ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis [FEB61008||2022 Quantitative Finance]

Beating the naive hedge for oil refineries

Student

E.D. Wessel

Student ID number

446744

Supervisor

Dr. C. Zhou

Second Assessor

Dr. H.J.W.G. Kole

April 18, 2023

Abstract

Oil refineries are exposed to price movements of crude oil, gasoline and heating oil during the refining process. This paper studies the multi-product hedging of downside risk for oil refineries with futures contracts for the period $1991 - 2020$. Herein, we compare our benchmark, the naive hedge, against three hedging frameworks: the single commodity, fixed proportion and flexible multi-commodity hedging frameworks. The latter two hedging frameworks model use a vine copula approach which captures the characteristics of price changes as well as the dependency structures between the prices of oil products. We propose a mixture R-vine copula model and compare it against three vine copula classes through an out-of-sample test. Our results show that the fixed proportion hedging framework achieves significantly more risk reduction than the naive hedge for seven of eight risk measures. Furthermore, we find that the R-vine copula model is preferred for the fixed proportion hedging framework, whereas for the flexible multi-commodity hedging framework our proposed mixture R-vine copula model is preferred for most risk measures.

Keywords: *Vine Copula*; *Multi-product hedging*; *Hedging framework*; *Optimal hedge ratio*; *Risk measures*

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Contents

1 Introduction

Towards the end of 2022, inflationary price changes became a more prominent news topic on the back of the COVID-19 crisis and Ukraine war, which led to increased commodity prices, especially in the energy sector. Moreover, in general, product prices became more volatile due to supply chain issues as well as personnel shortages. Manufacturing companies therefore were more exposed to the multi-product problem. This problem entails the risk of not being able to sell their manufactured product for the price imagined at the moment of procurement of raw materials.

This study focuses on the downside risk for oil refineries. More specifically, we focus on the conversion of three barrels of crude oil to two barrels of gasoline and one barrel of heating oil (Hale et al. (2002)). This $3:2:1$ ratio is also known as the crack spread. An oil refinery typically buys and sells futures of these commodities as hedging instruments. A futures contract is a financial contract to buy or sell an asset or commodity at a specified future date for a specified price. If, for example, a manufacturer shorts a futures contract, while the price of its manufactured goods decreases over time, then the manufacturer still receives the pre-determined price in the futures contract. As such, the profit margin of the manufacturer depends on: spot prices of its input materials and output products as well as futures prices changes. Herein, the quantity of futures purchased relies on the determined hedge ratios. Hence, improved hedge ratios can reduce the oil refinery's downside risk of its profit margin. Downside risk of the profit margin is the risk that spot and futures prices shift in such a way that the profit of an oil refinery decreases and becomes negative.

In this paper, we examine and compare the hedge effectiveness of three hedging frameworks: the single commodity, the fixed proportion and the flexible multi-commodity hedging frameworks, against our benchmark hedging strategy: the naive hedge. Within the fixed proportion and flexible multi-commodity hedging frameworks, we study which vine copula model achieves the most risk reduction and we propose a mixture vine copula model. For our study, we use the spot and futures prices of crude oil, gasoline and heating oil from 1987 to 2022.

The single commodity hedging framework is a simple hedging solution, where a commodity consumer generally takes a long futures position and a commodity producer a short position concerning one commodity (Pindyck (2001)). This framework disregards the possible predicting power of similar trends in the prices of other input and output products of the manufacturing process. It determines the hedge ratios through modelling the dependency relation between each commodity spot price and futures price with bivariate copulae. A bivariate copula is bivariate cumulative distribution function that defines the dependency structure between two random variables. An advantageous characteristic of a copula is that it can capture tail dependency between variables, which helps measuring the exposure to extreme events. With the 2-dimensional distribution drawn from bivariate copulae, we determine the hedge ratios which minimise eight downside risk measures based on Value-at-Risk (VaR), Expected Shortfall (ES) and Lower Partial Moments (LPM).

The fixed proportion hedging framework stems from the $3:2:1$ ratio which is typically used for production process oil refineries. In 1994, following this rationale a crack spread futures contract was

introduced on the New York Mercantile Exchange (NYMEX) to lower margin costs, as it would concern a single trade to hedge three commodities. To determine the single hedge ratio within this framework, we model the joint distribution of all six variables, meaning both the spot and futures prices of crude oil, gasoline and heating oil, with a vine copula approach. A vine copula is an approach to define multivariate distributions based on a constellation of bivariate copulae. After attaining the 6-dimensional distribution, we minimise the downside risk of the profit margin of the manufacturer through one hedge ratio, restricting that the purchased futures per commodity are a $3:2:1$ multiple of this hedge ratio.

We apply a flexible multi-commodity hedging framework, which also relies on the 6-dimensional distribution through the implementation of a vine copula. The difference with the fixed proportion hedging framework is that the flexible multi-commodity hedging framework relaxes the fixed $3:2:1$ proportion constraint. The flexible multi-commodity hedging framework, optimises three hedge ratios to minimise downside risk measures. We expect that the flexible multi-commodity hedging framework can hedge most effectively, as it combines the free three hedge ratios with modelling the joint distribution. Hedge effectiveness is the decrease in downside risk of the hedged profit margin over the downside risk of the unhedged profit margin.

Lastly, we compare our hedging frameworks against a generally hard to beat hedging strategy, namely the naive hedge (Wang et al. (2015)). The naive hedging strategy hedges all its positions by purchasing a futures contract equal to its production size. We test the hedging frameworks in an out-of-sample test. Herein, we examine the hedge ratios over the test windows running from 1991 to 2020.

Besides comparing the four hedging frameworks, we determine the best vine copula models within the fixed proportion and flexible multi-commodity frameworks. Herein, we evaluate our proposed mixture R-vine model against canonical (C-), drawable (D-) and regular (R-) vine models. These models are subclasses of vine copulae. In our proposed mixture vine copula model, we combine at most three bivariate copulae with a comparable fit to capture the dependency of a single variable pair in a vine copula construction. As such, we diversify the fit of the vine copula model and aim to improve the approximation of the joint distribution and subsequently the estimation of hedge ratios.

Contrary to our expectations, our results show that the fixed proportion hedging framework achieves the significantly highest mean hedge effectiveness for 7 of the 8 downside risk measures. Herein, the fixed proportion is the only hedging framework to significantly outperform the naive hedging framework. Additionally, we find that our proposed mixture R-vine copula model is the preferred model within the flexible multi-commodity hedging framework, while the R-vine copula model is preferred within the fixed proportion hedging framework.

Section 2 entails the existing literature regarding the topic and the relevance of our study. In Section 3 we report the dataset used for our research. In Section 4, we clarify the general framework of copulae and explain the methods used for our research. The results of our research are presented in Section 5. Lastly, the concluding remarks can be found in Section 6.

2 Literature

In this section we discuss the relevance and contribution of our research with respect to literature.

The multi-product problem has been studied extensively in risk management, where the oil refinery case is a primary example. Other studied multi-product hedging problems are encountered in the soybean and kettle industry (Awudu et al. (2016)) as well as in corn-based ethanol production (Dahlgran (2009)).

Initially, the hedging solutions for oil refineries were derived and applied within the mean-variance framework (Haigh & Holt (2002), Alexander et al. (2013)). Over time, new methods for risk management surfaced such as multivariate copulae. Joe (1996), Bedford & Cooke (2001) and Bedford & Cooke (2002) introduced the idea of pair-copula decomposition, which was the basis for vine copulae. Hereafter, Aas et al. (2009) implemented an algorithm to assign a unique copula type to each pair copula and an estimation procedure for small subclasses of vine copulae, namely C- and D-vine copulae. Estimation of the full vine copula class became more practical with the introduction of the sequential estimation algorithm in Dissmann et al. (2013). Moreover, Dissmann et al. (2013) have shown that R-vine models can result in a better fit than C- and D-vines, due to a greater flexibility.

Vine copulae are constellations of many bivariate copulae. Generally, one copula type is used to describe the dependency between two variables within a vine copula (Aas et al. (2009), Dissmann et al. (2013)). However, it is also possible to express a dependency relation between two variables by two or more copula types to specify the dependency structure more accurate. Existing papers have implemented mixtures in vine copulas already through the computationally expensive EM-algorithm to reveal hidden dependence patterns in multivariate data (D. Kim et al. (2013)) or to increase flexibility (Weiß & Scheffer (2015)).

Sukcharoen & Leatham (2017) and Liu et al. (2017) are closely related to our study. Both works study the hedging performance of copula models for the petroleum complex. The former compares the hedging performance of the C- and D-vine copulae against more rigid standard multivariate copulae for the weekly price changes during 1987-2015. Herein, the authors minimise eight measures of downside risk and find that the D-vine model achieves the best fit and hedging results within a flexible multi-commodity framework. The latter research focuses on the daily hedging performance between 2012-2016 of a fixed proportion hedging framework against a flexible multi-commodity hedging framework. This joint distribution is determined by means of a kernel copula approach, which fits the joint distribution non-parametrically. The authors find that the relaxation of the crack spread constraint improves the daily hedging performance for the time period 2012-2016. Furthermore, Ji & Fan (2011) apply a DCC-ECM-MVGARCH model that outperforms the weekly naive hedge for the oil refinery case. However, in this paper hedge effectiveness is defined as the reduction in variance.

By contrast, we contribute to the literature by hedging the petroleum complex with three different hedging strategies with the full range of copula models. We have found a hedging strategy that reduces more risk on average for seven downside risk measures, including Value-at-Risk, Expected Shortfall and Lower Partial Moments, than the naive hedge. Moreover, we have been able to deduce the added value of modeling

the joint distribution with respect to the bivariate distribution. Lastly, we propose an ad-hoc approach for a mixture vine copula model, which increases hedging effectiveness for the majority of the downside risk measures within the flexible multi-commodity framework.

3 Data

In this section we discuss and analyse the data and its characteristics used for our research. To study the hedging of input and output products of an oil refinery, we use the weekly spot and futures price changes of crude oil, gasoline and heating oil. More specifically, the time series used for the spot price (S^C) and futures price (F^C) of crude oil is the West Texas Intermediate (WTI) crude oil at Cushing, Oklahoma. The spot price of gasoline (S^G) and futures price of gasoline (F^G) are derived from the regular unleaded gasoline time series at New York Harbor. From 2006 onwards, we replace the futures prices of gasoline with the Reformulated Blendstock for Oxygen Blending (RBOB) due to the lack of data availability. Lastly, the spot price of heating oil (S^H) and futures price of heating oil (F^H) are represented by No. 2 heating oil at New York Harbor. Herein, gasoline and heating oil prices have been converted to barrel units.

The data is retrieved from the Datastream database, which started recording price positions in 1986. Our dataset covers the period from 7 January 1987 until 9 February 2022. Moreover, futures prices have been constructed by taking the prices of nearest expiring futures contract and rolling them over at expiry date. If a futures price is not available on Wednesday, we take the available price from the nearest day before. In Figure 1, we observe stable spot prices in the first half of our dataset and an increase in volatility from 2004 onwards. We also observe a discrepancy between crude oil and the other petroleum products after 2010, when crude oil trades at a lower price.

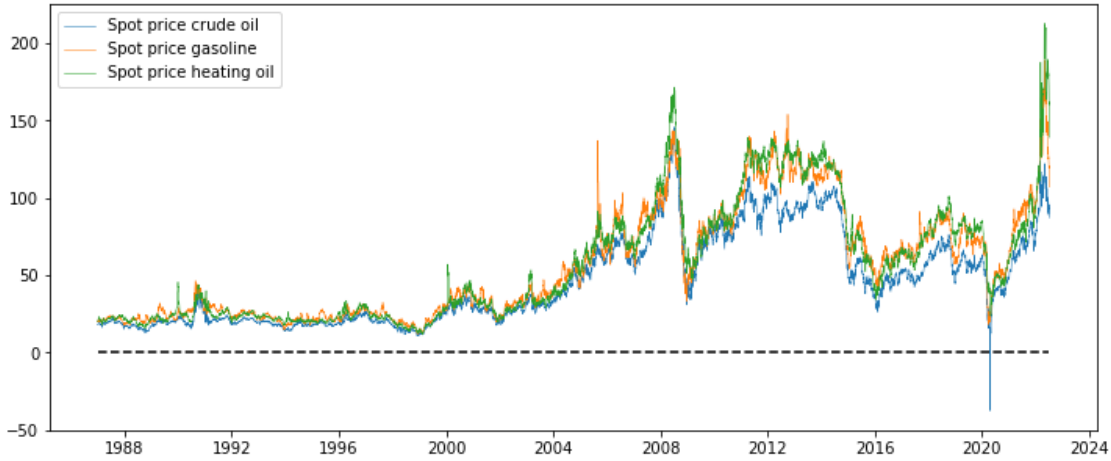

Figure 1: Daily spot prices of crude oil, gasoline and heating oil over the period 1987 – 2022.

We can link steep increases and declines of prices to macroeconomic events. The volatile price in 2008 can be linked to the financial crisis. The decrease around 2014 was caused by an oversupply of petroleum

(Mead & Stiger (2015)). The demand decline caused by COVID-19 produced the unique negative spike in April 2020. Lastly, the recent rise of petroleum prices is due to an increased energy demand after COVID-19 subsided combined with the Ukraine war and the subsequent obstructed supply of Russian oil and gas.

Table 1: Data characteristics regarding the weekly price changes of 6 variables during our sample period: spot price of crude oil (ΔS^C), spot price of gasoline (ΔS^G), spot price of heating oil (ΔS^H), futures price of crude oil (ΔF^C), futures price of gasoline (ΔF^G) and futures price of heating oil (ΔF^H). ADF is an abbreviation for the Augmented Dickey-Fuller statistic. A * in front of the statistic signals 1% significance.

	ΔS^C	ΔS^G	ΔS^H	ΔF^C	ΔF^G	ΔF^H
Mean	0.039	0.051	0.052	0.014	0.148	0.050
Median	0.070	0.042	0.073	0.120	0.126	0.084
Min.	-14.560	-44.318	-14.776	-14.410	-19.849	-14.784
Max.	14.130	57.456	18.732	14.080	28.938	18.018
Std. Dev.	2.460	3.622	2.883	2.395	3.008	2.736
Skewness	-0.325	0.801	-0.061	-0.539	0.019	-0.186
Kurtosis	5.174	47.752	5.141	5.474	8.058	4.787
Jarque-Bera	*2,076.0	*174,252.9	*2,018.8	*2,375.8	*4,957.0	*1,759.7
ADF	*-22.797	*-23.554	*-42.266	*-21.914	*-21.581	*-22.683

In Table 1, we observe that the spot and futures price changes of gasoline are the most volatile and the only variables which are positively skewed. Furthermore, the kurtosis of price changes in gasoline spot prices is relatively high, meaning fat tails in its distribution. Moreover, all variables reject the Jarque-Bera test at 1% significance, meaning that the price changes are not normally distributed. Lastly, all variables reject the Augmented Dickey Fuller test at 1%, which signals that the considered time series are stationary. Table 2 presents that for each spot price variable the level of association and correlation with the futures price changes of the same commodity is the highest, and vice versa. Moreover, all products are related positively.

Table 2: Correlation matrix (left) and the Kendall's τ matrix (right) of the weekly price changes.

	ΔS^C	ΔS^G	ΔS^H	ΔF^C	ΔF^G	ΔF^H
ΔS^C	1.000	-	-	-	-	-
ΔS^G	0.622	1.000	-	-	-	-
ΔS^H	0.789	0.638	1.000	-	-	-
ΔF^C	0.976	0.632	0.806	1.000	-	-
ΔF^G	0.753	0.892	0.739	0.765	1.000	-
ΔF^H	0.842	0.669	0.952	0.864	0.793	1.000

	ΔS^C	ΔS^G	ΔS^H	ΔF^C	ΔF^G	ΔF^H
ΔS^C	1.000	-	-	-	-	-
ΔS^G	0.503	1.000	-	-	-	-
ΔS^H	0.594	0.499	1.000	-	-	-
ΔF^C	0.887	0.518	0.609	1.000	-	-
ΔF^G	0.566	0.778	0.554	0.588	1.000	-
ΔF^H	0.638	0.525	0.846	0.665	0.600	1.000

4 Methodology

In this section, we set out the hedging frameworks and methods used for our research. Firstly, in Section 4.1, we explain the workings of an oil refinery and the problem setting. Moreover, we elaborate on the three hedging frameworks. Hereafter, in Section 4.2, we consider the theoretical aspect of the copula framework. This framework enables us to ultimately model the joint distribution of the six variables. Subsequently, we explain the estimation procedure of the bivariate and vine copula models in Section 4.3. In Section 4.4, we set out the sampling procedures of the models. Furthermore, we state how to convert draws from the copula density into draws from the joint distribution of the price change series in Section 4.5. Ultimately, we discuss the downside risk objectives of the hedging frameworks in Section 4.6 and the methods to gauge the fit of our models as well as the performance of our hedging strategies in Section 4.7.

