH process is Gaussian, the unconditional distribution exhibits some excess kurtosis.
2.4.2 RiskMetrics

The RiskMetrics\textsuperscript{21} approach is a special, convenient case of the GARCH process. Variances of the error term are modeled using an Exponentially Weighted Moving Average (EWMA) forecast. The EWMA forecast is a weighted average of the previous forecast, with weight $\lambda$, and of the latest squared innovation, with weight $(1 - \lambda)$:

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2$$

where $\sigma_t^2$ is the volatility forecast at time $t$, and $r_t^2$ is the squared return, which acts as the proxy for true volatility. The parameter $\lambda$, a.k.a. the decay factor, determines the relative weights placed on previous period’s forecast. The EWMA model places geometrically declining weights on past observations, assigning greater weights to more recent observations. Note that through backward substitution, the RiskMetrics volatility model can be rewritten as:

$$\sigma_t^2 = (1 - \lambda) \sum_{\tau=1}^{\infty} r_{t-\tau}^2$$

whereby the prediction of volatility is an exponentially weighted moving average of past squared returns.

Although in principle the decay factor $\lambda$ can be estimated, the RiskMetrics approach has shown $\lambda = 0.94$ to be the optimal choice for daily forecast (\textit{J.P. Morgan (1994)}). Therefore, it has become convention to set the decay factor $\lambda$ to be 0.94 in many RiskMetrics based risk applications. This has clear advantage for the volatility forecast of

\textsuperscript{21}The RiskMetrics variance method was first established in 1989, when the chairman of J.P. Morgan, asked for a daily report measuring and explaining the risks of his firm. Nearly four years later in 1992, J.P. Morgan launched the RiskMetrics methodology to the marketplace, making the substantive research and analysis widely available to all market participants.
a large portfolio because no estimation of the decay factor $\lambda$ is necessary. However, similar to the variance-covariance estimation, this approach suffers from the same drawbacks that it’s not able to capture the asymmetry and fat-tails behavior of the financial return time series, which results in the potential underestimation of VaR.

2.4.3 Historical Simulation (HS)

Parametric approaches for volatility forecasts, such as variance-covariance method and RiskMetrics approach, have often been criticized for its applicability and appropriateness regarding its distributional assumptions. Non-parametric approaches, specifically historical simulation (HS) have gained support since it essentially uses only the empirical distribution of the returns. The VaR estimated by HS is simply calculated by the $\alpha$-th quantile of the empirical sample returns distribution.

$$VaR_{t+1,\alpha} = \text{Quantile}\{X_t\}_{t=1}^n$$

Therefore, HS is easy to understand and to implement, while avoiding ad hoc assumptions on the returns distribution. Additionally, the HS allows the flexibility to capture asymmetry, fat-tails and other non-normality characteristics without knowing the underlying return distribution.

However, by construction, the HS approach assumes that all the information regarding the underlying returns distribution is captured by the sample period. Also, it assumes that historical distribution of returns will remain the same over the next periods. These are very strong assumptions because there is no guarantee that the historical sample period chosen is a full representation of the entire return distribution. As a result,
the out-of-sample VaR forecast based on HS may not be reliable when the distribution over the sample period does not represent the population distribution²².

2.4.4 Filtered Historical Simulation (FHS)

In order to remedy some of the shortcomings of the HS approach, we apply the Filtered Historical Simulation (FHS) approach²³, which combines the Historical Simulation and the GARCH model. Specifically, similar to HS approach, FHS does not make any distributional assumption about the standardized returns; it forecasts the variance through the GARCH volatility model. Hence, it is a mixture of parametric and non-parametric statistical estimation. Barone-Adesi and Giannopoulos (2001) demonstrated the superiority of FHS over HS in generating more robust VaR forecasts. The main advantage of FHS is its ability to produce risk measures consistent with the current volatility regime at any arbitrarily high confidence level.

FHS is a two step approach: first fit a GARCH-model to the return series and then use HS to infer the distribution of the residuals. By using the quantile of the standardized residuals, the conditional standard deviation, and the conditional mean forecasts from the GARCH volatility model, the one-step-ahead VaR forecast is given as:

\[ \text{VaR}_{t+1} = \mu_{t+1} + \sigma_{t+1} \text{Quantile}\{X_t^{\text{N}}\} \]

²² For a complete discussion on the use of historical simulation approach for VaR estimation, interested readers can refer to various articles such as Hendricks (1996) and Barone-Adesi et al. (2000).

²³ For a complete discussion on the use of filtered historical simulation approach for VaR estimation, interested readers can refer to various articles such as Hull and White (1998) and Barone-Adesi et al. (1999).
The combination of GARCH filtering and HS alleviate the limitations of classical approaches, since FHS can accommodate the volatility clustering, the observed ‘fat’ tails and the skewness of the empirical distribution.

2.4.5 Extreme Value VaR

Since VaR estimation is focused on the tails of a probability distribution, techniques from EVT are particularly effective. Appealing aspects of EVT have made convincing arguments for its use in calculating VaR. As we have discussed earlier, the modeling of extremes can be done in two ways: the block maxima approach and the peaks-over-threshold approach (POT). The POT approach is generally considered to be more appropriate for practical applications, due to its more efficient use of limited data as all observations above the threshold are utilized to estimate parameters of the tail. The POT method accommodates distributional asymmetry through the separate modeling of the distribution’s left and right tails. Given the observed skewness of most financial returns, separate evaluation of the left and right tails of the returns distribution is advantageous relative to symmetric models, such as Normal distribution or Student’s $t$ distribution. By convention, returns above the threshold $u$, for the left tail of the distribution are examined separately from the observations in the right tail after multiplying by $-1$.

If we fix a sufficiently high threshold $u$, and let $Y_1, \ldots, Y_n$ be the excesses above this threshold, i.e. $Y_i = X_i - u$. By Pickands-Balkema and de Haan Theorem, $F_u(y) = G_{\xi, \beta}(y)$ provided the threshold is sufficiently high. By setting $x = u + y$, an approximation of $F(x)$, for $x > u$, can be written as:
and the function $F(u)$ can be estimated non-parametrically using the empirical CDF:

$$F(u) = \frac{n - N_u}{n}$$

where $N_u$ represents the number of exceedances over the threshold $u$ and $n$ is the sample size. After substituting out the GPD distribution and the empirical CDF, $F(x)$ can be estimated by

$$\overline{F(x)} = 1 - \frac{N_u}{n} \left(1 + \hat{\xi} \left(\frac{x - u}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}}\right)$$

where $\hat{\xi}$ and $\hat{\beta}$ are estimates of $\xi$ and $\beta$, respectively, which can be obtained by MLE. For $\alpha > F(u)$, $VaR_\alpha$ can be obtained from the above equation by solving for $x$

$$VaR_\alpha = u + \hat{\beta} \left[\left(\frac{n}{N_u} (1 - \alpha)\right)^{-\frac{1}{\hat{\xi}}} - 1\right]$$

where $u$ is the threshold, $\hat{\xi}$ and $\hat{\beta}$ are estimates of the shape parameter and scale parameter, respectively. It’s worth mentioning that the main advantage of this unconditional GPD approach is that it focuses attention directly on the tail of the distribution. However, it doesn’t recognize the fact that returns are non-i.i.d.

### 2.4.6 GARCH-EVT Methodology

The Extreme Value Theory (EVT) focuses on the tail distribution of the returns. For that reason, it’s not surprising that the extreme value based VaR is superior to the traditional variance-covariance method and non-parametric approach in estimating extreme risk (Christoffersen (1998)). However, the presence of serial correlation in the
squared returns violates the basic i.i.d. assumption of EVT. In order to address the issue of stochastic volatility, we adopt McNeil and Frey’s approach (McNeil and Frey (2000)) to model agricultural commodity future returns and measure the associated tail risk.

McNeil and Frey’s solution to observed volatility clustering in financial returns is to first fit a GARCH-type volatility model to the returns data by pseudo-maximum likelihood (PML), and apply the EVT to the GARCH residuals in the second stage. The GARCH residuals have been shown to be closer to the i.i.d. assumption than the raw returns series while retaining the characteristic of fat-tails. The advantage of this GARCH-EVT two stage approach lies in its ability to capture conditional heteroscedasticity in the data through the GARCH framework, while at the same time modeling the extreme tail behavior through the EVT method.