4.1 The hedging framework

An oil refinery uses approximately three barrels of crude oil to produce two barrels of gasoline and one barrel of heating oil. The production process takes roughly one week, leaving the oil refinery exposed to the price changes of the volatile oil products market. The conceptual framework of Ji & Fan (2011) sets out the 2-stage hedging cycle, which covers three weeks, meaning 15 trading days, in total. In these three weeks, the buyer takes the following actions:

- Stage 1 - the planning stage covers the first 2 weeks
  - Day 1: Opens long position in crude oil futures, as the buyer plans ahead and prefers certainty about the price of its input product. The purchase price is wholly locked in, if the buyer acquires a futures contract of equal size as the crude oil input needed. Similarly, the buyer takes short positions in gasoline and heating oil futures contracts.
- Stage 2 - the operational stage concerns the last week
  - Day 1: Buys crude oil on the spot market to start the cracking process, subsequently closes the crude oil futures position. If, in the meantime, the crude oil price has decreased, buying on the spot market is cheaper, while the value of the futures contracts is likely to have declined. If prices have increased, an oil refinery makes up for the difference through the income generated by the futures contract.
  - Day 5 (last day of operational stage): Sells gasoline and heating oil against spot prices, subsequently closes futures positions of the corresponding products.

We can overlook the time from decision to operation and the production time for oil refinery process, because the actual sale of these contracts can be executed simultaneously. This concept will not alter the downside risk measures of the profit margin, but it will make the entire decision-making process clearer. This

sequence of actions results in the following profit margin for the refiner for a specific time window (rolling window sample) t

$$\pi(\lambda) = -S_2^C + \frac{2}{3}S_2^G + \frac{1}{3}S_2^H + \lambda^C(-F_1^C + F_2^C) - \frac{2}{3}\lambda^G(-F_1^G + F_2^G) - \frac{1}{3}\lambda^H(-F_1^H + F_2^H), \tag{1}$$

where $\pi(\lambda)$ is the profit margin of the oil refinery for the hedge ratios $\lambda = (\lambda^C, \lambda^G, \lambda^H)$. The indices represent the week $\{1, 2\}$ in which the product is bought or sold against the spot price or futures price. Because the changes in future and spot prices belong to our dataset, we can rewrite Equation (1) as

$$\pi(\lambda) = -S_2^C + \frac{2}{3}S_2^G + \frac{1}{3}S_2^H + \lambda^C\Delta F^C - \frac{2}{3}\lambda^G\Delta F^G - \frac{1}{3}\lambda^H\Delta F^H,$$

where $\Delta F^C = (-F_1^C + F_2^C)$, $\Delta F^G = (-F_1^G + F_2^G)$ and $\Delta F^H = (-F_1^H + F_2^H)$. We subtract the sum of spot prices in week 1 (S_1^{CGH}) from both sides, as

$$\pi(\lambda) - S_1^{CGH} = -S_2^C + \frac{2}{3}S_2^G + \frac{1}{3}S_2^H + \lambda^C\Delta F^C - \frac{2}{3}\lambda^G\Delta F^G - \frac{1}{3}\lambda^H\Delta F^H - S_1^{CGH},$$

where $S_1^{CGH} = -S_1^C + \frac{2}{3}S_1^G + \frac{1}{3}S_1^H$. This transformation allows us to establish the formula for the hedged portfolio $Y(\lambda)$ as sum of variable price changes, as follows

$$Y(\lambda) = -\Delta S^C + \frac{2}{3}\Delta S^G + \frac{1}{3}\Delta S^H + \lambda^C\Delta F^C - \frac{2}{3}\lambda^G\Delta F^G - \frac{1}{3}\lambda^H\Delta F^H, \tag{2}$$

where $Y(\lambda) = \pi(\lambda) - S_1^{CGH}$, $\Delta S^C = (-S_1^C + S_2^C)$, $\Delta S^G = (-S_1^G + S_2^G)$ and $\Delta S^H = (-S_1^H + S_2^H)$. Then, we define the objective function for time window t to minimise downside risk with hedge ratios as

$$\lambda^* = \underset{\lambda}{\operatorname{argmin}}\, Risk(Y(\lambda)), \tag{3}$$

where $Risk$ is the downside risk measure of choice, for example Value-at-Risk (VaR) or Expected Shortfall (ES). Ultimately, we aim to find these optimal hedge ratios by optimisation over the joint distribution of $Y(\lambda)$ for each time frame t. In this paper, we examine the performance of three hedging strategies searching the optimal hedging ratios in three unique hedging frameworks: the single commodity-, the fixed proportion- and the flexible multi-commodity hedging framework. These three models differ in the flexibility towards the hedging exercise and the information considered for determining the hedging ratios. We determine our hedge ratios based on a rolling window making steps of one week from t is 16 October 1991 to 7 August 2019. This construction provides 1452 models and hedge ratios per hedging framework to test, as shown in Figure 2.

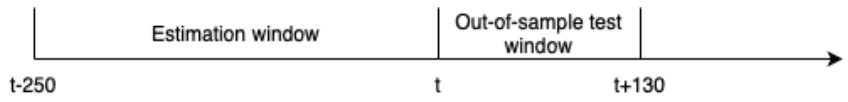

Figure 2: Estimation and test windows of our rolling window model for the hedge strategies. Herein, the estimation window used to estimate our models and hedge ratios covers 250 weeks and the test window covers 130 weeks.

4.1.1 The single commodity hedging framework

The single commodity hedging framework optimises hedge ratios based on the changes in spot and future prices of each commodity individually. The rationale behind this model stems from the fact that the spot and futures prices of the same petroleum product exhibit the strongest correlation. This hedging model is flexible per commodity, as λ^C, λ^G and λ^H are free parameters in Equation (2). When determining the hedge ratio for one of the petroleum products, the historical price changes of the other products will not be taken into account. We can define the objective functions of this model, as follows

$$\lambda^{C*} = \underset{\lambda^C}{\operatorname{argmin}}\, Risk(Y_C(\lambda^C)),$$
$$\lambda^{G*} = \underset{\lambda^G}{\operatorname{argmin}}\, Risk(Y_G(\lambda^G)),$$
$$\lambda^{H*} = \underset{\lambda^H}{\operatorname{argmin}}\, Risk(Y_H(\lambda^H)),$$

given hedged portfolio distributions $Y(\lambda^i)$ for each commodity i

$$Y_C(\lambda^C) = -\Delta S^C + \lambda^C \Delta F^C,$$
$$Y_G(\lambda^G) = \Delta S^G - \lambda^G \Delta F^G,$$
$$Y_H(\lambda^H) = \Delta S^H - \lambda^H \Delta F^H,$$

where the sign of crude oil contracts is opposed to gasoline and heating oil. This is due to the fact that we protect the crude oil position against rising prices, while we hedge the products that the oil refinery sells against declining prices. The total hedged portfolio in the single commodity hedging framework follows the formula,

$$Y(\lambda^*) = Y_C(\lambda^{C*}) + \frac{2}{3} Y_G(\lambda^{G*}) + \frac{1}{3} Y_H(\lambda^{H*}), \tag{4}$$

where $\lambda^* = (\lambda^{C*}, \lambda^{G*}, \lambda^{H*})$.

4.1.2 The fixed proportion hedging framework

Given that the proportions of the production process of an oil refinery are known in advance, we consider to optimise the hedge ratio for this fixed $3:2:1$ proportion. This fixed proportion assumes that hedge ratios $\lambda^C, \lambda^G, \lambda^H$ are multiples of each other. For implementation, we can thus remove two degrees of freedom by using λ^f, which provides the fixed proportion hedge ratio

$$\lambda^{f*} = \underset{\lambda^f}{\operatorname{argmin}}\, Risk(Y_f(\lambda^f)),$$

given the hedge portfolio distribution

$$Y_f(\lambda^f) = -\Delta S^C + \frac{2}{3}\Delta S^G + \frac{1}{3}\Delta S^H + \lambda^f(\Delta F^C - \frac{2}{3}\Delta F^G - \frac{1}{3}\Delta F^H), \tag{5}$$

where one degree of freedom is left in λ^f.

4.1.3 The flexible multi-commodity hedging framework

The flexible multi-commodity hedging framework optimises hedge ratios λ as described in Equation (2) and (3). Herein, three degrees of freedom are present in λ^C, λ^G and λ^H, while historical information of all variables is jointly considered to determine the joint distributions over which the hedge ratios are optimised.

4.2 Theoretical copula framework

From the previous section, we would like to investigate the multivariate distribution $Y(\lambda)$ for rolling window sample t. Herefore, we use vine copula methodology, which can model complex dependencies amongst higher dimensions. To understand a vine copula, we first describe the copula framework and the bivariate copula, in general.

Copulae are multivariate cumulative distribution functions (CDFs) that define the structure of dependencies between random variables $X = (X_1, X_2, ..., X_d)$. The d-dimensional copula can be decomposed to d univariate marginal distributions and a multi-dimensional copula, according to Sklar (1959). If we let $F(x)$ be a joint distribution with marginal distributions $F_1, ..., F_d$, there exists a copula C that links these marginal distributions as

$$F(x_1, ..., x_d) = C(F_1(x_1), ...F_d(x_d)), \tag{6}$$

$$F(x_1, ..., x_d) = Pr(X_1 \leq x_1, ..., X_d \leq x_d),$$

where copula C can be defined as C: $[0,1]^d \rightarrow [0,1]$, which is a CDF with d univariate marginal distributions. The copula is unique when the univariate marginal distribution F_i is continuous for each random variable, as

$$C(u_1, ..., u_n) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_d)),$$

where F_i^{-1} is the inverse of F_i and observations u_i are distributed $U(0,1)$.

For our multivariate problem, it is possible to use a multivariate copula from one copula type, similar to the multivariate copula models in Embrechts et al. (2001). These multivariate copulae strictly impose one copula type, such as Gaussian, Student-t or Clayton (see Appendix A), on all variables. However, these models are inflexible, as the dependencies amongst multiple variables are then to be modeled through one or two copula parameters, depending on the copula type. To achieve a more tailored fit, Aas et al. (2009) introduced a pair-copula construction method based on bivariate copula as building blocks, which can equip copula types and parameters to the dependency structures per variable pair.

4.2.1 Bivariate copula structure

A bivariate copula is a copula with $d = 2$ dimensions, such that it models the dependence structure between two variables. We define the bivariate cumulative distribution functions for the used 36 copula types in Appendix A. Different copula types and families can capture unique dependency characteristics. For example, the Gaussian copula is symmetric and not able to capture tail dependence, while the symmetrical Student-t

copula can capture tail dependence in both tails with a second copula parameter. Furthermore, a certain group of copula types, the Archimedean copula family comprises the tail-asymmetric Clayton and Gumbel copula types. Each copula type has at least one copula parameter, which shapes the bivariate distribution. The difference in symmetry can be observed from the scatter plots in Figure 3.

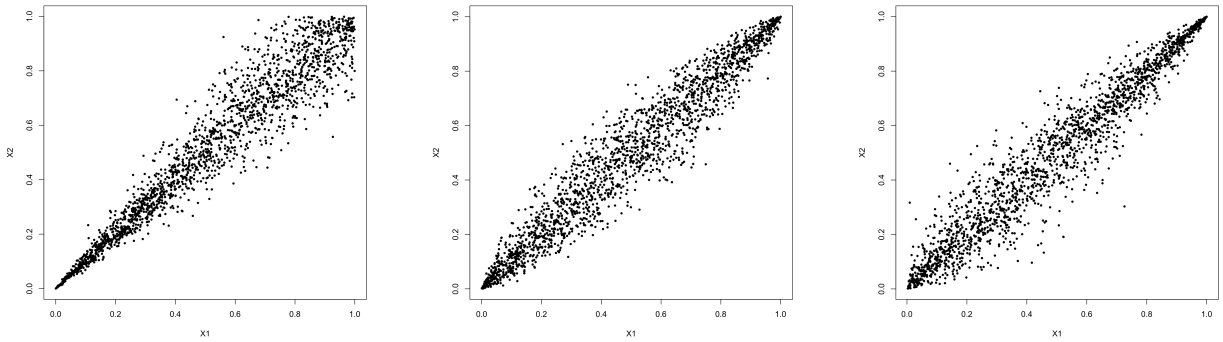

Figure 3: Three scatter plots with different measures of tail dependence where copula defines the dependence structure on 3,000 random observations on bivariate copulae. All scatter plots have a linear correlation of roughly 0.9. From left to right a.) Clayton copula that exhibits lower tail dependence b.) Gaussian copula that exhibits no tail dependence c.) Gumbel copula that exhibits upper tail dependence. The scatter plots of all the copula types applied in this research are provided in Appendix A.

From Sklar's theorem in Equation (6), it is possible to derive the joint density function $f_{12}(x_1, x_2)$ and conditional density $f_{1|2}(x_1|x_2)$, as

$$f_{1,2}(x_1, x_2) = c_{1,2}(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2),$$

$$f_{1|2}(x_1|x_2) = c_{1,2}(F_1(x_1), F_2(x_2))f_2(x_2),$$

where $c_{1,2}(F_1(x_1), F_2(x_2))$ represents the pair-copula density for marginal CDFs $F_i(x_i)$. Furthermore, $f_1(x_1)$ and $f_2(x_2)$ provide the marginal density functions. Through a constellation of these functions with more dimensions, we ultimately construct the joint density of a vine copula.

4.2.2 Vine copula structure

With the previously described bivariate copulae, we define a multivariate model to simulate $Y(\lambda)$ with six random variables for our oil refinery instance. From here on, we refer to the observed price changes in spot and future prices as random variables $X = (X_1, X_2, ..., X_6)$, which means in practice $X_1 = \Delta S^C$, $X_2 = \Delta S^G$, $X_3 = \Delta S^H$, $X_4 = \Delta F^C$, $X_5 = \Delta F^C$ and $X_6 = \Delta F^H$. Vine copulae allow us to construct multivariate distributions from bivariate copulae, by defining the marginal distributions of the random variables as well as the unconditional- and conditional dependence of the pairs of random variables of concern. Herein, we have the flexibility to assign the best fitting copula types to specify the dependence structure between random variable pairs.

To describe the 6-dimensional distribution of $(X_1, ..., X_6)$, we first define the 3-dimensional distribution (X_1, X_4, X_6) and its joint density function $f(x_1, x_4, x_6)$, as an example. Figure 4 shows that, when considering X_1, X_4 and X_6, we can construct multiple vine copula models. The difference between the two vines is the order of the variables in tree T_1, which consequently leads to a different T_2. A tree is a connected acyclic graph with node set N and edge set E, denoted as $T = (N, E)$. Herein, each edge e represents a pair-copula density $c_{i,j}(F_i(x_i), F_j(x_j))$ based on two (conditional) marginal CDFs: F_i, F_j from variable i and j, while the edge label provides its subscript. Each node n in tree T_1 provides a marginal CDF corresponding to its subscript variable. For example, node n_1 in Figure 4 expresses the marginal CDF of variable X_1: F_1. In higher trees $(T_r, r \geq 2)$, each node presents a conditional marginal CDF. In this case, the marginal CDF does not exactly correspond to the subscript of the node, but to a combination of the variables included in the subscript. Herein, the variable(s) in the common node of the previous tree are included in the conditioning set, while the remaining variable is in the conditioned set. For example, node $n_{4,6}$ in Figure 4a provides the conditional marginal CDF $F_{6|4}$, where variable 6 is in the conditioned set and variable 4, the common node in T_1, is in the conditioning set. A node in higher trees does not uniquely define the marginal CDF. For instance, node $n_{4,6}$ in Figure 4b provides the conditional marginal CDF $F_{4|6}$, as the bivariate copula (edge) is conditional on variable 6. In Figure 4b, variable 6 is the common node in T_1.