Let $P_t$ denote a sequence of commodity cash prices at each moment of time $t$. The continuously compounded log-return is defined as a log-change, i.e. $r_t = \ln(P_t) - \ln(P_{t-1})$. We assume that the dynamics of returns can be represented by:

$$r_t = a_0 + a_0 r_{t-1} + \sigma_t Z_t$$

where $Z_t \sim i. i. d. N(0,1)$

That is, the innovations $Z_t$ is a strict white noise process with zero mean, unit variance, and marginal distribution function $F_Z(z)$. We assume that the conditional variance $\sigma_t^2$ of the mean-adjusted series $\epsilon_t = \sigma_t Z_t = r_t - a_0 - a_1 r_{t-1}$ follows a $GARCH(p, q)$ process:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}$$

where the ARCH term coefficient $\{\alpha_i\}_{i=0,...,q}$ and GARCH term coefficient $\{\beta_j\}_{j=0,...,p}$ are all assumed to be positive to ensure that the conditional variance $\sigma_t^2$ exists and is
always positive. The $GARCH(p, q)$ model is fitted using a pseudo-maximum likelihood (PML) approach, which means that the likelihood for a $GARCH(p, q)$ model with Gaussian innovations is maximized to obtain parameter estimates. The assumption of Gaussian innovations contradicts our belief that financial returns distributions are generally skewed with fat-tails. However, the PML method has been shown to yield a consistent and asymptotically Normal estimator (Gouriéroux (1997)). The standardized residuals can be calculated as:

$$(z_{t-n+1}, \ldots, z_t) = \left( \frac{r_{t-n+1} - \hat{a}_0 - \hat{a}_1 r_{t-n}}{\hat{\sigma}_{t-n+1}}, \ldots, \frac{r_t - \hat{a}_0 - \hat{a}_1 r_{t-1}}{\hat{\sigma}_t} \right)$$

where $\hat{\alpha}_t$ and $\hat{\sigma}_j$ indicates estimated parameters using the PML approach. The one-step ahead forecast for the conditional variance in $t + 1$ is given by:

$$\hat{\sigma}_{t+1}^2 = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \hat{\varepsilon}_{t+1-i}^2 + \sum_{j=1}^{p} \hat{\beta}_j \hat{\sigma}_{t+1-j}^2$$

where $\varepsilon_t = r_t - a_0 - a_1 r_{t-1}$. For stage two of the GARCH-EVT approach, we estimate the tails of the standardized residuals computed in the stage one using EVT. The $\alpha$-th quantile of the innovations is given by:

$$\bar{VaR}_\alpha(Z) = u + \frac{\hat{\beta}}{\xi} \left[ \left( \frac{n}{N_a} (1 - \alpha) \right)^{-\frac{1}{\xi}} - 1 \right]$$

Therefore, for one-day horizon, an estimate of the VaR is:

$$VaR_{\alpha} = \hat{\alpha}_0 + \hat{\alpha}_1 r_t + \hat{\sigma}_{t+1} \bar{VaR}_\alpha(Z)$$
2.4.7 Expected Shortfall (ES)

As the most commonly used risk measure, VaR has been heavily criticized. Firstly, VaR does not indicate the size of the potential loss, given that the loss incurred exceeds VaR. Secondly, as Artzher et al. (1997; 1999) showed that VaR is not a coherent risk measure in that VaR is not sub-additive. That is, the total VaR of a portfolio may be either greater or smaller than the sum of underlying individual VaRs. This poses a problem if the risk management of a commodity portfolio is based on VaR limits of individual products.

A coherent risk measure $\rho(\cdot)$ is defined as one that satisfies the following four properties:

1. (Sub-additivity) $\rho(x) + \rho(y) \leq \rho(x + y)$
2. (Homogeneity) $\rho(tx) = t\rho(x)$, for $t \in \mathbb{R}$
3. (Monotonicity) $\rho(x) \geq \rho(y)$, if $x \geq y$
4. (Translational Invariance) $\rho(x + n) = \rho(x) + n$

To overcome these deficiencies, Artzher et al. (1997; 1999) introduced Expected Shortfall (ES) as an alternative risk measure, which is not only coherent, but also gauges the extent of the loss when VaR is exceeded. The Expected Shortfall is defined as:

$$ES_\alpha = E(r|r > VaR_\alpha)$$

That is, the Expected Shortfall measures the expected value of loss conditional on the fact that a VaR violation occurs. The above expression can be rewritten as:

$$ES_\alpha = VaR_\alpha + E(r - VaR_\alpha| r > VaR_\alpha)$$
The second term of the above equation can be interpreted as the excess distribution $F_{VaR_\alpha}(x)$ over the threshold $VaR_\alpha$. According to the Pickands-Balkema-de Haan Theorem, if the threshold $VaR_\alpha$ is carefully chosen to be high enough, the excess distribution follows GPD. Therefore, the mean of the excess distribution $F_{VaR_\alpha}(x)$ is given by:

$$\frac{1}{1-\xi} [\beta + \xi(VaR_\alpha - u)]$$

where $\xi$ is the shape parameter, $\beta$ is the scale parameter, and $u$ is the location parameter of the GPD. The Expected Shortfall calculated using EVT-based methods can be estimated as:

$$\tilde{ES}_\alpha = \frac{VaR_\alpha}{1-\xi} + \frac{\beta - \xi \hat{u}}{1-\xi}$$

2.5 Backtesting Risk Models

The best way to evaluate the competing VaR approaches is to assess the out-of-sample accuracy of the estimated VaR in forecasting extreme returns. The simplest method is to compare the out-of-sample VaR estimates to the actual realized return in the next period. A violation occurs if the realized return is more negative than the estimated VaR (left tail), or more positive than the estimated VaR (right tail) in any given day. The violation ratio is calculated by dividing the total number of violations by the total number of one-step-ahead forecasts. When forecasting VaR at a certain quantile $\alpha$, we expect the estimated VaR will be violated $100(1-\alpha)\%$ of the time if the model is correctly specified. In other words, the violation ratio converges to $\alpha$ as the sample size increases.
A violation ratio higher than $\alpha$ implies that the model consistently underestimates the return/loss at the tail; while a violation ratio lower than $\alpha$ implies that the model consistently overestimates the return/loss at the tail.

In order to determine whether the frequency of violation is in line with the expected significance level, we use the unconditional coverage test of Kupiec (1995). As Christoffersen (1998) pointed out, the Kupiec test only provides a necessary condition to classify a VaR model as adequate. In the presence of volatility clustering or volatility persistence, the conditional accuracy of VaR estimates becomes an important issue. Christoffersen and Pelletier (2004) proposed a conditional coverage test that jointly investigates (1) whether the number of violations is statistically consistent with the hypothesized number $100(1 - \alpha)\%$; and (2) whether the violations are independently distributed through time. An accurate VaR measure must exhibit both the unconditional coverage property and independence. Accordingly, tests that jointly examine the unconditional coverage and independence properties provide an opportunity to detect VaR measures which are deficient in one way or the other (Christoffersen and Pelletier (2004)). Both the Markov test of Christoffersen (1998) and the duration test of Christoffersen and Pelletier (2004) can be extended to jointly test for both independence and unconditional coverage.

2.5.1 Unconditional Coverage Test

Some of the earliest proposed VaR backtests, e.g. Kupiec (1995), focused exclusively on the property of unconditional coverage. In general, these tests are
concerned with whether or not the reported VaR is violated at $100(1 - \alpha)\%$ of the time. Kupiec (1995) proposed a proportion of failures (POF) test that examines how many times an estimated VaR is violated over a given time horizon. Assume the VaR estimates were accurate, the violations can be modeled as independent draws from a Bernoulli distribution. If the number of violations differs considerably from $100(1 - \alpha)\%$ of the sample, then the accuracy of the underlying VaR model is called into question. Define $N = \sum_{t=1}^{T} I_{t+1}$ as the number of violations over $T$ period, where $I_{t+1}$ is the sequence of VaR violations that can be described as:

Right tail: $I_{t+1} = \begin{cases} 1, & \text{if } r_{t+1} > VaR_{t+1|t} \\ 0, & \text{if } r_{t+1} \leq VaR_{t+1|t} \end{cases}$

Left tail: $I_{t+1} = \begin{cases} 1, & \text{if } r_{t+1} < VaR_{t+1|t} \\ 0, & \text{if } r_{t+1} \geq VaR_{t+1|t} \end{cases}$

The null hypothesis of Kupiec’s unconditional coverage test assumes that the probability of occurrence of the violations $\frac{N}{T}$ equals the expected significance level $\alpha$. That is, the VaR violation series $I_{t}\sim i. i. d. Bernoulli(\alpha)$. Let $p$ be the expected violation rate, i.e. $p = 1 - \alpha$. Using a sample of $T$ observations, the likelihood ratio statistic $LR_{uc}$ is:

$$LR_{uc} = 2 \log \left( \frac{N}{T} \right) - \log(p) \rightarrow \chi^2(1)$$

The null hypothesis is rejected if the likelihood ratio statistic $LR_{uc}$ is larger than the $\chi^2$ critical value at the given significance level.

While the unconditional coverage tests provide a useful benchmark for assessing the accuracy of a given VaR model, these tests suffer inevitably from two shortcomings.