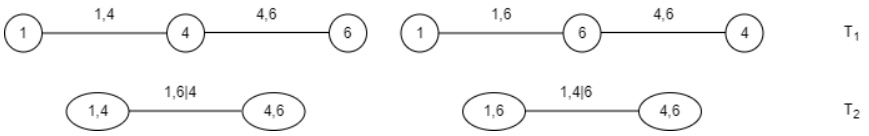

Figure 4: Graphical examples of an R-vine for a 3-dimensional vine for variables (X_1, X_4, X_6). Herein, the graph on the left a.) also resembles a part of the 6-dimensional R-vine in Figure 5, while the graph on the right b.) is an R-vine with a different structure, but equivalent joint density function $f(x_1, x_4, x_6)$ as a.).

We encounter these different conditional marginal CDFs in Equations (7) and (8), which provide the joint density for these three variables. In total, the joint density for three variables can be defined in three different formulas due to copulas being symmetrical. We give two examples: the graph in Figure 4a leads to

$$\begin{aligned} f(x_1, x_4, x_6) =& c_{1,4}(F_1(x_1), F_4(x_4)) \cdot c_{4,6}(F_4(x_4), F_6(x_6)) \\ & c_{1,6|4}(F_{1|4}(x_1|x_4), F_{6|4}(x_6|x_4)) \\ & f_1(x_1) \cdot f_4(x_4) \cdot f_6(x_6), \end{aligned} \tag{7}$$

where we make a simplifying assumption that all bivariate copulae of conditional distributions in vine copulae are not dependent on the variables that they are conditioned on. It is found in Stoeber et al. (2013) that this simplifying assumption leads to satisfactory vine copula approximations, such that we use $c_{i,j|q}(\cdot, \cdot)$ instead of $c_{i,j|q}(\cdot, \cdot|x_q)$. Using the factorisation of the vine copula depicted in Figure 4b, the same joint density can

also be defined as

$$\begin{aligned} f(x_1, x_4, x_6) =& c_{1,6}(F_1(x_1), F_6(x_6)) \cdot c_{4,6}(F_4(x_4), F_6(x_6)) \\ & c_{1,4|6}(F_{1|6}(x_1|x_6), F_{4|6}(x_4|x_6)) \\ & f_1(x_1) \cdot f_4(x_4) \cdot f_6(x_6). \end{aligned} \tag{8}$$

We have shown that the smallest vine copula ($d = 3$) has multiple factorisations. In both factorisations, all possible variable pairs are included in the conditioned sets of the edges, which is a characteristic of a vine copula.

We expand the 3-dimensional vine to a 6-dimensional vine copula through a graphical example provided in Figure 5. Herein, the nodes and edges in Trapezoid 1 resemble the vine copula model in Figure 4a. The

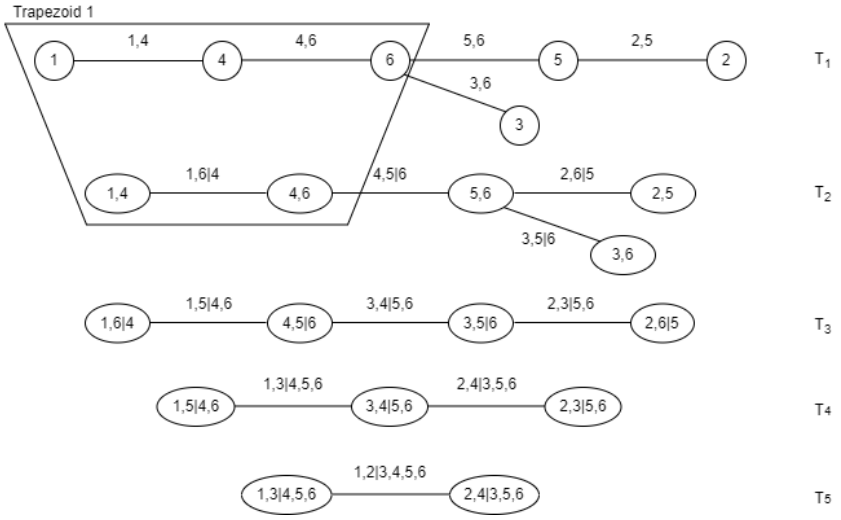

Figure 5: Graphical example of an R-vine for 6-dimensional vine, where Trapezoid 1 secludes a 3-dimensional vine.

graphical vine models in Figure 4 and 5 are called regular vines (R-vines). For example, the joint density function of the corresponding vine specification in Figure 5, $V = \{T_1, ..., T_5\}$, is mathematically defined as follows

$$\begin{aligned} f(x_1, x_2, x_3, x_4, x_5, x_6) =& c_{1,4}(F_1(x_1), F_4(x_4)) \cdot c_{4,6}(F_4(x_4), F_6(x_6)) \cdot c_{5,6}(F_5(x_6), F_5(x_6)) \cdot c_{3,6}(F_3(x_3), F_6(x_6)) \\ & c_{2,5}(F_2(x_2), F_5(x_5)) \cdot c_{1,6|4}(F_{1|4}(x_1|x_4), F_{6|4}(x_6|x_4)) \cdot c_{4,5|6}(F_{4|6}(x_4|x_6), F_{5|6}(x_5|x_6)) \\ & c_{3,5|6}(F_{3|6}(x_3|x_6), F_{5|6}(x_5|x_6)) \cdot c_{2,6|5}(F_{2|5}(x_2|x_5), F_{6|5}(x_6|x_5)) \\ & c_{1,5|4,6}(F_{1|4,6}(x_1|x_4, x_6), F_{5|4,6}(x_5|x_4, x_6)) \cdot c_{3,4|5,6}(F_{3|5,6}(x_3|x_5, x_6), F_{4|5,6}(x_4|x_5, x_6)) \\ & c_{2,3|5,6}(F_{2|5,6}(x_2|x_5, x_6), F_{3|5,6}(x_3|x_5, x_6)) \\ & c_{1,3|4,5,6}(F_{1|4,5,6}(x_1|x_4, x_5, x_6), F_{3|4,5,6}(x_3|x_4, x_5, x_6)) \\ & c_{2,4|3,5,6}(F_{2|3,5,6}(x_2|x_3, x_5, x_6), F_{4|3,5,6}(x_4|x_2, x_5, x_6)) \\ & c_{1,2|3,4,5,6}(F_{1|3,4,5,6}(x_1|x_3, x_4, x_5, x_6), F_{2|3,4,5,6}(x_2|x_3, x_4, x_5, x_6)) \\ & f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \cdot f_5(x_5) \cdot f_6(x_6), \end{aligned}$$

which is one of 23,040 possible R-vine tree structures for a 6-dimensional joint distribution. Following Morales-Napoles (2010), the number of different decompositions of R-vine tree sequences grows as $d!$ x $2^{(d-2)(d-3)/2-1}$ with dimension d.

In general, R-vines consist of linked trees, where the edges in one tree become the nodes of the next tree by means of conditioning. If the following conditions are met, the set of graphs $V = (T_1, ..., T_{d-1})$, is an R-vine tree sequence on d variables:

- Tree T_1 has an edge set E_1 and node set $N_1 = \{1, ..., d\}$
- For $r \geq 2$, T_r is a tree with node set N_r of which the number of nodes is equal to number of edges of the edge set E_{r-1}
- For $\{a, b\} \in E_r$ for $r = 2, ..., d-1$, it must hold that $|a \cap b| = 1$. According to this proximity condition, an edge between two nodes in tree T_r can only exist if the associated edges in T_{r-1} have a common node. We can deduce the association between these nodes by their common subscript

To build an R-vine V, it is necessary to specify $d-1$ unconditional bivariate copulae (edges) between variables in tree T_1. For T_r with $r \geq 2$ in the R-vine, it is necessary to specify $d - r$ bivariate copulae conditional on the variables in the conditioning and conditioned set of the common node.

Within the class of R-vines, there exist two subclasses: C-vines and D-vines. These vine copula constructions are known for their simple and interpretable construction. A C-vine tree sequence resembles a star, because in each tree there is one root node, which is connected with all remaining nodes. This subclass is often used when there is one pivot variable. The trees in a D-vine tree structure are paths, where each node has a maximum degree of 2. This structure is used when modelling variables with a clear sequence or order. Both subclasses comprise $\frac{d!}{2}$ unique vine tree sequences each. For our instance $d = 6$, this results in 360 different vine tree sequence per subclass, which means that each subclass represents approximately 1.5% ($\frac{360}{23,040}$) of the total possible R-vine tree sequences. We search all R-vine structures to find the best fitting vine, which means a substantial increase in the number of C- and D- vine structures. Moreover, for our instance, it is likely that our data contains three 'pilot' pairs, namely the spot-futures pair of each commodity. This could indicate that although C- and D- vines are more interpretable, they may not be optimal.

4.3 Estimation procedure

During the estimation procedure, we fit the model based on joint density, which concerns defining the vine tree sequence, the copula types and the copula parameters parameters for each edge in the vine tree. For the estimation procedure we differentiate across the three frameworks: the single commodity hedging framework with $Y(\lambda)$ having two dimensions in Section 4.3.1, the fixed proportion- and the flexible multi-commodity hedging framework where $Y(\lambda)$ has six dimensions in Section 4.3.2. Furthermore, in Section 4.3.3, we explain

the estimation procedure of our proposed mixture R-vine model used for the fixed proportion and the flexible multi-commodity hedging framework.

Firstly, we convert our observed price change series $(X_1, ..., X_6)$ into marginally uniformly distributed pseudo-observations $(Z_1, ..., Z_6)$ for copula inputs of all our applied copula models. In general, there are two options with regard to modelling the marginal distribution: parametric and non-parametric estimation methods. Because misspecification of the parametric model is non-robust and because the non-parametric approach performs better overall than the parametric approach (G. Kim et al. (2007)), we opt to apply the non-parametric approach of Genest et al. (1995).

In all three hedging frameworks, we estimate the pseudo-observations Z_i of commodity price changes in i by their scaled empirical distribution functions $\hat{F}_i(x)$, defined as follows

$$\hat{F}_i(x) = \frac{1}{T+1} \sum_{t=1}^{T} \mathbf{1}_{X_{i,t} \leq x}, \tag{9}$$

where $\{X_{i,t}\}_{t=1}^{T}$ represents the observations of X_i in a given period. In our instance, this period includes $T = 250$ price changes in commodity i. More specifically, we transform one price change observation $X_{i,t}$ to $Z_{i,t}$ as follows

$$Z_{i,t} = \hat{F}_i(X_{i,t}). \tag{10}$$

Subsequently, we can define $Z_i = \hat{F}_i(X_i)$ as the pseudo data for variable i, representing the normalised rank of the whole rolling window sample for the respective variable.

4.3.1 The bivariate copula models for the single-commodity framework

We discuss the estimation procedure for the bivariate copula of the 2-dimensional single commodity hedging model from Section 4.1.1. Herein, it is specified which variables are connected, as illustrated in Figure 6. Therefore, we estimate which copula type-parameter(s) combination best describes the bivariate distribution of the spot and future price changes of crude oil, gasoline and heating oil, respectively.

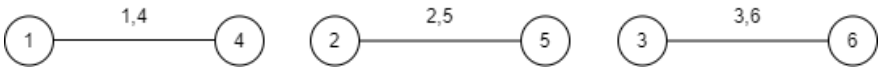

Figure 6: Graphical representation of the single commodity hedging framework.

We select the copula type-parameter(s) combination by applying maximum pseudo likelihood estimation for all copula types mentioned in our list in Appendix A. Herein, we pick the copula type with the lowest Akaike Information Criterion (AIC) based on the pseudo data derived in Equation (10), given by

$$\hat{\Theta}_{f;i,j} = \underset{\Theta_{f;i,j}}{\operatorname{argmin}}\, AIC(\Theta_{f;i,j}) \quad AIC(\Theta_{f;i,j}) = -2 \sum_{t=1}^{T} \log\{c_{f;i,j}(Z_{i,t}, Z_{j,t}; \Theta_{i,j})\} + 2k, \tag{11}$$

where k is the number of parameters in copula parameter(s) $\Theta_{i,j}$, which belongs to variable pair $\{i, j\}$. These parameters set the copula density function of copula type f: $c_f(\cdot, \cdot; \Theta_{i,j})$. Furthermore, $Z_{i,t}$ provides

the pseudo-observations of Z_i, where t iterates from the first to the last observation T within Z_i. $\hat{\Theta}_{f;i,j}$ represents the copula parameters of copula type f with the highest AIC for the pseudo-observations. In the single commodity hedging framework the variables $\{i, j\}$ are the pairs $\{\{1, 4\}, \{2, 5\}, \{3, 6\}\}$. Manner (2007) and Brechmann (2010) reported reliable performance of this information criterion for the selection of bivariate copulae. The papers note that identification performance is related positively with dependence and sample size. When weak dependence is present, the AIC might be unsatisfactory.

4.3.2 Estimation procedure vine copula

In this section, we expand the estimation procedure to six dimensions of the vine copula model for the fixed proportion and flexible multi-commodity hedging framework. Besides finding the best fitting copula type and parameters per variable pair, we determine the tree structure for all trees in the vine structure. We visually deduce this increase in complexity, when we compare the vine copula in Figure 5 to the bivariate copulae needed in the single hedging framework in Figure 6. In Section 4.3.3, we expand the selection possibilities by allowing to assign more than one copula type per variable pair, provided certain conditions are met.

A method to find the optimal order of the vine tree sequence is to estimate the log-likelihood of all possible tree sequences for all copula types for every edge and optimize the corresponding parameters. However, we know in advance that this is computationally too expensive. As such, we follow the algorithm of Dissmann et al. (2013), which has been found as a good indicator of the rank of full log-likelihood for vine copula models. This algorithm relies on repeating three steps for each tree: defining the tree structure, selecting and estimating the best copula types and parameters and, lastly, generating conditional pseudo-observations. In each step, we first describe how to initialise the algorithm for the first tree T_1. Subsequently, we explain how this step is performed in higher order trees.

Step 1: Defining a tree structure

Our starting point is to find the structure of the first layer, such as tree T_1 in Figure 5. For every R-vine, the estimation of the first tree, T_1, deviates slightly from the higher tree estimations T_i for $i = \{2, ..., 5\}$. This is due to the fact that the pseudo-observations of T_1 are the empirical marginal CDFs of the variables $(Z_1, ..., Z_6)$ from Equation (10) and no variables have been paired yet. The structure of T_1 is selected from a complete graph (see Figure 7a), whereas the structure of higher order trees is partially pre-defined by the conditions of a vine copula in Section 4.2.2 (see Figure 7b and 7c).

Copula type specifications in the first tree generally have the biggest impact on model fit (Dissmann et al. (2013)). Therefore, it is important to define the dependence relation between random variable pairs with large dependencies. The automated sequential method constructs a vine copula specification based on Kendall's τ. Kendall's τ measures dependence between bivariate samples independently from an assumed distribution. To measure the dependence in our sample, we calculate the empirical Kendall's $\hat{\tau}$ for the pseudo-observations of each pair $\{i, j\}$, which is based on concordant and discordant observation pairs C

and D, as

$$\hat{\tau}_{i,j} = \frac{C - D}{C + D},$$
$$C = \sum_{t_1=1}^{T-1} \sum_{t_2=t_1+1}^{T} \mathbb{1}_{sign(z_{i,t_1} - z_{i,t_2}) = sign(z_{j,t_1} - z_{j,t_2})},$$
$$D = \sum_{t_1=1}^{T-1} \sum_{t_2=t_1+1}^{T} \mathbb{1}_{sign(z_{i,t_1} - z_{i,t_2}) \neq sign(z_{j,t_1} - z_{j,t_2})},$$

where $(z_{1,t}, ..., z_{6,t})$ represent the pseudo-observations from the data set for a rolling window sample with t_1 being the first observation of the estimation window and T the last. Concordance is the level of association between two variables. Concordance is higher when large values of z_i are observed alongside large values of z_j. We select the edges by applying the minimum spanning tree (MST) algorithm (Prim (1957)), which is explained in Appendix B. The resulting path has the largest sum of absolute empirical Kendall's $\hat{\tau}_{ij}$ for a tree, while satisfying the R-vine conditions as given in Section 4.2.2. This can be exemplified for tree T_1 as

$$(N_1, E_1) = \underset{E_1}{\operatorname{argmax}} \sum_{e=\{i,j\} \text{in spanning tree}} |\hat{\tau}_{i,j}|,$$

where N_1 represents the node set and E_1 represents the edge set of T_1.