76
Firstly, these tests are known to have difficulty detecting VaR measures that systematically underestimate risk. Moreover, the size of the underestimation can be quite substantial. Secondly, these tests focus exclusively on the unconditional coverage property of a given VaR model without examining if the independence property is satisfied. In other words, the unconditional coverage test implicitly assumes the VaR violations are independent. However, given the stylized fact of volatility clustering and fat-tailed distribution of financial returns, relying solely on the tests of unconditional coverage in backtesting would be problematic.

2.5.2 Conditional Coverage Test

In light of the deficiencies for using unconditional coverage tests to detect the independence property of VaR models, a variety of tests have been developed that explicitly examine the independence property of the VaR violations series \( I_t(\alpha) \). An early and influential test is Christoffersen’s Markov test (1998). The first-order Markov test examines whether or not the likelihood of a VaR violation is observed today depends on if a VaR violation occurred on the previous day. If the VaR measure accurately reflects the underlying market risk, then the chances of today's VaR violation should be independent of yesterday's VaR violation.

The conditional coverage test is a joint test of two properties: correct unconditional coverage and serial independence:

\[
LR_{cc} = LR_{uc} + LR_{ind},
\]

which is asymptotically distributed as a \( \chi^2(2) \) under the null hypothesis of independence. The test is carried out
by creating a $2 \times 2$ contingency table that records VaR violations on adjacent days as in Table 8.

Table 8. Contingency Table for Markov Independence Test

<table>
<thead>
<tr>
<th></th>
<th>$I_t = 0$</th>
<th>$I_t = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{t+1} = 0$</td>
<td>$n_{0,0}$</td>
<td>$n_{1,0}$</td>
</tr>
<tr>
<td>$I_{t+1} = 1$</td>
<td>$n_{0,1}$</td>
<td>$n_{1,1}$</td>
</tr>
<tr>
<td></td>
<td>$n_{0,0} + n_{0,1}$</td>
<td>$n_{1,0} + n_{1,1}$</td>
</tr>
</tbody>
</table>

where $n_{i,j}$ is the number of observations with value $i$ followed by $j$, for $i, j = 0, 1$. $\pi_{i,j} = \frac{n_{i,j}}{\Sigma_{j} n_{i,j}} = \frac{n_{i,j}}{n}$ is the corresponding probability. The values $i, j = 1$ indicate that a VaR violation is observed, while $i, j = 0$ indicate otherwise.

The likelihood ratio test statistic $LR_{ind}$ is specified for the null hypothesis of serial independence against the alternative of first-order Markov dependence. Therefore, the overall likelihood ratio test statistics for the conditional coverage test is:

$$LR_{cc} = -2 \log[(1 - p)^{T-N} p^N] + 2 \log \left[ (1 - \pi_{0,1})^{n_{0,0} \pi_{0,0}} n_{0,1} (1 - \pi_{1,1})^{n_{1,0} \pi_{1,0}} n_{1,1} \right] \xrightarrow{d} \chi^2(2)$$

If the VaR model met the independence property, then the proportion of VaR violation occurs after a previous VaR violation, i.e. $I_{t+1} = 1$, should be the same as the proportion of VaR violation occurs after no violation was observed, i.e. $I_{t+1} = 0$. Otherwise, this calls the validity of the VaR model into question.
A more recent independence test suggested by Christoffersen and Pelletier (2004) cleverly combined the unconditional coverage test and independence test together. They pointed out that if VaR violations are completely independent from each other, the amount of time that elapses between VaR violations should be independent of the amount of time that has elapsed since the last violation. In that sense, the time elapsed between VaR violations should not exhibit any kind of “duration dependence”. Unlike the Christoffersen’s first-order Markov test, this duration based test cannot be constructed simply by computing a $2 \times 2$ contingency table. Define the duration of time (in days) between two VaR violations as $D_i = t_i - t_{i-1}$, where $t_i$ denotes the days of violation number $i$. Under the null hypothesis that the VaR model is correctly specified, the no-hit duration should have no memory and a mean duration of $\frac{1}{p}$ days. The only memory-free continuous random distribution is the exponential distribution. Hence under the null hypothesis, the distribution of the no-hit durations should be $f_{exp}(D; p) = p exp(-pD)$. The alternative parsimonious distribution that allows for duration dependence can be specified as the Weibull distribution, the Gamma distribution, or the Exponential Autoregressive Conditional Duration (EACD) model (Bhattacharyya and Ritolia (2008)). Similarly, a likelihood ratio test is conducted under both the null hypothesis- exponential distribution- and the specified alternative hypothesis for the duration between VaR violations, which must be done using numerical methods. Christoffersen and Pelletier (2004) also showed this test of independence has more power than the Markov test to detect a VaR model that violates the independence property.
2.5.3 Bootstrap Test for the Expected Shortfall (ES)

As the natural analogue to the backtesting for VaR, the backtesting for ES is also based on the difference between the next-day return $r_{t+1}$ and the ES estimate at time $t$ $ES_{p,t+1|t}(r_{t+1})$, conditional on the realized return $r_{t+1}$ exceeding the VaR estimate $VaR_{p,t+1|t}(r_{t+1})$. Assume the return dynamics are given by a location-scale family $r_t = \mu_t + \sigma_t Z_t$, where the innovations $Z_t$ are a strict white noise process (i.i.d.) with zero mean, unit variance and marginal distribution function $F_Z(z)$. Following McNeil and Frey (2000), we define residuals to be:

$$R_{t+1} = \frac{r_{t+1} - ES_{p,t+1|t}(r_{t+1})}{\sigma_{t+1}} = \frac{\mu_{t+1} + \sigma_{t+1} Z_{t+1} - (\mu_{t+1} + \sigma_{t+1} ES_{p,t+1|t}(Z))}{\sigma_{t+1}} = Z_{t+1} - E_t[Z|Z > z_p]$$

Conditional on $r_{t+1} > VaR_{t+1|t}$, or equivalently $Z_{t+1} > z_p$, $z_p$ being the $p$-quantile of $Z$. The $R_t$’s are then i.i.d. with an expected value of zero.

For the entire backtesting period $T$, the empirical version of the residual $R_t$’s on days when violation occurs, i.e. days on which $r_{t+1} < VaR_{t+1|t}$ (left tail) or $r_{t+1} > VaR_{t+1|t}$ (right tail) can be denoted as:

$$\{R_{t+1}: t \in T, r_{t+1} < VaR_{t+1|t} \text{ or } r_{t+1} > VaR_{t+1|t}\}$$

where $R_{t+1} = \frac{r_{t+1} - ES_{t+1|t}}{\sigma_{t+1}}$

Under the null hypothesis that we estimate the dynamics of the process- $\mu_{t+1}, \sigma_{t+1}$ - and the first moment of the truncated innovation distribution $E_t[Z|Z > z_p]$ correctly, these “exceedance residuals” should behave like an i.i.d. sample from a random variable with mean zero. The 1-sided alternative hypothesis is that the residuals have a mean greater than zero, i.e. the ES is systematically underestimated, which as McNeil and Frey (2000)
noted is the more likely direction of failure. To test the hypothesis of mean zero, we use a bootstrap test that makes no assumption about the underlying distribution of the residuals.

2.6 Copula and CD-Vine

Copula modeling has become increasingly popular in many fields of application, because it provides an alternative way by which joint distributions of random variables can be modeled with greater flexibility in terms of marginal distributions and dependence structure. Chen (2006) and Patton (2006) extended the application to the stochastic time series setting, assuming certain conformity conditions are satisfied. This includes in particular stationary Markov processes, of which martingales are a special case. The Markov assumption is hereby satisfied if one takes a two-step method, i.e., first estimate the GARCH model, and then model the copula on the dependence structure of spot and futures market (Power and Vedenov (2008)).

The connection between copulas and joint distributions is established by Sklar’s Theorem (Nelson (2006)), which states that any $n$-dimensional distribution function $\mathcal{F}(\cdot)$ with continuous marginal distributions $F_1(x_1), ..., F_n(x_n)$ can be represented as:

$$\mathcal{F}(x_1, ..., x_n) = \mathcal{C}(F_1(x_1), ..., F_n(x_n)) \quad (6)$$

Then, assuming that the $F_i$’s are differentiable, and that $\mathcal{C}$ and $\mathcal{F}$ are $n$ times differentiable, the derivative of both sides results in

$$\frac{\partial^n \mathcal{F}(x_1, ..., x_n)}{\partial x_1 \cdots \partial x_n} = \frac{\partial^n \mathcal{C}(F_1(x_1), ..., F_n(x_n))}{\partial x_1 \cdots \partial x_n} \times f_1(x_1) \times \cdots \times f_n(x_n) \quad (7)$$

That is, the density of $\mathcal{F}$ has been decomposed as the product of the Copula density, which essentially captures the dependence structure, and a sequence of the univariate
marginal densities as the canonical representation (Cherubini, Luciano, and Vecchiato (2004)).