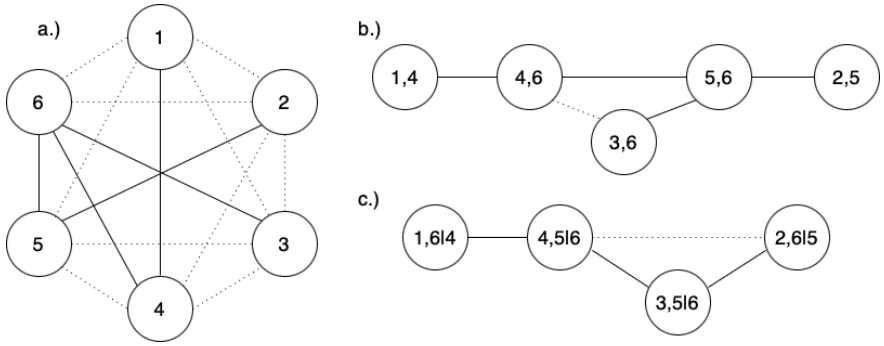

Figure 7: One graphical example of potential model selection of a.) T_1, b.) T_2 and c.) T_3 of the 6-dimensional R-vine given in Figure 5. Herein, the solid lines represent the selected edges, whereas the dashed lines show the other possible connections.

This procedure is also applicable for higher order trees, where Kendall's $\hat{\tau}$ is estimated over conditional pseudo-observations $z_{i|D}$ instead of z_i, where D represents the conditioning set.

Step 2: Selection of best copula type and parameter(s) for each edge in the tree

After defining the tree structure of T_1 similar to the example in Figure 7a, we determine the best fitting copula type and parameter(s) for each of the edges $\{i, j\}$ in (N_1, E_1) in tree T_1. An edge of tree T_1 represents the joint distribution between variable i with empirical marginal distribution Z_i and j with Z_j. For each edge, we apply the maximum pseudo likelihood estimation of Equation (11) for all copula types as described in Section 4.3.1. The difference with the procedure of Section 4.3.1 is that for T_1, we estimate the parameter(s)

for the five variable pairs in E_1, which include $\{\{1,4\},\{2,5\},\{3,6\},\{4,6\},\{5,6\}\}$ in the example of Figure 7a.

The copula types and parameters in higher trees are determined in a similar way. The difference is that the pseudo-observations are conditional on the variables in the common node in previous tree T_{r-1}, see Step 3. We group these common node variables in conditioning set D. Therefore, we estimate the optimal copula parameters for each copula type f in higher trees, as

$$\begin{aligned} \hat{\Theta}_{f;i,j|D} &= \underset{\Theta_{f;i,j|D}}{\operatorname{argmin}} AIC(\Theta_{f;i,j|D}) \\ AIC(\Theta_{f;i,j|D}) &= -2\sum_{t=1}^{T} \log\{c_{f;i,j|D}(Z_{i|D,t}, Z_{j|D,t}; \Theta_{i,j|D})\} + 2k, \end{aligned} \tag{12}$$

where the subscript $(f;i,j|D)$ refers to copula type f for edge $e_{i,j|D}$. $Z_{i|D,t}$ represents the conditional pseudo-observations with t being the observation index within $Z_{i|D}$. The remaining parameters are defined in Equation (11). For this edge, the copula type f and corresponding parameters $\hat{\Theta}_{f;i,j|D}$ with the lowest AIC are estimated and selected.

Step 3: Transforming pseudo data with copula specifications

The specifications of the tree structure and bivariate copulae of T_1 allow us to estimate the conditional distribution $Z_{i|j}$. These conditional marginal distributions are necessary to detect the structure and copula types directly in T_2 and indirectly in the higher order trees T_r for $r = \{3,4,5\}$. The marginal conditional distributions can be expressed as partial differentiated copulae. Based on the pairs in T_1, we define the CDF of pseudo-observations pair (Z_i, Z_j) as

$$C_{f;i,j}(z_i, z_j; \hat{\Theta}_{f;i,j}), \tag{13}$$

where $(f;i,j)$ provides the copula type f for edge $e_{i,j}$ selected in Step 2 and $\hat{\Theta}_{f;i,j}$ presents the copula parameters estimated in Equation (11). Hereafter, we define the CDF of Z_i conditional on $(Z_j = z_j)$ with the $h_{f;i,j}(\cdot,\cdot)$-function as

$$h_{f;i,j}(z_i, z_j; \hat{\Theta}_{f;i,j}) = \frac{\partial\, C_{f;i,j}(z_i, z_j; \hat{\Theta}_{f;i,j})}{\partial\, z_j}, \tag{14}$$

where the first term in the h-function is always the conditioned variable and the second term the conditioning variable. The h-function differs per edge. Finally, we apply the $h_{f;i,j}(\cdot,\cdot)$-function to (Z_i, Z_j) such that

$$Z_{i|j} = h_{f;i,j}(Z_i, Z_j; \hat{\Theta}_{f;i,j}). \tag{15}$$

$Z_{j|i}$ can be derived in a similar way with the same copula $C_{f;i,j}$ due to copula symmetry.

As mentioned before, all edges from T_1 are nodes in T_2 and these nodes can only be connected if the corresponding edges in T_1 share a node. For T_2, this leads to less possible pairs than in the complete graph of T_1, see Figure 7a and 7b. Only when there are multiple possible tree structures in higher order trees, we apply the MST algorithm from Step 1. The algorithm selects the path with the largest cumulative absolute

empirical Kendall's $\hat{\tau}_{ij}$ over the conditional pseudo-observations. If the tree is a path, where each node has a maximum degree of 2, then the tree structure is uniquely defined for higher trees.

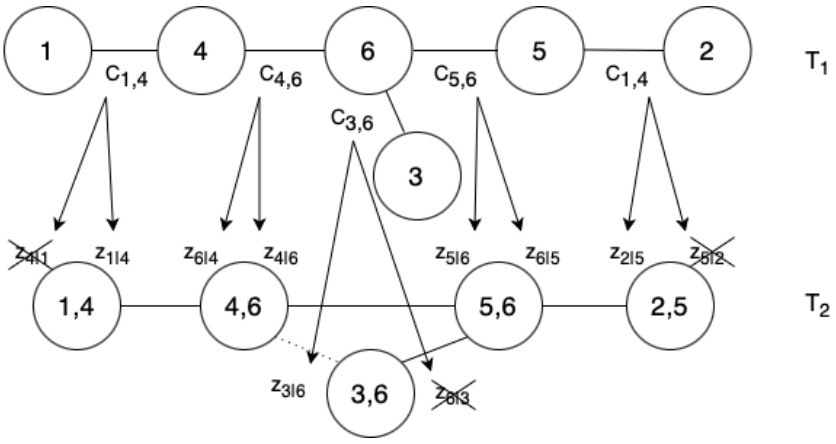

Figure 8: In depth example of model selection step from T_1 to T_2 in Figures 7a and 7b. Herein, the arrows represent the marginal conditional distributions that can be derived from each pair via Equations (13)-(15).

Figure 8 illustrates how the procedure decides which node pairs should be connected in T_2. The MST algorithm computes the Kendall's $\hat{\tau}$ of possible edges between nodes $n_{3,6}$ and $n_{4,6}$ over pseudo-observations $(Z_{3|6},\ Z_{4|6})$, between $n_{3,6}$ and $n_{5,6}$ over $(Z_{3|6},\ Z_{5|6})$ and between $n_{4,6}$ and $n_{5,6}$ over $(Z_{4|6},\ Z_{5|6})$. In our example, the tree with the largest cumulative Kendall's $\hat{\tau}$ excludes the edge between nodes $n_{3,6}$ and $n_{4,6}$. Furthermore, the crossed-out marginal conditional distributions are the pseudo-observations which could also have been derived from the copula specifications of T_1. However, these pseudo-observations are conditional on a node which is not shared in T_1 and therefore not used to determine higher order trees.

In higher trees $T_r, r \geq 3$, Kendall's $\hat{\tau}$ is estimated over the conditional marginal distributions $Z_{i|j,D}$, derived from the tree T_{r-1}. This is achieved in a similar way as in described by Equations (13)-(15). In this case $C_{f;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f;i,j|D})$ is the CDF of $(Z_{i|D}, Z_{j|D})$ as derived in Equation (12) and subsequently applying the corresponding h-function to $(Z_{i|D}, Z_{j|D})$ as follows

$$\begin{aligned} h_{f;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f;i,j|D}) &= \frac{\partial\, C_{f;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f;i,j|D})}{\partial\, z_{j|D}}, \\ Z_{i|j,D} &= h_{f;i,j|D}(Z_{i|D}, Z_{j|D}; \hat{\Theta}_{f;i,j|D}), \end{aligned} \tag{16}$$

where conditioning set D includes $r-2$ variables.

The path of T_3 defined over the the marginal conditional distributions generated by the edges of T_2 is shown in Figure 7c. The resulting path has a D-vine structure, as every node has at most two edges. Hence, this vine tree structure is uniquely defined in higher trees.

To specify the whole vine tree structure, we iterate over Step 1,2 and 3 until the tree consists of one edge for which we then define the copula type and estimate parameters. For our oil refinery instance, we end up with the specification of the full vine $V = \{T_1, ..., T_5\}$ with 15 defined edges. We estimate full log-likelihood and AIC of V by summing the log-likelihood and AIC of all edges in V as determined in Step 2.

4.3.3 Bivariate mixture in mixture R-vine model

The underlying dependency pattern between two variables in a vine copula is unknown, therefore we approximate the bivariate relation by assigning the best fitting copula type based on AIC in Sections 4.3.1 and 4.3.2. In this section, we explain the estimation procedure of our proposed mixture R-vine copula model. This mixture R-vine copula model uses multiple bivariate copulae to approximate the dependency structure between two variables.

In Manner (2007) and Brechmann (2010), it is found that AIC is the best indicator for copula type selection. In some cases, it is found that the blanket test based on empirical copula is superior (Brechmann (2010)). However, this selection test is computationally too expensive to implement. Therefore, AIC is a reliable selection criterion, but not strictly optimal. Although distinction between copula types increases with increasing dependence, we encounter copula types with similar fit to the pseudo-observations with strong dependence. In some cases, we come across copula types with higher log-likelihood, which are not selected as its AIC is higher due to the use of more parameters. As such, we assign multiple copula types with the lowest AIC to a variable pair. Our proposed mixture R-vine copula model is also motivated by Bates & Granger (1969) which concluded that a mixture of projections typically outperforms an individual projection.

We propose two conditions that copula types must meet, before we assign them to a variable pair in addition to the copula type with the lowest AIC. Firstly, at most three copula types may be assigned to each edge. Secondly, the copula type must score at most 1% higher AIC than the copula type with the lowest AIC. We substantiate this $100\% - 1\% = 99\%$ inclusion criterion by our expectation that adding copula types with higher AICs to deteriorate the approximation of the dependency structure. Furthermore, we choose to assign bivariate copula combination within R-vine structures, as R-vines also allow C- and D-vine structures. With this set-up, we incorporate distinct differences of copula types in estimation.

We translate these conditions into our estimation procedure by adapting Step 2 and 3 from Section 4.3.2. In Step 2, we label the best fitting copula type $f_1 = f$ with corresponding parameters $\hat{\Theta}_{f_1;i,j|D}$ based on the AIC. Copula type f_1 is assigned to the corresponding edge $e_{i,j|D}$. Moreover, we examine whether copula types with the second and third lowest AIC, f_2 and f_3, fall within the 1% range from f_1 in terms of AIC. We assign the copula type(s) that suffice(s) these conditions also to edge $i, j|D$.

In Step 3, we adjust the manner that we derive pseudo-observations. In Equation 16, we defined $Z_{i|j,D}$ by transforming all pseudo-observations $(Z_{i|D}, Z_{j|D})$, with function $h_{f;i,j|D}(\cdot,\cdot;\hat{\Theta}_{f;i,j|D})$. If there are multiple copula types assigned to one edge, we derive the pseudo-observations from these copula types with equal weight. This does not aggravate the computational burden, while we add copula types which we consider to have an equivalent fit. If two copula types are assigned to an edge, we transform half of the pseudo-observations with $h_{f_1;i,j|D}(\cdot,\cdot;\hat{\Theta}_{f_1;i,j|D})$ and the other 125 pseudo-observations with $h_{f_2;i,j|D}(\cdot,\cdot;\hat{\Theta}_{f_2;i,j|D})$.

For $(Z_{i|D}, Z_{j|D})$, we define CDF and corresponding h-function for f_1 as

$$C_{f_1;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f_1;i,j|D}), \quad \text{with } h_{f_1;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f_1;i,j|D}) = \frac{\partial\, C_{f_1;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f_1;i,j|D})}{\partial\, z_{j|D}}.$$

Moreover, we define the CDF and corresponding h-function for f_2 as

$$C_{f_2;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f_2;i,j|D}), \quad \text{with } h_{f_2;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f_2;i,j|D}) = \frac{\partial\, C_{f_2;i,j|D}(z_{i|D}, z_{j|D}; \hat{\Theta}_{f_2;i,j|D})}{\partial\, z_{j|D}},$$

On observation-level, we assign half of the observations t randomly to set T_{f_1} and the remaining observations to T_{f_2}, such that conditional pseudo-observations are derived from $h_{f_1;i,j|D}(\cdot,\cdot)$ or $h_{f_2;i,j|D}(\cdot,\cdot)$ as follows

$$Z_{i|j,D,t} = \begin{cases} h_{f_1;i,j|D}(Z_{i|D,t}, Z_{j|D,t}; \hat{\Theta}_{f_1;i,j|D}), & \text{if } t \in T_{f_1} \\ h_{f_2;i,j|D}(Z_{i|D,t}, Z_{j|D,t}; \hat{\Theta}_{f_2;i,j|D}), & \text{if } t \in T_{f_2}\ . \end{cases} \tag{17}$$

Similarly, we derive one third of the pseudo-observations from each copula type, if three copula types are assigned to one edge.

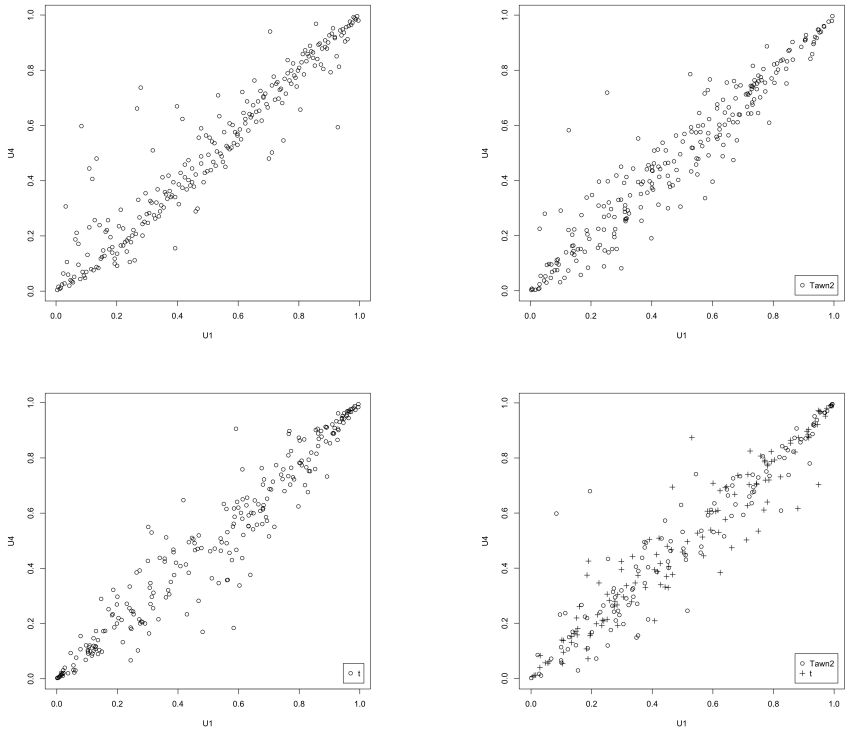

Figure 9: Scatter plots based on variable pair (X_1, X_4), where a.) shows the normalised ranked pairs of the respective variables (upper left panel), b.) provides the simulated dependency structure for the best fitting single copula type-parameters (upper right panel), c.) exhibits the simulated dependency structure for the second best fitting single copula type-parameters (bottom left panel) and d.) presents the simulated dependency structure for the best fitting copula type-parameters mixture.