Note that the estimation of the copula function can be applied to any pair of marginal distributions, not limited to those implied by the original joint distribution. For instance, the Gaussian density can be combined with a beta distribution and a Student’s t distribution to result in a joint density function $h$. This final density function is neither bivariate normal, nor beta, nor Student's t.

One problem often encountered in practical applications is how to identify the Copula function properly, especially in an arbitrarily higher dimensional space. Initially proposed by Joe (1996) and developed in more detail in Bedford and Cooke (2001; 2002), vine is “a flexible graphical model for describing multivariate copulas built up using a cascade of bivariate copulas, so called “pair-copula construction (PCC)” (Aas et al. (2009); Brechmann and Schepsmeier (2013)). Nevertheless, CD-Vine tree structure has a path dependency (see previous chapter for detailed reading). Hence, this leads to the decomposition of a multivariate density, the CD-Vine density (Brechmann and Schepsmeier (2013)):

\[
\begin{align*}
f(x) &= \\
\prod_{k=1}^{\frac{n}{p}} f_k(x_k) \times \\
\prod_{i=1}^{n-1} \prod_{j=1}^{n-i} c_{j,j+i+1}(j+1):(j+i-1) \left( F(x_j|x_{j+1}, \ldots, x_{j+i-1}), F(x_{j+i}|x_{j+1}, \ldots, x_{j+i-1}) \right) \theta_{j,j+i+1}(j+1):(j+i-1)
\end{align*}
\]

where $c$ is the corresponding density of the $n$-dimention Copula function $\mathbb{C}$.

In this paper, a sequential method is chosen to select a CD-Vine copula specification based on Kendall’s tau (Dissmann et al. (2011)).
2.7 Value-at-Risk Monte Carlo Simulation using Pair-Copula Decomposition

In the past decade, many studies in the literature have explored the advantage of applying the copula theory to the estimation of VaR, such as Caillault and Guegan (2005), Palaro and Hotta (2006), Fantazzini (2008), and Bastianin (2009). Although many studies have demonstrated that copula functions can improve VaR estimation, most of these approaches are based on bivariate copulas or very restricted multivariate copula functions, due to a lack of construction schemes for higher dimension copulas. Therefore, the innovative pair-copula estimation of VaR adopted in this chapter provided much needed flexibility and better fit to multivariate financial data than the traditional multivariate copula constructions.

In order to forecast the VaR of the given portfolio for the next day, we need to simulate from the estimated pair-copula decomposition using the algorithm proposed in Aas et al. (2009). For simplicity, we assume that the margins of the distribution are uniform. To simulate from a chosen pair-copula, we first sample \( w_1, ..., w_n \) independent uniform variables. Then set

\[
x_1 = w_1
\]

\[
x_2 = F^{-1}(w_2|x_1)
\]

\[
x_3 = F^{-1}(w_3|x_1, x_2)
\]

\[
\vdots
\]

\[
x_3 = F^{-1}(w_n|x_1, x_2, ..., x_{n-1})
\]

Each \( F(x_j|x_1, x_2, ..., x_{j-1}) \) can be determined using the relationship in (8) recursively for the vine structures. The simulated vine copula observations are then transformed using
the inverse cumulative distribution function of the chosen marginal distribution. The resulting standardized error terms together with the estimated parameters of the GARCH model are used to compute the log returns of each asset in the portfolio. Finally, we calculate the value of the portfolio for each of the simulation and use the empirical quantile function to estimate Value-at-Risk and Expected Shortfall at different confidence levels.

2.8 Empirical Evaluation and Results

2.8.1 Data Description and Preliminary Tests

To study the tail behavior of agricultural commodity futures, we select all futures contracts that are most liquid and widely held. End-of-day prices of the front month futures of corn, soybean, and wheat traded on Chicago Board of Trade (a.k.a. CME Group) have been obtained from Commodity Research Bureau (CRB). The daily returns are defined as \( r_t = \log \frac{p_t}{p_{t-1}} \times 100 \). Figure 15 plots daily prices, returns and squared returns for each series analyzed. Each plot of daily returns confirms the typical empirical time series properties. All series of returns exhibit mean-reverting and volatility clustering, i.e. periods of low volatility were followed by periods of high volatility. The squared daily returns also confirm the observation of volatility clustering with large changes tend to be followed by large changes. For risk management purposes, any measure of risk exposure should be conditional on the current volatility regime.
Basic statistics for the front month futures are summarized in Table 9. As is the case with most financial time series, the front month agricultural commodity futures exhibit evidence of fat tails, with the kurtosis of soybean and wheat far larger than 3. The Jarque-
Bera normality test\textsuperscript{24} also confirms that all returns are not normally distributed. The Augmented Dickey-Fuller test rejects the null hypothesis of a unit root for all series. Among all three agricultural commodity futures, wheat is more volatile than others. The standard deviation of wheat front month futures returns is 50\% higher than the standard deviation of soybean front month futures returns. Interestingly, all front month agricultural commodity futures returns exhibit positive skewness. The squared daily returns shown in Figure 15 exhibit volatility clustering, i.e. large changes tend to be followed by large changes in bursts. The Ljung-Box Q-statistics\textsuperscript{25} reported in Table 9 also reject the null hypothesis of no autocorrelation through 20-lags at a 1\% significance level for all series.

\textsuperscript{24} The Jarque-Bera test statistic JB is defined as $JB = \frac{n}{6} S^2 + \frac{1}{4} K^2$, where $n$ is the number of observations. $S$ is the sample skewness and $K$ is the sample kurtosis. The statistic $JB$ has an asymptotic chi-square distribution with two degrees of freedom and can be used to test the null hypothesis that the data are from a Normal distribution. The null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being 0.

\textsuperscript{25} The Ljung-Box test tests whether any of a group of autocorrelations of a time series are different from zero. The test statistics can be calculated as $Q = n(n + 2) \sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{n-k}$, where $n$ is the sample size, $\hat{\rho}_k$ is the sample autocorrelation at lag $k$, and $h$ is the number of lags being tested. For significance level $\alpha$, the critical region for rejection of the hypothesis of randomness is $Q > \chi^2_{1-\alpha,h}$, where $\chi^2_{1-\alpha,h}$ is the $\alpha$-quantile of the chi-square distribution with $h$ degrees of freedom.
Table 9. Summary Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Corn</th>
<th>Soybean</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start Date</td>
<td>24-Feb-89</td>
<td>24-Feb-89</td>
<td>24-Feb-89</td>
</tr>
<tr>
<td>End Date</td>
<td>15-Jan-10</td>
<td>15-Jan-10</td>
<td>15-Jan-10</td>
</tr>
<tr>
<td>Sample Size</td>
<td>5206</td>
<td>5196</td>
<td>5201</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0061</td>
<td>0.0047</td>
<td>0.0036</td>
</tr>
<tr>
<td>Median</td>
<td>-2.0013</td>
<td>-0.5020</td>
<td>-3.7469</td>
</tr>
<tr>
<td>Max</td>
<td>16.1557</td>
<td>23.2138</td>
<td>37.7102</td>
</tr>
<tr>
<td>Min</td>
<td>-16.3992</td>
<td>-26.6629</td>
<td>-31.4843</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>3.0627</td>
<td>2.4963</td>
<td>3.6590</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.1004</td>
<td>0.0119</td>
<td>0.2466</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.3307</td>
<td>19.3834</td>
<td>6.9454</td>
</tr>
<tr>
<td>Jarque-Bera Test</td>
<td>1189.1816</td>
<td>81571.5028</td>
<td>10528.0869</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>ADF Unit-Root Test</td>
<td>-17.061</td>
<td>-17.374</td>
<td>-18.814</td>
</tr>
<tr>
<td>p-value</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Auto Corr- $r$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>-0.343</td>
<td>-0.328</td>
<td>-0.369</td>
</tr>
<tr>
<td>Lag 5</td>
<td>-0.003</td>
<td>-0.067</td>
<td>-0.043</td>
</tr>
<tr>
<td>Lag 10</td>
<td>-0.031</td>
<td>0.034</td>
<td>0.024</td>
</tr>
<tr>
<td>Lag 20</td>
<td>0.012</td>
<td>0.016</td>
<td>0.006</td>
</tr>
<tr>
<td>Ljung-Box (20)</td>
<td>645.88</td>
<td>831.65</td>
<td>748.49</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Auto Corr- $r^2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>0.259</td>
<td>0.473</td>
<td>0.338</td>
</tr>
<tr>
<td>Lag 5</td>
<td>0.228</td>
<td>0.267</td>
<td>0.143</td>
</tr>
<tr>
<td>Lag 10</td>
<td>0.187</td>
<td>0.263</td>
<td>0.198</td>
</tr>
<tr>
<td>Lag 20</td>
<td>0.171</td>
<td>0.277</td>
<td>0.057</td>
</tr>
<tr>
<td>Ljung-Box (20)</td>
<td>4046.5</td>
<td>10164</td>
<td>2992.8</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

$^{26}$ An intercept term (no trend) is included in the integrating regression for the ADF Unit-Root test.
Given the above test results, the next step is to test the (G)ARCH effect using an ARCH-LM test\(^\text{27}\) on the residuals from a simple OLS regression of each returns series on an intercept and its own lagged values. The test statistics shown in Table 4 suggest the null hypothesis of no ARCH effect is rejected for all series and the null hypothesis of no GARCH effect is also rejected for all series. This result is not sensitive to the number of lags included in the ARCH-LM and GARCH tests.