In Figure 9 we observe the impact of combining bivariate copulae to approximate variable pair (X_1, X_4). Herein, the simulated dependency structure in Figure 9b is derived from the Tawn type 2 copula that had an

AIC of -629.7. The simulated observations in Figure 9c are generated from the Student-t copula that scored an AIC of -629.3 on the normalised pairs of X_1 and X_4 in Figure 9a. Transforming pseudo-observations with both copula types, as described in Equation (17), results in Figure 9d. These figures show that copula types with similar AICs can have different dependency structures (see Figures 9b and 9c). Figure 9d resembles both the outer pseudo-observations above the diagonal as in Figure 9b as well as the tail density as in Figure 9c.

With respect to the vine tree structure, the first tree is always identical to the standard R-vine copula, as this is defined on the Kendall's $\hat{\tau}$ metric over the same empirical marginal CDFs. However, higher order tree structures could differ, as these are also based on the different pseudo-observations possibly derived with the copulae combinations in the earlier trees.

4.4 Sampling

In the previous section, we defined the bivariate copulae, vine copula and vine copula mixture models, as well as the corresponding estimation procedures. In this section, we provide the sampling procedure for these three instances. These sampled multivariate distributions follow the dependence structure of the six variables as $(Z_1, ..., Z_6)$ as in the proposed models.

4.4.1 Bivariate copula sampling in the single commodity hedging framework

With the specifications derived in Section 4.3.1, we sample the 2-dimensional joint distribution of each of the commodity spot and futures price changes. We apply a Monte Carlo simulation method by drawing observations six times from a standard uniform distribution, meaning $U_i \sim U(0,1)$ for $i = \{1, ..., 6\}$. These observations are not tied to a marginal distribution of one of the variables. Subsequently, we apply the inverse of the $h(\cdot)$-function given in Equation 16 to impose proper cross-sectional dependence. The inverted h-function is the probability integral transform of Rosenblatt (1952), which generates samples from the copula, as follows

$$Z_i = U_i,$$

$$Z_j = h^{-1}_{f;i,j}(U_j, U_i),$$

where U_i, U_j are the simulated distributions of variables i, j which are transformed such that $(Z_i, Z_j) \sim C_{f;i,j}$.

We provide an example of how to draw the two marginal distributions from edge $e_{1,4}$ in Figure 6. The marginal distribution of, in this case, the spot price changes of crude oil is simulated as $Z_1 = U_1$. We simulate Z_4 with copula type $C_{f;1,4}$ and copula parameter(s) $\hat{\Theta}_{f;1,4}$ determined during the estimation procedure in Section 4.3.1, from the inverted $h(\cdot)$-function,

$$Z_4 = h^{-1}_{f;1,4}(U_2, U_1; \hat{\Theta}_{f;1,4}), \qquad \text{where } h_{f;1,4}(z_i, z_j) = \frac{\partial\, C_{f;1,4}(z_i, z_j; \hat{\Theta}_{f;1,4})}{\partial\, z_j}.$$

It does not matter whether we use U_2 or U_4, as both present observations standard uniformly distributed. The standard uniform variables Z_2, Z_5 and Z_3, Z_6 of variable pairs X_2, X_5 and X_3, X_6 can be simulated in similar fashion.

4.4.2 Vine copula sampling

We draw the 6-dimensional joint distribution from the determined model consisting of vine tree sequence V, its copula types and corresponding copula parameters. To derive the univariate standard uniform samples Z_i for variable i, we apply the sampling algorithm from Dissmann et al. (2013). This sampling algorithm can be applied from multiple starting nodes and is therefore not unique.

To demonstrate the sampling procedure, we explain it in a 4-dimensional setting, as shown in Figure 10. This vine copula has both a star-shaped in T_1 graph and a path-graph in T_2, giving a complete overview. Similar to Section 4.4.1, we draw observations four times from a standard uniform distribution, meaning $U_i \sim U(0,1)$ for $i = \{1, ..., 4\}$. These observations drive the randomness, while the sampling algorithm ensures the multivariate dependency structure of the transformed draws mimics the historical dependency.

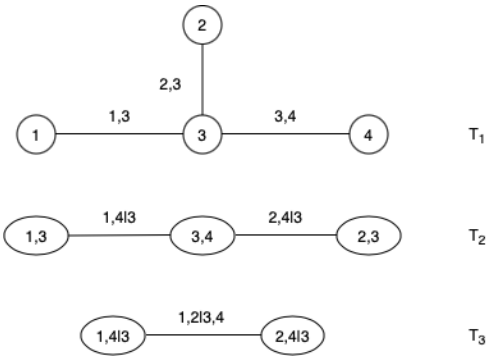

Figure 10: An example of a 4-dimensional R-vine copula.

We initialise the sampling algorithm in T_1 by assigning a simulated observations vector U_1 to standard uniform variable 4, such that

$$Z_4 = U_1.$$

Hereafter, we draw the second variable sharing edge $e_{3,4}$ with variable 4 by applying the probability integral transform, as

$$Z_3 = h^{-1}_{f;3,4}(U_2, U_1), \qquad \text{where } h_{f;3,4}(z_i, z_j) = \frac{\partial\, C_{f;3,4}(z_i, z_j; \hat{\Theta}_{f;3,4})}{\partial\, z_j}. \tag{18}$$

Furthermore, $\hat{\Theta}_{f;3,4}$ presents the estimated copula parameter(s) of this edge. To derive the other draws, we determine the conditional standard uniform variable $Z_{4|3}$ in higher tree T_2 at node $n_{3,4}$ as

$$Z_{4|3} = h_{f;3,4}(Z_4, Z_3),$$

where the same h-function as in Equation (18) is used. From node $n_{3,4}$, we traverse over edge $e_{2,4|3}$ by

deriving $Z_{2|3}$, representing node $n_{2,3}$, with our third simulated observations vector U_3, as

$$Z_{2|3} = h^{-1}_{f;2,4|3}(U_3, Z_{4|3}), \qquad \text{where } h_{f;2,4|3}(z_i, z_j) = \frac{\partial\, C_{f;2,4|3}(z_i, z_j; \hat{\Theta}_{f;2,4|3})}{\partial\, z_j}. \tag{19}$$

With $Z_{2|3}$, we can derive Z_2, as follows

$$Z_2 = h^{-1}_{f;2,3}(Z_{2|3}, Z_3), \qquad \text{where } h_{f;2,3}(z_i, z_j) = \frac{\partial\, C_{f;2,3}(z_i, z_j; \Theta_{f;2,3})}{\partial\, z_j}.$$

Moreover, we derive $Z_{2|3,4}$ belonging to $n_{2,4|3}$ with the same h-function as in Equation (19), as

$$Z_{2|3,4} = h_{f;2,4|3}(Z_{2|3}, Z_{4|3}).$$

With the derivation of $Z_{2|3,4}$ we have arrived in T_3. We derive the other node in tree T_3 with remaining observations U_4 as

$$Z_{1|3,4} = h^{-1}_{f;1,2|3,4}(U_4, Z_{2|3,4}), \qquad \text{where } h_{f;1,2|3,4}(z_i, z_j) = \frac{\partial\, C_{f;1,2|3,4}(z_i, z_j; \Theta_{f;1,2|3,4})}{\partial\, z_j}.$$

With these transformed observations, we can move down the vine to node $n_{1,3}$ in T_2 and derive $Z_{1|3}$ as

$$Z_{1|3} = h^{-1}_{f;1,4|3}(Z_{1|3,4}, Z_{4|3}), \qquad \text{where } h_{f;1,4|3}(z_i, z_j) = \frac{\partial\, C_{f;1,4|3}(z_i, z_j; \Theta_{f;1,4|3})}{\partial\, z_j}.$$

Ultimately, Z_1 in T_1 is derived as follows

$$Z_1 = h^{-1}_{f;1,3}(Z_{1|3}, Z_3), \qquad \text{where } h_{f;1,3}(z_i, z_j) = \frac{\partial\, C_{f;1,3}(z_i, z_j; \Theta_{f;1,3})}{\partial\, z_j}.$$

All standard uniform variables $(Z_1, ..., Z_4)$ follow the dependency structure of the vine copula model and approximate that of the observed variables $(X_1, ..., X_4)$.

In a similar fashion, we can generate a random sample from 6-dimensional vine copulae. We sample six independent standard uniform variables $(U_1, ..., U_6)$ and set the first node of the variable equal to $Z_1 = U_1$. Then draw the second standard uniform by the probability transform integral, $Z_2 = h^{-1}_{f;1,2}(U_2|U_1)$. Hereafter, proceed to the next tree by deriving the conditional marginal distribution $Z_{1|2} = h_{f;1,2}(h^{-1}_{f;1,2}(U_1|U_2)|U_2)$ and the node that shares an edge with this node $n_{1,2}$: $Z_{3|2} = h^{-1}_{f;1,3|2}(U_3|h_{f;1,2}(h^{-1}_{f;1,2}(U_1|U_2)|U_2))$. Subsequently, the next standard uniform variable is defined as: $Z_3 = h^{-1}_{f;2,3}(h^{-1}_{f;1,3|2}(U_3|h_{f;1,2}(h^{-1}_{f;1,2}(U_1|U_2)|U_2))|U_2)$. We proceed until all six standard uniform variables are simulated.

4.4.3 Sampling from bivariate copula mixtures in the mixture R-vine model

To generate the draws of the six standard uniform variables for our mixture R-vine model, we traverse the vine structure similarly as in Section 4.4.2. The difference lies in the sampling on the bivariate level, which is similar to the procedure of drawing conditional pseudo-observations in Section 4.3.3. From this estimation procedure, we know the fitted bivariate copulae combination entailing $C_{f_1,i,j}$ and possibly $C_{f_2,i,j}$ and $C_{f_3,i,j}$ for each edge. Herein, when sampling an edge, we assign each draw $z_{i,s}$ from Z_i to a copula type f_v through s. As such, similar to the procedure in Section 4.3.3, half of the samples s are randomly assigned to set

S_{f_1} and the other half are assigned to set S_{f_2}, in case two copula types are assigned to one edge. If three copula types are assigned to one edge, then one third of the samples are assigned to each copula type C_{f_v} for $v \in \{1, 2, 3\}$. A sample for Z_i is derived with copula type C_{f_v} as follows

$$Z_{i,s} = h^{-1}_{f_v;i,j}(Z_{i|j,t}, Z_{j,t}), \qquad \text{where } \frac{\partial\, C_{f_v,ij}(z_{i|j,s}, z_{j,s}; \hat{\Theta}_{f_v,ij})}{\partial\, z_{j,s}}, \quad \text{if } s \in S_{f_v}.$$

We apply this for every s, until all paired draws in $(Z_{i|j}, Z_j)$ are transformed.

4.5 Conversion to weekly price changes

In Section 4.4, we have shown how to generate the draws $(Z_1, ..., Z_6)$ from the copula density for both the bivariate copula and vine copula model. We apply the quantile function from the price change time series in Equation (10) to convert these draws to simulated observations from the joint distribution of price changes $(\tilde{X}_1, ..., \tilde{X}_6)$, as

$$\tilde{X}_i = \hat{F}_i^{-1}(Z_i)$$

where $\hat{F}_i^{-1}$ is the inverse of the empirical distribution function of Equation (10).

4.6 Downside risk measure objectives

In this section, we estimate the hedge ratios λ introduced in the hedging frameworks in Section 4.1. For the single commodity hedging framework, we determine λ^{C*}, λ^{G*} and λ^{H*} for commodity draws separately. In the fixed proportion and flexible multi-commodity framework, the hedge ratios are determined over the profit margins $Y(\lambda)$ from Equations 5 and 2, respectively. The observations of these profit margins are based all six of the price changes draws $(\tilde{X}_1, ..., \tilde{X}_6)$ jointly.

We optimise the hedge ratios for eight downside risk-objectives: minimum-VaR (at 1%, 5% and 10%), minimum-ES (at 1%, 5% and 10%), minimum-Lower Partial Moments (LPM_2 and LPM_3). Herein, we minimise the downside risk measures of $Y(\lambda)$ by performing the numerical Nelder-Mead search method (Nelder & Mead (1965)), where only the hedge ratios can be adjusted. This results in optimal hedge ratios.

The downside risk measures are defined in Equations (20), (21), (22) and (23) for confidence level α. This allows us to evaluate the hedge ratios' sensitivity to the downside risk measures. Value-at-Risk (VaR) is a quantile and measures the largest potential loss over a certain period of time, which is one week in our paper, as

$$VaR_\alpha = F_L^{-1}(\alpha), \tag{20}$$

where α is the confidence level $(0, 1)$. Expected Shortfall (ES) measures the expected loss given that losses exceed the VaR, therefore this risk measure is also known as the conditional VaR, as

$$ES_\alpha = \mathbb{E}(Y | Y \leq VaR_\alpha) = \frac{1}{1-a} \int_\alpha^1 VaR_u du, \tag{21}$$

where α is the confidence level $(0, 1)$. The second Lower Partial Moment (LPM_2), which is also known as semivariance, measures the variability of $Y(\lambda)$ that falls below zero, as

$$LPM_2 = \int_{-\infty}^{0} -Y^2 dF(Y). \tag{22}$$

Furthermore, we measure the third Lower Partial Moment (LPM_3) as

$$LPM_n = \int_{-\infty}^{0} (-Y)^n dF(Y), \tag{23}$$

where $n = 3$ signals the is the risk tolerance of the hedger. Herein, $n = 1$ represents risk-neutral behavior and $n > 1$ represents risk-averse behavior. Thus, LPM_3 penalises losses in the profit margin $Y(\lambda)$ stronger than LPM_2.

4.7 Performance measures

With the model specifications derived in Section 4.3 and derived hedge ratios from Section 4.6, we can assess which hedge strategy has the best performance for a 130-week out-of-sample test window.

We test the optimal hedge ratios out-of-sample by measuring hedge effectiveness (Ederington (1979)) for a test period of 130 weeks following the estimation period, as

$$\text{Hedge Effectiveness} = \left(1 - \frac{Risk(Y(\hat{\lambda}^*))}{Risk(Y(0))}\right),$$

where $Y(\hat{\lambda})$ is the income of the hedged portfolio (as given in Equations (4), (5) and (2)) for the optimised hedge ratios $\hat{\lambda}$ and $Y(0)$ the income of the portfolio that was not hedged. Furthermore, $Risk$ is the downside risk measure of concern.

Although hedge effectiveness against the unhedged portfolio allows us to compare the hedge strategies, we do not deem this strategy a competing benchmark. Therefore, we also compare our strategies against a naive hedging strategy $Y(1)$, which holds a hedge ratio of $\lambda = 1$, for every commodity and every rolling time window. This strategy states that a trader with a long (short) position in the spot market should simultaneously sell (buy) an equal amount of futures.

To test if a hedging strategy is significantly better, we apply the paired-t test. Via this paired t-test, we examine the performance in hedge effectiveness between hedging strategies per downside risk measure. We state the null hypothesis that the out-of-sample hedging effectiveness for two hedging strategies is equal, as

$$H_0 : \mu_d = 0 \quad H_1 : \mu_d \neq 0,$$

where H_1 gives the alternative hypothesis, μ_d represents the true mean difference between the hedge effectiveness of the hedging strategies. We compute the test statistic as follows

$$t = \frac{\bar{d} - 0}{\hat{\sigma}}\sqrt{n} \sim t_{n-1}(\alpha/2),$$

where $\bar{d}$ represents the mean of 1452 sample differences between the hedging effectiveness of strategy A and strategy B. Furthermore, $\hat{\sigma}$ provides the sample standard deviation of the differences between the hedge effectiveness of the tested hedging strategies. Moreover, $n = 1452$ is the number of out-of-sample test observations. The critical value for the test statistic with two-tailed significance $\alpha/2$ is derived from the t-distribution with $n - 1$ degrees of freedom.

As we roll through our time window, we attain a sample to compare three hedging frameworks for eight downside risk measures over time periods differing in volatility and trends. Herein, the fixed and flexible multi-commodity hedging frameworks are modeled with four different vine copula models.

5 Results

In this section we examine which of the hedging strategies discussed in the previous section performs best. Firstly, in Section 5.1, we discuss the in-sample fit of our models. Hereafter, we analyse which vine copula model is preferred for modelling the joint distribution in Section 5.2 by comparing out-of-sample HE. In Section 5.3, we show which hedging strategy has superior out-of-sample performance. Ultimately, we provide further analysis of our models in Section 5.4. Herein, we observe the drivers of the difference between the fixed proportion and flexible hedging frameworks, the sensitivity to an extreme observation, the impact of reducing the rebalancing of hedge ratios, the consequence of reducing the number of draws and the sensitivity of the mixture R-vine copula model to the inclusion criterion.