### Table 10. (G)ARCH Effect Test

<table>
<thead>
<tr>
<th></th>
<th>Corn</th>
<th>Soybean</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Testing for ARCH(1)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-Statistic</td>
<td>350.19</td>
<td>1166.5</td>
<td>595.59</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td><strong>Testing for GARCH(1,1)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>11785.62</td>
<td>10312.33</td>
<td>12694.79</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

---

\(^{27}\) The ARCH-LM test is a Lagrange multiplier (LM) test for ARCH effect in the residuals (Engle 1982). The ARCH-LM test statistic is computed from an auxiliary test regression. To test the null hypothesis that there is no ARCH up to order \(q\) in the residuals, we estimate the model: \(e_t^2 = \lambda_0 + \sum_{i=1}^{q} \lambda_i e_{t-i}^2 + \nu_t\), where \(e\) is the residual obtained from the OLS regression. The F-statistic is an omitted variable test for the joint significance of all lagged squared residuals. The LM test statistic is asymptotically distributed \(\chi^2_q\) under quite general conditions.
Therefore, the GARCH effects need to be filtered out from the data before the extreme value analysis can be applied. At its simplest form, $GARCH(1,1)$ has been proven empirically to be effective in removing the GARCH effect (So and Philip (2006); Degiannakis and Xekalaki (2007)); hereby the rest of the chapter will focus on the residuals of the GARCH-type conditional volatility model fitted to the front month agricultural commodity futures.

### 2.8.2 The Models for the Marginal Distributions

We adopt the Inference Function for Margins (IFM) method due to the large number of parameters in the time-varying models, where the estimation of the parameters is done in two steps. The first step is to estimate the marginal models. The model used for the marginal distribution is $AR(1) - GARCH(1,1) - t$ specified as below:

$$r_t = \mu + ar_{t-1} + \varepsilon_t$$

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2$$

$$\varepsilon_t \cdot \frac{\nu}{\sqrt{\sigma_t^2(\nu - 2)}} \sim iid \ t_n$$

where $r_t$ is the return at time $t$. The parameter estimates and standard errors for the marginal distribution models are presented in Table 11. All parameters for the variance equation are highly significant. The autoregressive effect in the volatility specification is strong as it is in the range of 0.90 (soybean) to 0.92 (corn). In addition, the condition for covariance stationary is satisfied because it appears that the sum of $GARCH(1)$ coefficient $\beta$ and $ARCH(1)$ coefficient $\alpha$ is smaller than 1 for all the series. Note that
$GARCH(1) + ARCH(1)$ is really close to 1, which implies the shocks have persistent effects. Table 11 also presents the ARCH effect test statistics on the residuals. As the ARCH-LM test confirms there are no ARCH effects left in the residuals, the analysis in the rest of the chapter will focus on these standardized residuals.

Table 11. Summary Results for the Marginal Distribution

<table>
<thead>
<tr>
<th></th>
<th>Corn</th>
<th>Soybean</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.0414*</td>
<td>0.0264</td>
<td>0.0122</td>
</tr>
<tr>
<td></td>
<td>(0.0210)</td>
<td>(0.0184)</td>
<td>(0.0262)</td>
</tr>
<tr>
<td>AR(1)</td>
<td>-0.2177*</td>
<td>-0.1642*</td>
<td>-0.2347*</td>
</tr>
<tr>
<td></td>
<td>(0.0143)</td>
<td>(0.0143)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>GARCH Constant</td>
<td>0.0072*</td>
<td>0.0296*</td>
<td>0.0185*</td>
</tr>
<tr>
<td></td>
<td>(0.0031)</td>
<td>(0.0071)</td>
<td>(0.0064)</td>
</tr>
<tr>
<td>GARCH(1)</td>
<td>0.9255*</td>
<td>0.8947*</td>
<td>0.9180*</td>
</tr>
<tr>
<td></td>
<td>(0.0065)</td>
<td>(0.0083)</td>
<td>(0.0082)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.0709*</td>
<td>0.1046*</td>
<td>0.0814*</td>
</tr>
<tr>
<td></td>
<td>(0.0078)</td>
<td>(0.0092)</td>
<td>(0.0095)</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>7.6124*</td>
<td>6.6496*</td>
<td>8.1674*</td>
</tr>
<tr>
<td></td>
<td>(0.7512)</td>
<td>(0.5919)</td>
<td>(0.8345)</td>
</tr>
</tbody>
</table>

Residual ARCH Test  | Corn   | Soybean | Wheat   |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>8.0792</td>
<td>7.8361</td>
<td>3.3273</td>
</tr>
<tr>
<td>p-value</td>
<td>0.7789</td>
<td>0.7978</td>
<td>0.9928</td>
</tr>
</tbody>
</table>

* Indicates significance at the 5% level; standard errors in parentheses

One important exploratory tool used in extreme value analysis is the quantile-quantile plot (QQ-plot). The QQ-plot is a graphical way to determine whether two data sets come from populations with a common distribution. The graph of the quantile makes it easy to informally assess the goodness of fit of the data to the assumed underlying
distribution. Let $X_{n,n} < X_{n-1,n} < \cdots < X_{1,n}$ be the ordered data and let $F$ be the empirical distribution. The graph of quantile is defined by the set of the points:

$$X_{k,n}, F^{-1}\left(\frac{n-k+1}{n+1}\right), k = 1, \ldots, n$$

Basically the QQ-plot compares the quantile of the dataset with the quantile of the reference distribution $F$. A perfect fit shows a $45^\circ$ linear line. If the plot curves to the right, the data has a right tail that is heavier than the reference distribution $F$; on the contrary, if the plot curves to the left, the data has a left tail that is heavier than the reference distribution $F$. This is the case when comparing the original returns series to the Normal distribution function, where the data have heavier tails than expected from a Normal distribution. The second QQ-plot compares the GARCH-filtered residuals to the Normal distribution function, where the Normal distribution fits the standardized residuals visibly better, albeit imperfect either. It suggests the distribution of the GARCH-filtered residuals have slightly heavier tails than the standard Normal distribution; nevertheless it's a much better starting point for the extreme value analysis than the otherwise unfiltered returns.
2.8.3 Exploratory Analysis

In order to form an initial opinion on the correlation and dependence relationship among the agricultural commodity futures, we present conditional correlation measures in Table 12. As expected, the conditional correlation measures matrix indicates there exists a strong dependence between the agricultural commodity futures. Note that
although the underlying assumptions are different, both the linear correlation and Spearman statistics measure the proportion of the variation in one variable that might be considered as being associated with the variation in the other variable. On the other hand, although sharing similar underlying assumptions, the Kendall statistics measure the difference between the probability that the observed data is on the same order (concordance) vs. the probability that the observed data is not of the same order (discordance). Not surprisingly, the pair-wise correlation coefficients between the agricultural commodity futures are significantly different from zero at 5% significance level. In addition, corn is more strongly associated with soybean than the other two pairs.

Table 12. Conditional Correlation Measures Matrix

<table>
<thead>
<tr>
<th></th>
<th>Linear Correlation</th>
<th>Spearman's Rho</th>
<th>Kendall's Tau</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Corn</td>
<td>Soybean</td>
<td>Wheat</td>
</tr>
<tr>
<td>Corn</td>
<td>0.3958*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soybean</td>
<td>0.3973*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p Indicates significance at the 5%; p-values are presented in parentheses
In order to explore the potentially asymmetric behavior between the risk and return, the residual series have been split into right tails and left tails. That is to say, we focus on the computation of the tail-risk measures—such as VaR and ES—for both long and short trading positions in the agricultural commodity market. In the long case, the risk comes from a fall in the price of the agricultural commodity; while vice versa in the short case, the trader loses money when the price increases. Consequently, one focuses in the long case on the left tail of the distribution of returns, and on the right tail of the distribution in the short case. In the left tail case, we switch the sign of the residuals to make them
positive values. Table 13 below lists the number of observations included in each case.