5.1 In-sample fit

We observe that on average the R-vine copula models have the highest log-likelihood with 1343.7 to the estimation sample in Table 3. In Section 4.3, we discussed the differences between the applied models, wherein we noted that vine copulae estimate 15 bivariate dependency structures against 3 in the single commodity hedging framework. In Table 3, we determine that the number of parameters increases faster than the log-likelihood and information criteria on average. This emphasises that modelling the dependencies between the commodities spot and futures price changes is at the core of the in-sample fit of the models.

Table 3: Average log-likelihood, AIC, BIC and average number of parameters of the single commodity hedging framework, C-,D-, R- and mixture R-vine copula models over the rolling window samples.

	Log-likelihood	AIC	BIC	Number of parameters
Single commodity	1013.4	-2014.8	-1998.1	6
C-vine	1314.4	-2584.7	-2523.3	22
D-vine	1337.8	-2632.3	-2571.7	22
R-vine	1343.7	-2645.0	-2585.8	21
Mixt. R-vine	1342.1	-2629.1	-2531.9	28

Table 4: Average hedge ratios and standard deviations (in brackets) for each commodity within the hedging frameworks with 3 hedge ratios. The hedge ratio trajectories of the single commodity and flexible multi-commodity framework for each risk measure are presented in Appendix C, respectively.

	VaR 1%	VaR 5%	VaR 10%	ES 1%	ES 5%	ES 10%	LPM_2	LPM_3
	Crude oil							
Single commodity	1.009	1.034	1.044	0.932	1.017	1.027	1.021	0.957
	(0.073)	(0.038)	(0.035)	(0.109)	(0.050)	(0.038)	(0.032)	(0.063)
C-vine	1.049	1.083	1.088	0.945	1.037	1.056	1.044	0.972
	(0.165)	(0.124)	(0.122)	(0.288)	(0.176)	(0.156)	(0.161)	(0.255)
D-vine	1.047	1.084	1.091	0.940	1.041	1.064	1.044	0.961
	(0.150)	(0.104)	(0.103)	(0.296)	(0.169)	(0.140)	(0.154)	(0.259)
R-vine	1.085	1.094	1.094	0.991	1.081	1.090	1.073	1.011
	(0.154)	(0.120)	(0.114)	(0.326)	(0.175)	(0.151)	(0.160)	(0.273)
Mixt. R-vine	1.077	1.095	1.096	0.999	1.082	1.090	1.074	1.019
	(0.143)	(0.119)	(0.114)	(0.272)	(0.163)	(0.143)	(0.148)	(0.228)
	Gasoline							
Single commodity	1.036	1.068	1.074	1.019	1.061	1.068	1.070	1.082
	(0.113)	(0.072)	(0.064)	(0.182)	(0.085)	(0.071)	(0.069)	(0.170)
C-vine	1.042	1.094	1.089	0.998	1.069	1.084	1.070	1.060
	(0.130)	(0.112)	(0.100)	(0.142)	(0.112)	(0.103)	(0.101)	(0.122)
D-vine	1.044	1.101	1.099	1.005	1.078	1.092	1.077	1.053
	(0.195)	(0.110)	(0.090)	(0.242)	(0.146)	(0.117)	(0.122)	(0.182)
R-vine	1.056	1.097	1.095	1.023	1.087	1.096	1.086	1.076
	(0.159)	(0.111)	(0.091)	(0.176)	(0.122)	(0.107)	(0.106)	(0.135)
Mixt. R-vine	1.058	1.102	1.097	1.032	1.092	1.100	1.090	1.085
	(0.156)	(0.112)	(0.092)	(0.180)	(0.121)	(0.108)	(0.108)	(0.142)
	Heating oil							
Single commodity	1.071	1.063	1.032	1.055	1.072	1.056	1.046	1.041
	(0.120)	(0.082)	(0.042)	(0.161)	(0.099)	(0.070)	(0.064)	(0.123)
C-vine	1.075	0.987	0.975	1.057	0.996	0.986	0.991	0.983
	(0.289)	(0.257)	(0.223)	(0.487)	(0.367)	(0.332)	(0.320)	(0.421)
D-vine	1.038	1.042	1.045	0.965	1.009	1.032	1.021	0.976
	(0.313)	(0.205)	(0.186)	(0.557)	(0.360)	(0.290)	(0.298)	(0.456)
R-vine	1.107	1.072	1.067	1.102	1.095	1.092	1.082	1.081
	(0.263)	(0.220)	(0.200)	(0.498)	(0.340)	(0.293)	(0.293)	(0.412)
Mixt. R-vine	1.122	1.070	1.061	1.155	1.113	1.098	1.091	1.115
	(0.300)	(0.218)	(0.191)	(0.541)	(0.360)	(0.300)	(0.299)	(0.418)

The C-vine copula has one variable in each tree that is connected with all other variables, hence the C-vine only models one spot-futures pair directly in the first tree. Moreover, we observe that the difference between the D- and R-vine is small, although the R-vine allows for $22,680$ more vine tree structures. We contribute this to the fact that the D-, R- and mixture R-vine include all three spot-futures pair directly in the first tree for all 1452 samples. We conclude that limiting the C-vine from modelling the spot-futures

Table 5: Average hedge ratios and standard deviations (in brackets) for the fixed proportion hedging framework with one hedge ratio. The hedge ratio trajectories of the fixed proportion hedging framework derived over the R-vine copula for each risk measure are presented in Appendix C.

	VaR 1%	VaR 5%	VaR 10%	ES 1%	ES 5%	ES 10%	LPM_2	LPM_3
C-vine	1.057	1.078	1.071	1.030	1.063	1.067	1.058	1.074
	(0.137)	(0.106)	(0.098)	(0.165)	(0.119)	(0.108)	(0.111)	(0.154)
D-vine	1.049	1.097	1.094	1.001	1.069	1.083	1.071	1.060
	(0.182)	(0.106)	(0.094)	(0.233)	(0.143)	(0.119)	(0.125)	(0.204)
R-vine	1.084	1.102	1.094	1.069	1.097	1.098	1.091	1.113
	(0.163)	(0.114)	(0.100)	(0.201)	(0.137)	(0.121)	(0.120)	(0.178)
Mixt. R-vine	1.093	1.104	1.096	1.077	1.102	1.101	1.094	1.118
	(0.164)	(0.117)	(0.100)	(0.203)	(0.137)	(0.122)	(0.122)	(0.185)

pairs in the first tree hampers its log-likelihood score. Based on the scores in Table 3, we expect the D-, R- and mixture R-vine models to determine the joint dependency structure of the variables the best. Therefore, the D-, R- and mixture R-vine model are expected to forecast the most accurate draws.

In Tables 4 and 5, we observe that the average hedge ratios are slightly higher than 1 for most models and risk measures. This means that on average the oil refinery holds more futures than the commodities it eventually buys or sells against the spot price. As such, an average hedge ratio of 1.02 means that it holds a position that is 2% larger in futures (either long or short) than in the commodity.

We note that for ES 1% the hedge ratios are the lowest on average for all commodities except heating oil in Tables 4 and 5. Furthermore, we find that the standard deviation of the heating oil hedge ratios is the highest for every vine copula, which is likely to follow from the fact that heating oil has the lowest weight (one third) in the joint distribution of the profit margin in Table 4. Subsequently, a more articulate position in heating oil is needed for similar impact on the profit margin. For the single commodity hedging framework, we do not observe this phenomenon, because the hedge ratios are determined separately. Within this framework, we also notice that the volatility of the hedge ratios is related to the volatility of the price changes of the underlying commodity, as the least volatile commodity, crude oil, also has the least volatile hedge ratios for each risk measure objectives in Table 4. With regard to the mixture R-vine model, we observe that the average hedge ratios are close to the average hedge ratios of the R-vine model, especially for crude oil and gasoline in Table 4 and the fixed proportion framework in Table 5. The average hedge ratios and standard deviation for heating oil in Table 4 are generally higher than those of the R-vine model. This suggests that through the mixture R-vine model, the model holds a more protective position in heating oil futures. Lastly, we note for the fixed proportion hedging framework in Table 5 that all average hedge ratios are greater than 1 combined with low standard deviation relative to the hedge ratios of the corresponding vine copula in the flexible multi-commodity hedging framework in Table 4.

5.2 Vine copula comparison

In this section, we examine which of the vine copulae attains the best hedging results by modelling the multivariate joint distribution for the fixed proportion and flexible multi-commodity hedging framework. In Section 5.1, we deduced that the R-vine copula has the best in-sample fit in terms of average AIC followed closely by the mixture R-vine and D-vine copulae. We observe a similar ranking in out-of-sample performance in within the flexible multi-commodity framework in Table 6 and the fixed proportion framework in Table 7.

Table 6: Mean and median hedge effectiveness of the C-, D-, R- and mixture R-vine model for eight downside risk measures, where the hedge ratios were determined by the flexible multi-commodity framework. Bold numbers are the highest values per downside risk measure.

		VaR 1%	VaR 5%	VaR 10%	ES 1%	ES 5%	ES 10%	LPM_2	LPM_3
C-vine	Mean	20.333	39.315	42.397	18.351	30.677	35.726	55.180	52.029
	Median	*19.769*	*39.008*	*42.963*	*20.927*	*32.702*	*36.304*	*56.406*	***66.213***
D-vine	Mean	21.809	**39.607**	**43.003**	**19.436**	30.877	35.952	55.310	52.824
	Median	***20.823***	*38.984*	*42.716*	*20.212*	*32.369*	***36.889***	*57.053*	*64.387*
R-vine	Mean	21.803	39.525	42.876	19.257	30.933	35.995	55.406	52.213
	Median	*20.450*	*39.066*	***43.222***	*19.921*	*32.620*	*36.739*	*56.776*	*64.350*
Mixt. R-vine	Mean	**21.906**	39.601	42.896	19.911	**30.966**	**36.075**	**55.580**	**54.026**
	Median	*20.629*	***39.101***	*43.079*	***21.002***	***32.951***	*36.784*	***57.135***	*64.951*

A mean hedge effectiveness (HE) of 20.333 achieved by the C-vine model for the risk measure VaR 1% means that the C-vine model on average reduces VaR 1% with 20.333% more than the unhedged portfolio, where the hedge ratios equal zero. We compare the HE metrics across models directly, because the measures are relative to the same unhedged portfolio. For example, if we compare mean HE between the C- and D-vine for ES 5%, we find that the hedge based on D-vine copula results in $30.877 - 30.677 = 0.200$ HE percentage points (pp) higher. Moreover, the median HE represents the HE observation separating the higher half from the lower half. Although we focus on the mean HE, we are able to recognise a heavily skewed mean by a period of extremely low HE, if the mean is much lower than the median. This is the case for LPM_3 in Table 6, for example.

In the flexible multi-commodity framework, we see that the mixture R-vine copula achieves the highest average HE for five risk measures, while the D-vine performs the best for the three remaining VaR risk measures (see Table 6). The t-statistics in Table 8 show that the HE of the C-vine copula model is significantly worse than the all other vine copula models for all risk measures, except LPM_3. Furthermore, we observe that the performance of the D- and R-vine is close, such that we can not state that the mean HE is unequal for five of the risk measures. Although the D-vine achieves the highest average HE for the three VaR risk measures, it is only significantly better for VaR 10% with respect to the R- and mixture R-vine models. We observe that the mixture R-vine model scores significantly better than all other vine copula models on four non-VaR risk measures in Table 8. As such, if we are restricted to pick one vine copula model for all risk

measures within the flexible commodity framework, then we prefer the mixture R-vine model.

Table 7: Mean and median HE of the C-, D- and R-vine model for eight risk measures, where the hedge ratios were determined following the fixed proportion framework. Bold numbers are the highest mean and median values per risk measure.

		VaR 1%	VaR 5%	VaR 10%	ES 1%	ES 5%	ES 10%	LPM_2	LPM_3
C-vine	Mean	21.487	40.214	43.505	22.332	31.911	36.753	56.838	63.817
	Median	*21.318*	*40.233*	*44.154*	*22.269*	*32.429*	*37.583*	*58.299*	*70.493*
D-vine	Mean	**22.883**	40.190	**43.569**	**23.704**	32.255	36.828	56.860	64.040
	Median	***23.107***	*39.962*	***44.671***	***25.307***	*33.135*	***37.880***	*58.605*	*69.749*
R-vine	Mean	22.718	**40.251**	43.506	23.590	**32.432**	**36.984**	**57.054**	**64.549**
	Median	*22.567*	***40.256***	*44.308*	*24.246*	*33.566*	*37.661*	***58.819***	***70.460***
Mixt. R-vine	Mean	22.734	40.186	43.534	23.570	32.400	36.964	57.010	64.348
	Median	*22.631*	*40.028*	*44.437*	*24.652*	***33.670***	*37.660*	*58.798*	*70.365*

In the fixed proportion hedging framework, we also observe the inferior out-of-sample hedging performance of the C-vine relative to the other vine copulae in Table 7. Contrary to the flexible multi-commodity hedging framework, we examine that the R-vine achieves on average the highest HE for five of the risk measures, whereas the D-vine achieves this for three risk measures. For the VaR 1%, 5% and 10% as well as the ES 1% risk measures we do not reject the null hypothesis that the mean HE of the D-, R- and mixture R-vines are equal in Table 8.

Although the absolute differences in mean HE between the R- and the mixture R-vine models for the remaining four risk measures are small, the t-test indicates a significantly better hedging performance by the R-vine model. This is due to the fact that the differences in HE between these two models are tested throughout our time period of 1452 weeks.

Thus, for the fixed proportion hedging framework, we conclude that the R-vine either achieves better or equivalent hedging performance than the other vine copula models for each risk measure. Therefore, we determine the hedge ratios in the fixed proportion commodity hedging framework over the joint distribution forecasted from the R-vine model. Furthermore, we derive the hedge ratios in the flexible multi-commodity framework from the mixture R-vine copula, when comparing hedging strategies in the next section.

Table 8: T-statistics from two-tailed paired t-test, where the positive (negative) t-statistic indicates a better performance for the model left (top) of the statistic. Left hand side provides the t-statistic based on the HE in flexible multi-commodity hedging framework, whereas the right hand side shows the HE in the fixed proportion hedging framework. The asterisks, *, ** and ***, represent the rejection of the equal HE at 10%, 5% and 1% significance, respectively.

VaR 1%	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-7.095***	-7.830***	-8.747***	-9.431***	-9.968***	-10.609***
D-vine	-	0.046	-0.643	-	1.458	1.290
R-vine	-	-	-0.920	-	-	-0.202
VaR 5%	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-3.179***	-2.462**	-3.196***	0.369	-0.569	0.461
D-vine	-	1.087	0.079	-	-1.103	0.080
R-vine	-	-	-1.107	-	-	1.390
VaR 10%	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-6.376***	-5.226***	-5.357***	-1.005	-0.003	-0.495
D-vine	-	1.785*	1.543	-	1.443	0.731
R-vine	-	-	-0.281	-	-	-0.693
ES 1%	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-4.068***	-3.681***	-6.575***	-6.786***	-8.479***	-9.205***
D-vine	-	1.029	-2.537**	-	0.853	0.904
R-vine	-	-	-4.851***	-	-	0.310
ES 5%	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-2.122**	-3.013***	-3.400***	-6.855***	-12.376***	-11.634***
D-vine	-	-1.058	-1.542	-	-5.220***	-4.133***
R-vine	-	-	-1.166	-	-	2.402**
ES 10%	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-4.298***	-5.963***	-7.462***	-2.673***	-9.899***	-8.944***
D-vine	-	-1.330	-3.615***	-	-8.787***	-7.345***
R-vine	-	-	-4.173***	-	-	2.691***
LPM_2	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-1.649*	-3.326***	-5.936***	-0.596	-8.166***	-6.422***
D-vine	-	-2.296**	-5.627***	-	-7.825***	-5.798***
R-vine	-	-	-6.262***	-	-	4.600***
LPM_3	D-vine	R-vine	Mixt. R-vine	D-vine	R-vine	Mixt. R-vine
C-vine	-1.946*	-0.581	-4.677***	-1.848*	-8.452***	-6.083***
D-vine	-	2.012**	-3.559***	-	-6.493***	-3.574***
R-vine	-	-	-5.922***	-	-	5.902***

5.3 Out-of-sample performance of hedging strategies

In this section, we discuss and compare the hedging performance of the hedge ratio estimations within the single commodity hedging framework based on 3 paired copulae, the fixed proportion hedging framework based on the R-vine copula and flexible multi-commodity hedging framework based on the mixture R-vine copula. We abbreviate each of these hedging strategies as the single, fixed and flexible hedge. Moreover, we

compare these hedges against our benchmark hedging strategy: the naive hedge.