The following empirical analysis will use the split samples.

<table>
<thead>
<tr>
<th></th>
<th>Corn</th>
<th>Soybean</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Corn_Left Tail</td>
<td>Soybean_Left Tail</td>
<td>Wheat_Left Tail</td>
</tr>
<tr>
<td># of Obs.</td>
<td>2713</td>
<td>2609</td>
<td>2588</td>
</tr>
<tr>
<td>Name</td>
<td>Corn_Right Tail</td>
<td>Soybean_Right Tail</td>
<td>Wheat_Right Tail</td>
</tr>
<tr>
<td># of Obs.</td>
<td>2493</td>
<td>2598</td>
<td>2619</td>
</tr>
</tbody>
</table>

2.8.4 Pair-Copula Decomposition Estimation

The major advantage of vine-based pair-copula models is their flexibility in modeling multivariate dependence. They allow for flexible specification of the dependence between different pairs of marginal distributions using bivariate copula functions, while specifying an overall dependence between all of the marginal distributions. We utilize both the C-Vine and the D-Vine structures in this chapter to test the forecasting accuracy of the VaR and ES at different confidence levels. For the three-dimensional case, the full decomposition is:

\[
f(x_1, x_2, x_3) = f_1(x_1) \cdot f_{2\mid 1}(x_2 \mid x_1) \cdot f_{3\mid 1,2}(x_3 \mid x_1, x_2)
\]

\[
= f_1(x_1) \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2) \cdot c_{23\mid 1}(F(x_2 \mid x_1), F(x_3 \mid x_1))
\]

\[
\cdot c_{13}(F_1(x_1), F(x_3)) \cdot f_3(x_3)
\]
That is to say, in a three-dimensional vine copula, there are three different decompositions\(^28\) and each of the three decompositions is both a C-Vine and a D-Vine.

In order to further investigate the dependence relationship between agricultural commodity futures, we map the standardized residuals into the unit square applying the probability-integral transformation\(^29\) to estimate the pair-copulas via maximum likelihood (MLE). The parameters of the marginal distribution and the pair-copula can be estimated simultaneously using MLE; however, the computation becomes very complex very quickly as the dimension of the problem increases. Therefore, we use the Inference Function for Margins (IFM) method, where the parameter estimation is performed in two steps: the parameters in the marginal distributions are estimated first; and the copula parameters are estimated conditioned on the marginal distributions estimates in the previous step.

For a C-Vine decomposition, the log-likelihood is given by:

\[
l(x; \Theta) = \sum_{j=1}^{d-1} \sum_{i=1}^{d-j} \sum_{t=1}^{\tau} \log \left( c_{j, j+i, i+1, \ldots, j-1} \left( F(x_{j,t} | x_{1,t}, \ldots, x_{j-1,t} ), F(x_{j+1,t} | x_{1,t}, \ldots, x_{j-1,t} ) \right) \right)
\]

For a D-Vine decomposition, the log-likelihood is given by:

\[
l(x; \Theta) = \sum_{j=1}^{d-1} \sum_{i=1}^{d-j} \sum_{t=1}^{\tau} \log \left( c_{i, i+j, i+1, \ldots, i+j-1} \left( F(x_{i,t} | x_{i+1,t}, \ldots, x_{i+j-1,t} ), F(x_{i+j,t} | x_{i+1,t}, \ldots, x_{i+j-1,t} ) \right) \right)
\]

\(^28\) In general, for the \(d\)-dimensional case, there are \(n! / 2\) different C-Vines and \(n! / 2\) distinct D-Vines.

\(^29\) If \(X\) is a continuous random variable with \(CDF F_X(x)\) and if \(Y = F_X(X)\), then \(Y\) is a uniform random variable on the interval \([0,1]\).
We consider different copulas described in Chapter 1 for the selection and estimation of pair-copula, namely Gaussian copula, Student’s $t$ copula, Gumbel copula, Clayton copula, Frank copula, and Joe copula. In addition, we need to determine which key variable to use that governs interactions in the dataset at the root of the C-Vine. The key variable is identified by selecting the series that has the strongest dependence relationship with other series, using the C-Vine structure selection criterion described by Czado et al. (2012). As shown in Table 5, corn futures exhibit the strongest dependence among the series. Therefore, we have decided to locate corn futures at the root of the canonical vine. By selecting all further C-Vine root nodes in the similar fashion, the root node order in the data set is determined as corn, soybean, and wheat. The chosen pair-copula decomposition, along with best copula fits and parameter estimates using the full sample for the C-Vine is shown in Figure 18 as an illustration. In the case of the Student’s $t$ copula, the parameter is the degrees of freedom. The chosen D-Vine pair-copula decomposition, along with best copula fits and parameter estimates using the full sample is shown in Figure 19 as an illustration.

![Figure 18. C-Vine Tree Plot of Uniformed GARCH-filtered Residuals](image)

97
Figure 19. D-Vine Tree Plot of Uniformed GARCH-filtered Residuals

2.8.5 Dynamic Backtest

In order to assess the accuracy of the CD-Vine copula with conditional GARCH margins approach and alternative methods of computing VaR and ES, we backtested the methods on a portfolio of equally weighted agricultural commodity futures return series using the following procedure. The competing 13 risk models included in this study are:

1. The traditional unconditional benchmark models on portfolio returns such as the Normal model, the Student’s $t$ model, the Historical Simulation model, and the unconditional Extreme Value Theory model;

2. The conditional GARCH process based models such as RiskMetrics model, the conditional Normal model, the conditional Student’s $t$ model, the Filtered Historical Simulation model, and the conditional Extreme Value Theory model;

3. The multivariate copula model such as the multivariate Gaussian copula and the multivariate Student’s $t$ copula;

4. The pair-copula models such as the CD-Vine copula models based on six different pair-copula families: Gaussian copula, Student’s $t$ copula, Gumbel
copula, Clayton copula, Frank copula, and Joe copula, where the Clayton copula exhibits lower tail dependence while both the Gumbel copula and the Joe copula exhibits upper tail dependence.

The performance of different risk models is evaluated in terms of their one-step-ahead out-of-sample VaR and ES forecasts. For all models, we utilize a rolling sample of 1000 observations to forecast the VaR and ES for both the left tail (long position) and the right tail (short position) at the confidence level of 95%, 99%, and 99.9%. The dynamic nature of this process allows us to capture the time-varying GARCH characteristics of the data in different time periods. For the conditional volatility model, we use $AR(1) - GARCH(1,1) - t$ model consistent with the GARCH-EVT framework; as documented by McNeil and Frey (2000) and Gençay et al. (2003), within the backtesting period, it is practically impossible to examine the fitted model carefully every day and to choose the best parameterization, so suppose that the $AR(1) - GARCH(1,1) - t$ specification is adequate on each rolling window. A similar constraint is also related to the GPD modeling. It's not feasible to visually examine the mean excess plot for each backtesting sample to determine the threshold value for day and choose a different value for $u$.

Therefore, we set the number of exceedance equal to 100-10% of each backtesting sample- as supported by the simulation study conducted by McNeil and Frey (2000) and Marimoutou, Raggad and Trabelsi (2009). As a result, observations above the 90$^{th}$ percentile for the right tail and observations below the 10$^{th}$ percentile for the left tail are estimated using Extreme Value theory. For each copula models, we dynamically re-estimate the GARCH filter parameters for each 1000 observations to obtain the
standardized residuals. Next, the standardized residuals are transformed to standard uniform series using the empirical CDF. Finally, we fit multivariate copula and/or CD-Vine copula on the uniform margins and simulate 50,000 Monte Carlo scenarios to calculate one-step-ahead VaR and ES forecast.

Table 14 summarizes the VaR violation ratios and the competing models are compared and contrasted, along with the unconditional and conditional coverage tests results in Table 15. We show the p-value of each hypothesis test generated using the bootstrap method to alleviate potential bias with respect to the underlying distribution assumption. The null hypothesis of Kupiec's unconditional coverage test assumes the probability of the expected violation occurrence equals the significance level. A p-value above 5% for the unconditional coverage test indicates the number of violations is statistically no different from the significance level. The null hypothesis of Christoffensen's conditional coverage test assumes the probability of VaR violation occurrences equals the significance level and the violations are independently distributed through time. A p-value above 5% for the conditional coverage test indicates the occurrence of VaR violation is statistically independent at the significance level. Table 16 summarizes the bootstrap backtesting results of ES. The null hypothesis of the McNeil and Frey’s ES test assumes ES behaves like an i.i.d. sample from a random variable with mean zero. The 1-sided alternative hypothesis is that the residuals have a mean greater than zero, i.e. the ES is systematically underestimated. A p-value above 5% for the ES test indicates the ES is independently and identically distributed with a zero mean, given the occurrence of VaR violation at the significance level.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>6.92% 2.54% 1.02%</td>
<td>5.28% 1.45% 0.40%</td>
<td>4.38% 1.16% 0.24%</td>
<td>5.18% 0.78% 0.02%</td>
</tr>
<tr>
<td>Student’s t</td>
<td>7.32% 1.28% 0.14%</td>
<td>7.08% 1.57% 0.14%</td>
<td>5.82% 0.90% 0.05%</td>
<td>5.14% 0.78% 0.05%</td>
</tr>
<tr>
<td>Historical Simulation</td>
<td>7.01% 1.95% 0.28%</td>
<td>Filtered Historical Simulation</td>
<td>6.09% 1.38% 0.28%</td>
<td>MV Gaussian Copula</td>
</tr>
<tr>
<td>EVT</td>
<td>7.13% 1.59% 0.31%</td>
<td>Conditional EVT</td>
<td>6.16% 1.33% 0.19%</td>
<td>4.99% 0.76% 0.05%</td>
</tr>
<tr>
<td>Normal</td>
<td>6.54% 2.64% 1.02%</td>
<td>5.66% 1.66% 0.50%</td>
<td>4.80% 1.23% 0.29%</td>
<td>4.99% 0.76% 0.05%</td>
</tr>
<tr>
<td>Student’s t</td>
<td>7.08% 1.57% 0.14%</td>
<td>Conditional Normal</td>
<td>8.13% 2.59% 0.64%</td>
<td>5.87% 0.95% 0.02%</td>
</tr>
<tr>
<td>Historical Simulation</td>
<td>7.44% 1.78% 0.28%</td>
<td>Conditional Student’s t</td>
<td>5.83% 1.02% 0.10%</td>
<td>C-Vine Copula</td>
</tr>
<tr>
<td>EVT</td>
<td>7.13% 1.93% 0.21%</td>
<td>Filtered Historical Simulation</td>
<td>6.78% 1.52% 0.26%</td>
<td>5.11% 0.81% 0.05%</td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>5.28% 1.45% 0.40%</td>
<td>Condition EVT</td>
<td>6.56% 1.47% 0.19%</td>
<td>D-Vine Copula</td>
</tr>
</tbody>
</table>
Table 15. Bootstrap Unconditional/Conditional Coverage Test

<table>
<thead>
<tr>
<th></th>
<th>5.00% Uncond.</th>
<th>5.00% Cond.</th>
<th>1.00% Uncond.</th>
<th>1.00% Cond.</th>
<th>0.10% Uncond.</th>
<th>0.10% Cond.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Left Tails</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>0.007</td>
<td>0.006</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Student’s t</td>
<td>0.001</td>
<td>0.001</td>
<td>0.224*</td>
<td>0.209^</td>
<td>0.400*</td>
<td>-</td>
</tr>
<tr>
<td>Historical Simulation</td>
<td>0.003</td>
<td>0.006</td>
<td>0.003</td>
<td>0.004</td>
<td>0.063*</td>
<td>0.000</td>
</tr>
<tr>
<td>EVT</td>
<td>0.002</td>
<td>0.003</td>
<td>0.057*</td>
<td>0.051^</td>
<td>0.048</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Panel A: Traditional Benchmark Models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Right Tails</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>0.016</td>
<td>0.012</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Student’s t</td>
<td>0.002</td>
<td>0.003</td>
<td>0.049</td>
<td>0.036</td>
<td>0.411*</td>
<td>-</td>
</tr>
<tr>
<td>Historical Simulation</td>
<td>0.000</td>
<td>0.000</td>
<td>0.013</td>
<td>0.010</td>
<td>0.065*</td>
<td>0.000</td>
</tr>
<tr>
<td>EVT</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>0.003</td>
<td>0.181*</td>
<td>0.005</td>
</tr>
<tr>
<td><strong>Panel B: GARCH Processes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>0.419*</td>
<td>0.060^</td>
<td>0.064*</td>
<td>0.070^</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>Conditional Normal</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Conditional Student’s t</td>
<td>0.449*</td>
<td>0.005</td>
<td>0.281*</td>
<td>0.416^</td>
<td>0.464*</td>
<td>-</td>
</tr>
<tr>
<td>Filtered Historical Simulation</td>
<td>0.030</td>
<td>0.004</td>
<td>0.096*</td>
<td>0.134^</td>
<td>0.059*</td>
<td>0.001</td>
</tr>
<tr>
<td>Conditional EVT</td>
<td>0.019</td>
<td>0.003</td>
<td>0.150*</td>
<td>0.138^</td>
<td>0.241*</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Panel C: Multivariate Copula Models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MV Gaussian Copula</td>
<td>0.177*</td>
<td>0.104^</td>
<td>0.354*</td>
<td>0.319^</td>
<td>0.118*</td>
<td>-</td>
</tr>
<tr>
<td>MV Student’s t Copula</td>
<td>0.107*</td>
<td>0.039</td>
<td>0.415*</td>
<td>0.382^</td>
<td>0.334*</td>
<td>-</td>
</tr>
<tr>
<td><strong>Panel D: Pair-copula Models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-Vine Copula</td>
<td>0.435*</td>
<td>0.133^</td>
<td>0.250*</td>
<td>0.333^</td>
<td>0.170*</td>
<td>-</td>
</tr>
<tr>
<td>D-Vine Copula</td>
<td>0.458*</td>
<td>0.147^</td>
<td>0.244*</td>
<td>0.339^</td>
<td>0.338*</td>
<td>-</td>
</tr>
</tbody>
</table>

* Indicates fail to reject $H_0$ at the 5%: the number of VaR violations is statistically no different than the significance level

^ Indicates fail to reject $H_0$ at the 5%: the occurrence of VaR violations is statistically independent of each other
Table 16. Bootstrap Backtesting Results of Expected Shortfall

<table>
<thead>
<tr>
<th></th>
<th>5.00%</th>
<th>1.00%</th>
<th>0.10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p Value</td>
<td>p Value</td>
<td>p Value</td>
</tr>
<tr>
<td><strong>Left Tails</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel A: Traditional Benchmark Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Student’s t</td>
<td>0.630*</td>
<td>0.455*</td>
<td>0.011</td>
</tr>
<tr>
<td>Historical Simulation</td>
<td>0.117*</td>
<td>0.703*</td>
<td>0.372*</td>
</tr>
<tr>
<td>EVT</td>
<td>0.223*</td>
<td>0.275*</td>
<td>0.397*</td>
</tr>
<tr>
<td>Panel B: GARCH Processes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>0.999*</td>
<td>0.999*</td>
<td>0.063*</td>
</tr>
<tr>
<td>Conditional Normal</td>
<td>0.999*</td>
<td>0.999*</td>
<td>0.036</td>
</tr>
<tr>
<td>Conditional Student’s t</td>
<td>0.999*</td>
<td>0.999*</td>
<td>0.948*</td>
</tr>
<tr>
<td>Filtered Historical Simulation</td>
<td>0.424*</td>
<td>0.516*</td>
<td>0.831*</td>
</tr>
<tr>
<td>Conditional EVT</td>
<td>0.509*</td>
<td>0.448*</td>
<td>0.525*</td>
</tr>
<tr>
<td>Panel C: Multivariate Copula Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MV Gaussian Copula</td>
<td>0.000</td>
<td>0.003</td>
<td>0.066*</td>
</tr>
<tr>
<td>MV Student’s t Copula</td>
<td>0.065*</td>
<td>0.489*</td>
<td>0.953*</td>
</tr>
<tr>
<td>Panel D: Pair-copula Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-Vine Copula</td>
<td>0.067*</td>
<td>0.404*</td>
<td>-</td>
</tr>
<tr>
<td>D-Vine Copula</td>
<td>0.048</td>
<td>0.314*</td>
<td>0.618*</td>
</tr>
<tr>
<td><strong>Right Tails</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel A: Traditional Benchmark Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>Student’s t</td>
<td>0.038</td>
<td>0.988*</td>
<td>0.013</td>
</tr>
<tr>
<td>Historical Simulation</td>
<td>0.216*</td>
<td>0.235*</td>
<td>0.339*</td>
</tr>
<tr>
<td>EVT</td>
<td>0.057*</td>
<td>0.530*</td>
<td>0.142*</td>
</tr>
<tr>
<td>Panel B: GARCH Processes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>0.999*</td>
<td>0.999*</td>
<td>0.014</td>
</tr>
<tr>
<td>Conditional Normal</td>
<td>0.999*</td>
<td>0.999*</td>
<td>0.061*</td>
</tr>
<tr>
<td>Conditional Student’s t</td>
<td>0.999*</td>
<td>0.999*</td>
<td>0.638*</td>
</tr>
<tr>
<td>Filtered Historical Simulation</td>
<td>0.658*</td>
<td>0.541*</td>
<td>0.301*</td>
</tr>
<tr>
<td>Conditional EVT</td>
<td>0.565*</td>
<td>0.362*</td>
<td>0.115*</td>
</tr>
<tr>
<td>Panel C: Multivariate Copula Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MV Gaussian Copula</td>
<td>0.000</td>
<td>0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>MV Student’s t Copula</td>
<td>0.006</td>
<td>0.196*</td>
<td>-</td>
</tr>
<tr>
<td>Panel D: Pair-copula Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-Vine Copula</td>
<td>0.011</td>
<td>0.189*</td>
<td>-</td>
</tr>
<tr>
<td>D-Vine Copula</td>
<td>0.010</td>
<td>0.234*</td>
<td>-</td>
</tr>
</tbody>
</table>