Table 9 shows that the fixed hedge achieves the best out-of-sample hedging performance for all risk measures, except VaR 10%. Furthermore, we see that the naive hedge is a competitive strategy, as it ranks best for VaR 10% and second-best for ES 1%, ES 5%, ES 10%, LPM_2 and LPM_3. Least preferred is the single hedge, because it has the lowest mean HE for five risk measures.

Table 9: Mean and median HE of the four hedging strategies for eight risk measures, where the single hedging strategy is based on 3 paired copulae, the fixed hedging strategy is based on the R-vine copula model and the flexible hedging strategy is based on the mixture R-vine model.

	Naive		Single		Fixed		Flexible	
	Mean	*Median*	Mean	*Median*	Mean	*Median*	Mean	*Median*
VaR 1%	20.865	*19.910*	20.261	*20.185*	**22.718**	***22.567***	21.906	*20.629*
VaR 5%	39.203	*39.253*	39.057	*38.632*	**40.251**	***40.256***	39.601	*39.101*
VaR 10%	**43.657**	***44.643***	43.363	*43.666*	43.506	*44.308*	42.896	*43.079*
ES 1%	20.865	*19.910*	18.023	*20.354*	**23.590**	***24.246***	19.911	*21.002*
ES 5%	31.086	*31.679*	30.476	*31.827*	**32.432**	***33.566***	30.966	*32.951*
ES 10%	36.172	*36.770*	35.910	*36.380*	**36.984**	***37.661***	36.075	*36.784*
LPM_2	56.216	*56.782*	55.454	*56.766*	**57.054**	***58.819***	55.580	*57.135*
LPM_3	63.078	*68.744*	56.324	*57.085*	**64.549**	***70.460***	54.026	*64.951*

We conclude from Table 9 that risk reduction with respect to the unhedged portfolio is more difficult when measuring risk over more extreme downside observations, such as VaR 1% and ES 1%. For both VaR and ES, we see a positive relation between observations considered and mean HE. In our instance, VaR 1%, 5% and 10% consider the profit margin of the lowest, 6th lowest and 13th lowest observation in the test window, respectively. Whereas, ES 1%, 5% and 10% measure the average profit margin of the lowest, 6 lowest and 13 lowest observations in the test window, respectively. The LPM_2 and LPM_3 risk measures consider the moments of all negative profits in the test windows, where LPM_3 penalises extreme negative values from the profit margin stronger than LPM_2. Supposing half of the profit margins is negative for the observations in the test window, then the LPM risk measures are determined over these 65 negative observations. Hence, it appears that measuring with a risk metric that concerns more negative observations leads to a larger risk reduction relative to no hedging.

Due to the fact that we traverse 1452 test windows, the seemingly small differences in average HE between the hedges are almost all significantly unequal with 90%, 95% and 99% confidence for the paired t-tests in Table 10.

Table 10: T-statistics from two-tailed paired t-test, comparing four hedging strategies in pairs, which consist of the strategies to the left and directly above the statistic. The positive (negative) t-statistic indicates a better performance for the model left (top) of the statistic. The asterisks, *, ** and ***, represent the rejection of the equal HE at 1%, 5% and 10% significance, respectively.

VaR 1%	Single	Fixed	Flexible
Naive	4.457***	-15.086***	-8.216***
Single	-	-12.229***	-11.537***
Fixed	-	-	5.440***
VaR 5%	Single	Fixed	Flexible
Naive	1.333	-9.395***	-3.393***
Single	-	-12.802***	-5.831***
Fixed	-	-	8.275***
VaR 10%	Single	Fixed	Flexible
Naive	3.444***	1.992**	7.697***
Single	-	-1.632	5.161***
Fixed	-	-	6.985***
ES 1%	Single	Fixed	Flexible
Naive	7.898***	-15.167***	2.890***
Single	-	-13.042***	-4.816***
Fixed	-	-	11.995***

ES 5%	Single	Fixed	Flexible
Naive	8.335***	-20.575***	1.586
Single	-	-20.554***	-6.019***
Fixed	-	-	24.657***
ES 10%	Single	Fixed	Flexible
Naive	6.608***	-18.390***	1.795*
Single	-	-26.602***	-3.488***
Fixed	-	-	26.437***
LPM_2	Single	Fixed	Flexible
Naive	11.223***	-19.164***	7.094***
Single	-	-24.589***	-1.965**
Fixed	-	-	21.610***
LPM_3	Single	Fixed	Flexible
Naive	16.252***	-14.489***	11.657***
Single	-	-19.692***	4.453***
Fixed	-	-	13.943***

5.3.1 Value-at-Risk

For the VaR 1% risk measure, we observe that the HE for the naive hedge is constant for many consecutive test windows with approximately one shift per year in Figure 11 (top). The other hedges follow these shocks, which suggests that the shifts originate from change in risk measured in the unhedged portfolio. Besides the simultaneous shifts, we notice negative HE during two stints: from June 2001 to March 2003 and from April 2013 to March 2015. During the first stint, there were two occurrences of the worst unhedged returns, where the spot price of crude oil rose harder than petroleum products combined. Both returns driving the HE below zero happened simultaneously with a crude oil futures price increase that was weaker than the futures price changes of the petroleum products. Hence, the margins of the spot and futures prices moved in opposite directions, which is called basis risk. In these occasions, the only possibility to compensate losses is by taking a negative position in futures, which is a speculative position. We find that these constant negative HE stints only occur for the VaR 1% and ES 1% as these both exclusively measure the single worst observation of the hedged and unhedged portfolios.

We see the three strategies achieving distinctive results at the end of our rolling window period in Figure 13 (top). The hedging performances of the strategies are driven by the hedge ratios for the worst observation in the test windows: the week of 22 April 2020. During this week the price of crude oil tumbled below zero US dollars ($) for the first time ever. Simultaneously, the spot price of gasoline decreased less substantially than crude oil and heating oil. As a result, the unhedged refining margin improved. However, the change for

the futures margin overshadowed this increase, as crude oil decreased with 12$, while gasoline and heating oil declined with 3$ and 8$, respectively. For this instance, a lower position in crude oil futures improves the hedge. We observe in the hedge ratio trajectories in Appendix C that the fixed hedge results in the highest HE during this period (see Figure 11) with the fixed hedge ratio between $0.8 - 0.9$. The flexible hedge realises HE similar to the naive hedge with λ_C around 0.9, λ_H at approximately 1 and λ_G also between $0.8 - 0.9$. The single hedge performs worst, with λ_C over 1, while λ_G and λ_H were set to roughly 0.95.

We examine that all four hedging strategies result in strictly positive HE for VaR 5% in Figure 11 (middle). Moreover, we observe a steep upward trend peaking in mid-2005, which subsequently decreases quickly. We contribute this phenomenon to the occurrence of a few extreme losses for the unhedged portfolio on the back of hurricane Katrina, which we further investigate in Section 5.4.2. These losses are hedged adequately with positions in futures, hence the VaR 5% is substantially lower for the unhedged portfolio and the HE peaks.

For the VaR 10% risk measure, we find in Tables 9 and 10 that the naive hedge has a significantly higher HE than the three 'actively' managed hedging strategies. The naive hedge is most often the strategy with the highest HE (31%) for this risk measure, while he single and flexible hedge attain the lowest HE most often for 25% and 36% of the samples, respectively (see Figure 21 in Appendix D). Herein, we note that the average difference in HE between the strategy with the lowest HE and highest HE is the biggest for the flexible hedge at $4.760pp$. The lower risk reduction from the 'actively' managed hedging strategies suggests that it is difficult to converse the optimal in-sample hedge ratio(s) for the 10% quantile into out-of-sample HE. We contribute this to the fact that it is rarely the case that this 10% quantile observation happens with similar price movements as the 13th lowest observation in the test window. Moreover, we observe a steep upward trend from early-2010 to late-2011 in Figure 11 (bottom). This trend is caused by the unhedged portfolio losing more than 5$ in five weeks between September 2012 and February 2013. These losses enter the test windows for the rolling window samples in early-2010, which boost HE in the following period. In comparison to the VaR 10% hedged portfolios, we notice a less steep climb for the VaR 5% HE, as for the VaR 5% the hedged portfolio is also affected by the losses.

Figure 11: Out-of-sample HE in percentages over time for four hedging strategies: the naive, single, fixed and flexible hedge. The graph displays the reduction in Value-at-Risk (VaR) at 99% (top), 95% (middle) and 90% (bottom) confidence levels for the test windows of 130 weeks starting from the rolling window sample date on the horizontal axis.

5.3.2 Expected Shortfall

When minimising for ES, we reduce the average loss of the 1%, 5% and 10% worst observations of the profit margin, respectively. Tables 9 and 10 present that the fixed hedge achieves a higher mean HE with 99% confidence with respect to every other hedging strategy for each of the ES risk measures.

Figure 12 (top) shows that the strategies with free hedge ratios, the single and flexible hedge, score lower HE for ES 1% in the period spanning $2005 - 2009$. This is caused by the price movements on the back of hurricane Katrina in the last week of August 2005. More specifically, gasoline spot prices decreased with 44\$ in combination with gasoline futures prices declining just 9\$, which triggered the hedge ratios to move strongly for the single and flexible hedges (see Appendix C). Herein, the crude and heating oil hedge ratios of the flexible hedge are even set below zero. Again, we examine the effect of excluding the Katrina observation in the estimation window in Section 5.4.2.

For the ES 5% and 10% risk measures, we observe that the hedging strategies follow a similar pattern in HE in Figure 12 (middle and bottom) as well as in hedge ratios in Appendix C. Herein, the risk reduction with respect to the respective unhedged portfolios is approximately $5pp$ higher for ES 10%. For both risk measures, we find that the fixed hedge has the lowest mean HE approximately 5% and the second lowest for roughly 12% of the samples, while the fixed hedge achieves the highest risk reduction for more than 40% of the samples (see Figure 21 in Appendix D).

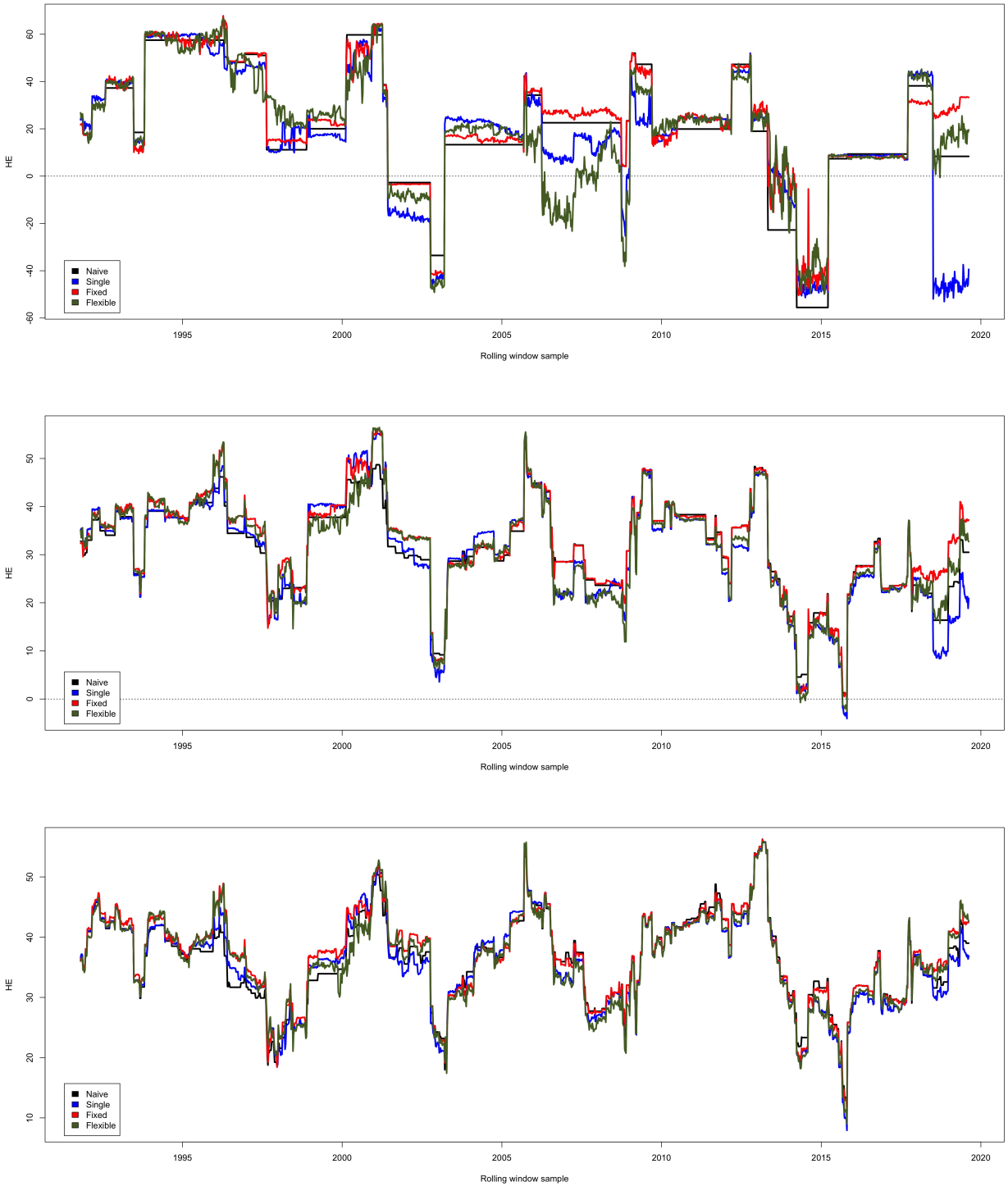

Figure 12: Out-of-sample HE in percentages over time for four hedging strategies: the naive, single, fixed and flexible hedge. The graph displays the reduction in Expected Shortfall (ES) at 99% (top), 95% (middle) and 90% (bottom) confidence levels for the test windows of 130 weeks starting from the rolling window sample date on the horizontal axis.

5.3.3 Lower Partial Moments

The LPM risk measures assess the variability of the negative profit margins. Herein, losses of the profit margin are taken to the power of 2 and 3 for LPM_2 and LPM_3, respectively. This means that LPM_3 is more sensitive to extreme losses than LPM_2. We examine that for both risk measures the fixed hedge has a significantly higher mean HE. We find that the fixed hedge has the lowest HE in 6% and 8% of the samples for LPM_2 and LPM_3, respectively.

The single and flexible hedge have significantly lower HE than the naive hedge. Visually, we can derive from Figure 13 that these hedging strategies' risk reductions lack on the back of the extreme losses when the hurricane Katrina observation enters the estimation window. For the flexible hedge minimising LPM_3, this causes gasoline hedge ratios to rise up to 1.5, whereas the crude and heating oil hedge ratios take even more extreme positions below 0.0 based on in-sample loss reduction. This leads to the flexible hedge to add more than 125*pp* LPM_3 risk with respect to the unhedged portfolio at its minimum around February 2007.