* Indicates fail to reject $H_0$ at the 5%: the excess conditional expected shortfall (excess of the actual series when VaR is violated), is i.i.d. and has zero mean
It is clear that except for the copula-based models, the majority of the risk models evaluated underestimate the tail risk. That is, the corresponding VaR violation ratios at the specified significance levels at both tails are higher than expected. This issue is more pronounced for traditional unconditional models, such as Normal, Student’s $t$, Historical Simulation, and Extreme Value Theory (EVT) models, rather than their corresponding conditional counterparts. The models with GARCH filters— which take autoregressive conditional heteroscedasticity into consideration—tend to perform better than those that do not allow for it. In addition, the copula-based models, including the multivariate Gaussian/Student’s $t$ copula and CD-Vine copula, perform significantly better than traditional unconditional models and conditional GARCH-based models at both tails, especially at higher confidence levels ($\alpha = 1.00\%, 0.10\%$). Finally, the ES backtesting results reinforces the VaR backtesting results in that it confirms when VaR violation occurs, the conditional models perform better than unconditional ones. Note that the performance of the copula-based models is very comparable to that of the Conditional Student’s $t$ model. This shouldn’t come as a shock since the $AR(1) – GARCH(1,1) – t$ conditional process is the marginal model used for all copula-based risk models.

Among traditional benchmark models, Student’s $t$ and EVT model generally outperforms Normal and Historical Simulation model across significance levels for both tails. This is demonstrated by lower VaR violation ratio, as well as the passing of Kupiec’s unconditional coverage test and Christoffensen’s conditional coverage test at higher confidence levels ($\alpha = 1\%$ and $0.1\%$). However, one interesting observation is at 5% significance level, Normal model actually generates the lowest VaR violation ratio,
albeit not passing either Kupiec’s nor Christoffensen’s test; this seemingly “outperformance” reverses by a wide margin at higher confidence levels ($\alpha = 1.00\%, 0.10\%$). Moreover, when we overlay the ES backtesting results, EVT is favored slightly over the Student’s $t$ model.

Among the conditional GARCH-based models, Conditional Student’s $t$ model clearly outperforms its peers in terms of VaR violation ratio and unconditional/conditional coverage tests. However, the industry standard RiskMetrics has commendable performance at significance level of 5% and 1%, but not at the highest significance level of 0.1%. On the contrary, the Filtered Historical Simulation model and Conditional EVT model outshined their peers—except for Conditional Student’s $t$ model—at higher significance level of 1% and 0.1%. By construction, these two models were able to take into account of recent volatility scheme and better capture the embedded extreme tail risk. The ES backtesting results also offer consistent relative model performance ranking as the VaR backtesting results.

Among the multivariate copula models, the multivariate Student’s $t$ copula generally outperforms the multivariate Gaussian copula across significance levels for both tails. However, the same observation echoes the earlier Normal vs. Student’s $t$ model comparison. At 5% significance level, the multivariate Gaussian copula model actually generated lower VaR violation ratio than the multivariate Student’s $t$ copula. As anticipated, this “advantage” quickly fades away at higher significance levels of 1% and 0.1%. The ES backtesting results also confirm the multivariate Student’s $t$ copula as the front runner.
Among the CD-Vine copula models, C-Vine copula and D-Vine copula has very comparable performances across significance levels for both tails, though D-Vine copula has shown a minor upper hand. Additionally, the ES backtesting results offer consistent performance ranking as the VaR backtesting results, i.e. C-Vine copula performs on par vs. the D-Vine copula. This is to be expected though as demonstrated by Figure 18 and Figure 19 in the prior section of pair-copula decomposition estimation. The "star" structure of C-Vine and "path-dependency" feature of D-Vine is not very distinctive in a three-dimensional case. Finally, both C-Vine copula and D-Vine copula outperforms the multivariate Gaussian copula and Student’s $t$ copula.

2.9 Conclusion

Most traditional risk management methods are based on the assumption that the risk factors are multivariate normally distributed. However, previous empirical evidence has shown the financial return distributions are asymmetric with tail dependence. In this chapter, we have proposed a novel method to forecast VaR and ES of a multivariate agricultural commodity futures portfolio based on the theory of CD-Vine copula. A comprehensive empirical analysis is conducted on a panel of 13 competing risk models on a portfolio of equally weighted Corn, Soybean, and Wheat futures at 5%, 1%, and 0.1% significance levels for left tail (long position) and the right tail (short position). The competing risk models evaluated are: traditional unconditional benchmark models such as the Normal model, the Student’s $t$ model, the Historical Simulation model, and the Extreme Value Theory model; the conditional GARCH process based models such as
RiskMetrics model, the conditional Normal model, the conditional Student’s $t$ model, the Filtered Historical Simulation model, and the conditional Extreme Value Theory model; the multivariate copula models such as the multivariate Gaussian copula and the multivariate Student’s $t$ copula; the pair-copula models such as the CD-Vine copula models based on six different pair-copula families: Gaussian copula, Student’s $t$ copula, Gumbel copula, Clayton copula, Frank copula, and Joe copula. Kupiec’s and Christoffensen’s unconditional and conditional coverage test and McNeil and Frey’s ES test is used to assess the relative performance of the GARCH-filtered Copula VaR/ES estimation approach against the alternative risk models using Monte Carlo simulation.

The present paper proposes the use of CD-Vine copulas as a means of estimating multivariate GARCH-based models without imposing the restrictive assumption of joint-normality on the marginal distribution. With the careful characterization of the GARCH marginals, taking into account of the non-elliptical higher order moment relationships, the dynamic estimation outcomes from the proposed Copula approach were able to accommodate much more flexible dependence structure with the time varying volatility. However, there are several interesting side findings that shouldn’t be overlooked. Firstly, it clearly holds that conditional risk models fare better in an extended backtesting period versus unconditional models; however, at 1% and 0.1% significance levels, the backtesting results of simple Student’s $t$ model is very comparable to that of the conditional models. Second, at loose significance level (5%), the VaR violation ratio may not be the best judge of model performance; for instance, both Filtered Historical Simulation and Conditional EVT produce higher VaR violation ratios which fail
Kupiec’s and Christoffensen’s test, this observation reverses at 1% and 0.1% confidence level. Lastly, for practitioners with limited computational resources, conditional Student’s $t$ model is also a viable option with comparable performance and only a fraction of computational efforts.
References


Application with Energy Futures (Doctoral dissertation).


Appendix A: The Reflection Principle

For every sample path of $X_t$ that hits level $y$ before time $t$ but finishes below level $x$ at time $t$, there is another equally probable path that hits $y$ before $t$ and then travels upward at least $y - x$ units to finish above level $y + (y - x) = 2y - x$ at time $t$. Thus

$$P\{X_t \leq x, M_t > y\} = P\{X_t \geq 2y - x\} = P\{X_t \leq x - 2y\} = \Phi\left(\frac{x - 2y}{\sqrt{t}}\right)$$
Appendix B: Sklar’s Theorem

Let $\mathcal{F}$ be an $n$-dimensional distribution function with margins $F_1, \ldots, F_n$. Then there exists an $n$-dimensional Copula $\mathcal{C}$ s.t. for all $x$ in $\mathbb{R}^n$,

$$\mathcal{F}(x_1, \ldots, x_n) = \mathcal{C}(F_1(x_1), \ldots, F_n(x_n)) \quad (10)$$

If $F_1, \ldots, F_n$ are all continuous, then $\mathcal{C}$ is uniquely determined on the space $\text{Range}(F_1) \times \ldots \times \text{Range}(F_n)$. Conversely, if $\mathcal{C}$ is an $n$-dimensional Copula and $F_1, \ldots, F_n$ are probability distribution functions, the function $\mathcal{F}$ defined above is an $n$-dimensional distribution function with margins $F_1, \ldots, F_n$.\[30]

\[30\] The full proof of the Sklar’s Theorem is subsumed in here. Interested readers can refer to Nelson (2006) for complementary reading.