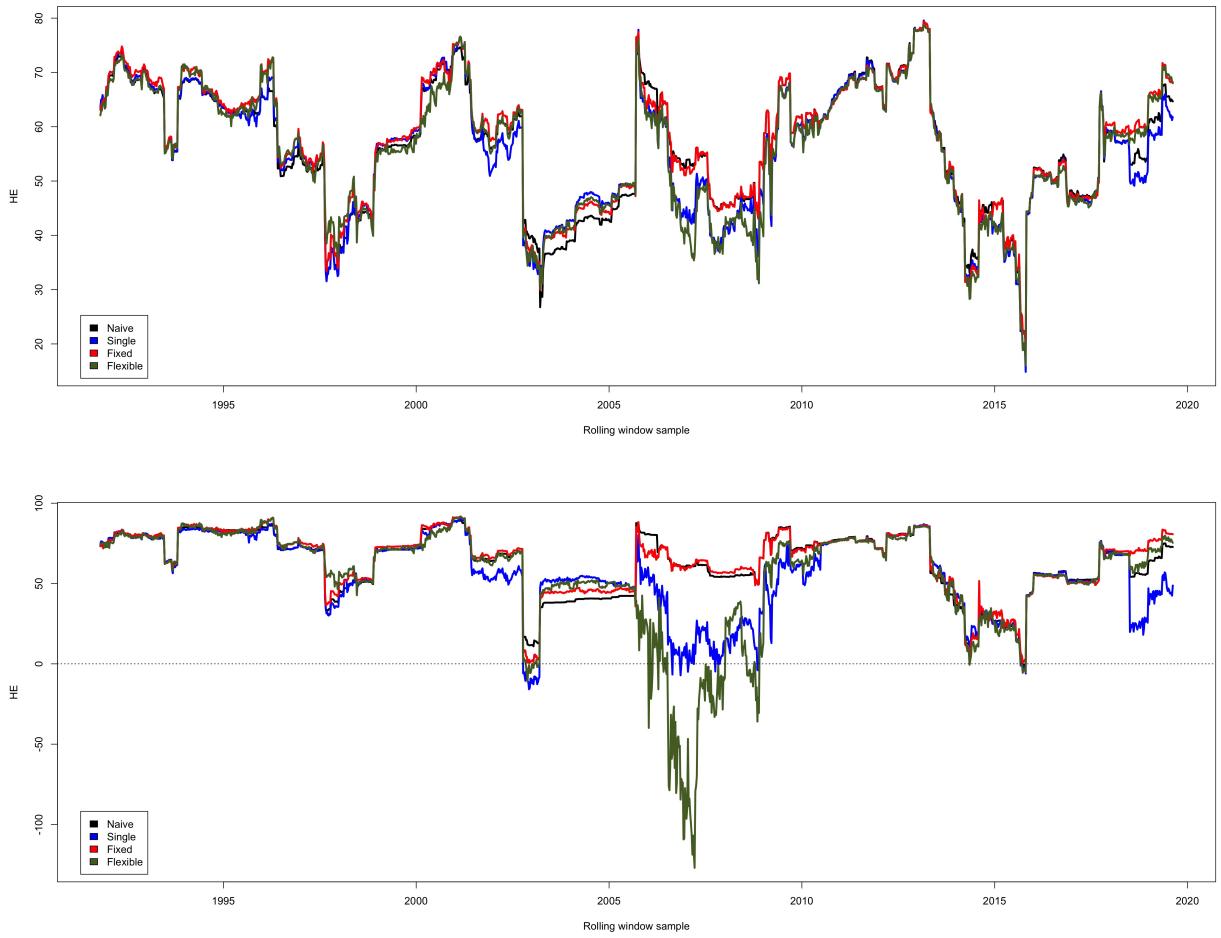

Figure 13: Out-of-sample HE in percentages over time for four hedging strategies: the naive, single, fixed and flexible hedge. The graph displays the reduction in Lower Partial Moments (LPM), LPM_2 (top) and LPM_3 (bottom), for the test windows of 130 weeks starting from the rolling window sample date on the horizontal axis.

5.4 Sensitivity analysis

In this section, we analyse the consequences of the decisions made in the process of conducting the research. Firstly, we focus on the impact of modelling with vine copula models instead of 3 paired copulae. Hereafter, in Section 5.4.2, we observe the impact of removing an extreme observation from our estimation window. Subsequently, we examine reducing the times of rebalancing the hedge ratios in Section 5.4.3. Moreover, in Section 5.4.4, we observe the effect of determining the hedge ratios on $10,000$ sampled draws instead of $100,000$. Lastly, we discuss the inclusion criterion in Section 5.4.5.

5.4.1 Impact of vine copulae and the fixed proportion constraint

Table 9 shows that the hedging strategies based on vine copulae, the fixed and flexible hedge, achieve more risk reduction than the strategy based on 3 paired copulae, the single hedge, for most risk measures. However, we can not directly gauge the impact of the vine copulae, as the fixed hedge is also affected by its $3:2:1$ fixed proportion constraint on the hedge ratios and the flexible hedge is derived over the mixture R-vine copula model. Hence, in this section, we compare the HE based on the R-vine copula as described in Section 4.3.2 against the HE based on 3 paired copulae as described in Section 4.3.1. We compare the performance of both of these copula models, with and without the fixed proportion constraint.

Table 11: Mean and median HE of four hedging strategies for eight risk measures. The hedging strategies are either based on 3 paired copulae or the R-vine copula and are either subject to the fixed proportion constraint (Fixed) or not (Flexible). Means are bold for the copula model with the highest HE for each risk measure within the Fixed and Flexible group.

	3 paired copulae				R-vine copula			
	Flexible		Fixed		Flexible		Fixed	
	Mean	*Median*	Mean	*Median*	Mean	*Median*	Mean	*Median*
VaR 1%	20.261	*20.185*	21.522	*20.869*	**21.803**	*20.450*	**22.718**	*22.567*
VaR 5%	39.057	*38.632*	39.943	*40.512*	**39.525**	*39.066*	**40.251**	*40.256*
VaR 10%	**43.363**	*43.666*	**43.768**	*45.265*	42.876	*43.222*	43.506	*44.308*
ES 1%	18.023	*20.354*	22.524	*22.116*	**19.257**	*19.921*	**23.590**	*24.246*
ES 5%	30.476	*31.827*	31.681	*32.557*	**30.933**	*32.620*	**32.432**	*33.566*
ES 10%	35.910	*36.380*	36.598	*37.215*	**35.995**	*36.739*	**36.984**	*37.661*
LPM_2	**55.454**	*56.766*	56.577	*57.697*	55.406	*56.776*	**57.054**	*58.819*
LPM_3	**56.324**	*57.085*	63.585	*69.671*	52.213	*64.350*	**64.549**	*70.460*

Table 11 presents three phenomena. Firstly, applying the fixed proportion constraint on the respective copula models increases the average HE significantly (see t-statistics in Table 22 in Appendix E.1). Secondly, hedging performance based the R-vine copula results in higher mean HE than for similar models based on 3 paired copulae for the vast majority of risk measures. Lastly, the hedging strategy based on 3 paired copulae without fixed proportion constraint improves more in HE by applying the fixed proportion constraint than by replacing the 3 paired copulae with the R-vine copula model. Herein, we also observe that the median HE is at least $0.6pp$ and at most $12.6pp$ higher for the models with fixed proportion constraint compared

to the same copula model without the proportion constraint. This suggests that the lower mean HE for models without the proportion constraint is not only due to extremely low HE results.

5.4.2 Sensitivity to the extreme observation hurricane Katrina

Sections 5.3.2 and 5.3.3 presented a low reduction of risk from the single- and flexible hedges for the ES 1%, LPM_2 and LPM_3 risk measures in the rolling window samples from September 2005 to June 2010. This period corresponds to the period in which the extreme spot price movement (-44$) after hurricane Katrina is present in the estimation window. The essence of risk management is to reduce the risk exposure for any situation by being prepared for extreme events. Hence, we do not eliminate such an observation from the estimation and test windows. However, in hindsight, we could deem the Katrina observation to be an unrealistic outlier with respect to the contemporary price movements. As such, given the notable hedging performance, we examine the effect of removing this observation from the estimation windows.

Table 12: Mean and median HE of the four hedging strategies for eight risk measures based on the estimation set without the Katrina observation.

	Naive		Single		Fixed		Flexible	
	Mean	*Median*	Mean	*Median*	Mean	*Median*	Mean	*Median*
VaR 1%	20.874	*19.910*	20.414	*20.225*	**22.753**	***22.515***	22.101	*20.875*
VaR 5%	39.199	*39.253*	39.073	*38.603*	**40.216**	***40.274***	39.579	*39.013*
VaR 10%	**43.651**	***44.632***	43.363	*43.689*	43.498	*44.396*	42.976	*43.079*
ES 1%	20.874	*19.910*	19.472	***22.974***	**23.306**	*21.998*	22.792	*22.071*
ES 5%	31.086	*31.685*	30.788	*31.839*	**32.414**	***33.560***	31.315	*33.011*
ES 10%	36.170	*36.770*	35.992	*36.457*	**36.991**	***37.611***	36.205	*36.853*
LPM_2	56.210	*56.780*	56.243	*57.134*	**57.109**	***58.725***	56.494	*57.790*
LPM_3	63.072	*68.693*	62.163	*66.357*	**64.523**	***71.048***	63.518	*68.868*

We analyse to what extent the occurrence of this extreme downward price movement of a spot price and smaller decline in futures price after hurricane Katrina drives the hedging performance of each hedging strategy. We compare the HE without the Katrina observation in our estimation Table 12 to the model that includes it for the estimation of hedge ratios Table 9. The hedging performance for the naive is equivalent for both estimation sets as this strategy is not affected by extreme price movements.

For the flexible hedge, we see that the mean risk reduction for ES 1%, LPM_2 and LPM_3 improved with approximately $3pp$, $1pp$ and $11pp$, respectively, by removing the Katrina observation from the estimation windows. For the single hedge, we observe the HE increase with $1pp$ and $6pp$ for ES 1% and LPM_3. Visually, these changes are evident for LPM_3, when comparing the period $2005 - 2010$ in Figure 14 against Figure 13 (bottom). We see a similar improvements for ES 1% and LPM_2 in Figure 20 Appendix E.2, while the changes in the remaining risk measures are smaller. Moreover, Table 23 in Appendix E.2 shows that for the single and flexible hedges the HE is significantly higher for all ES, LPM risk measures as well as the VaR 1% risk measure. By contrast, we observe that for the fixed hedge the estimation set including the Katrina

observation leads to significantly higher HE for ES 1% and significantly lower for ES 10% and LPM_2. This is extra motivation to use the fixed hedge when extreme price movements are observed.

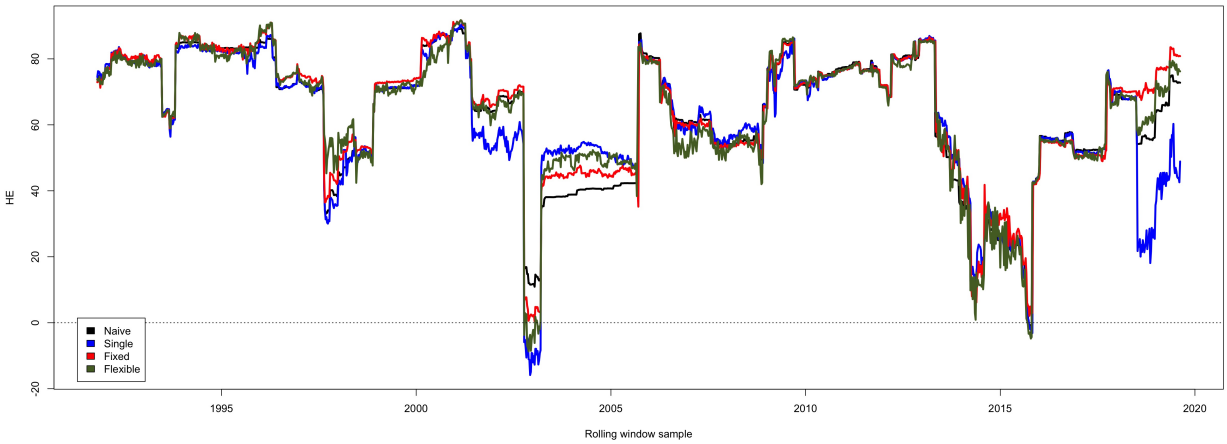

Figure 14: Out-of-sample HE in LPM_3 for the four hedging strategies, where the observation of hurricane Katrina is removed from the estimation windows.

In the ranking amongst hedge strategies, the removal of the Katrina observation results in the flexible hedge surpassing the naive hedge in overall hedging performance. Table 24 in Appendix E.2 shows that the flexible hedge attains significantly higher HE than the naive hedge for six risk measures with 99% confidence. This is a shift from Table 10, which presents that for the risk measures ES 1%, ES 10%, LPM_2 and LPM_3 the naive hedge reduced significantly more risk than the flexible hedge. Moreover, the flexible hedge outperforms the single hedge for seven of the risk measures with 99% confidence based on the estimation set without the Katrina observation. This suggests that the flexible hedge performs significantly worse when an extreme observation is included in the estimation set. Ultimately, the fixed proportion hedging strategy remains the hedging strategy with the highest average HE for all risk measures except VaR 10%.

5.4.3 Monthly rebalancing

In this section, we compare the performance of weekly rebalancing of hedge ratios against 4-weekly rebalancing. This idea is driven by the fact that with weekly rebalancing only one observation of 250 is replaced per rolling window sample. We examine how decreasing rate of adjusting hedge ratios to every four rolling window samples affects the hedging performance.

Table 13 provides that the fixed hedge strategy still achieves the highest mean HE of the hedge strategies that rebalance their hedge ratios, which does not include the naive hedge. Moreover, we note that the 4-weekly HE in Table 13 is similar to results based on weekly rebalancing in Table 9. We find three significant differences in HE in Table 25 in Appendix E.3. The single hedge rejects the null hypothesis of equal mean HE for weekly and 4-weekly rebalancing for ES 1%, preferring 4-weekly rebalancing with a 95% confidence

interval. The flexible hedge rejects the null hypothesis for LPM_3 with a 99% confidence interval and achieves a higher mean HE by rebalancing weekly. For the majority of the risk measures there is no significant difference in hedging performance between weekly and 4-weekly rebalancing. In practice, there is another argument to be made for 4-weekly rebalancing, namely the transaction costs related to adjusting hedge ratios.

Table 13: Mean and median HE of the four hedging strategies, which rebalance every four weeks, for eight risk measures based on the estimation set without the Katrina observation.

	Single		Fixed		Flexible	
	Mean	*Median*	Mean	*Median*	Mean	*Median*
VaR 1%	20.287	*20.063*	**22.755**	***22.364***	21.787	*20.433*
VaR 5%	39.093	*38.339*	**40.200**	***40.044***	39.516	*38.968*
VaR 10%	43.308	*43.500*	**43.496**	***44.571***	42.850	*42.985*
ES 1%	18.190	*20.464*	**23.512**	***24.154***	19.918	*20.920*
ES 5%	30.503	*31.713*	**32.424**	***33.620***	30.911	*32.938*
ES 10%	35.908	*36.312*	**36.984**	***37.672***	36.051	*36.823*
LPM_2	55.472	*56.713*	**57.055**	***58.842***	55.599	*57.154*
LPM_3	56.457	*58.231*	**64.598**	***70.584***	54.529	*65.343*

5.4.4 Downsizing vine copula draws

In this section, we examine the effect of downsizing the number of observations drawn during sampling in Section 4.4. The results presented thus far were based on $100,000$ draws from the vine copula model for each variable, which each presents a forecast. Our expectation is that in general a larger forecast sample helps to estimate hedge ratios more accurately, such that the determined hedge ratios pronounce the model better.

Table 14: Mean and median HE of the four hedging strategies for eight risk measures based on the hedge ratios optimised over $10,000$ in sample draws instead of $100,000$ as in Table 9.

	Single		Fixed		Flexible	
	Mean	*Median*	Mean	*Median*	Mean	*Median*
VaR 1%	20.211	*20.126*	**22.147**	***22.038***	20.461	*19.179*
VaR 5%	39.067	*38.668*	**40.076**	***40.047***	39.305	*38.507*
VaR 10%	43.212	*43.779*	**43.462**	***44.529***	42.781	*43.154*
ES 1%	18.071	*20.287*	**23.326**	***24.054***	19.459	*20.625*
ES 5%	30.497	*31.895*	**32.391**	***33.567***	30.897	*32.843*
ES 10%	35.910	*36.346*	**36.965**	***37.557***	36.029	*36.786*
LPM_2	55.473	*56.745*	**57.029**	***58.803***	55.549	*57.068*
LPM_3	56.429	*57.717*	**64.405**	***70.474***	54.310	*65.225*

Table 14 provides that the fixed hedge ranks the best amongst actively managed hedges in average HE in line with Table 9. When comparing the single hedge over $10,000$ draws to the hedges over $100,000$ draws, we find that the downsizing only impacts the hedge significantly for VaR 10% (see Table 15), for which $100,000$

draws are preferred. Moreover, when analysing the same comparison for the fixed hedge, we observe that the hedges over 100,000 draws result in significantly higher mean HE for every risk measure, except VaR 10%. Lastly, for the flexible hedge, we also prefer to determine the hedge ratios over 100,000 draws instead of 10,000 draws for VaR 1%, 5% and ES 1%, 5%, 10%, while for the remaining risk measures we do not observe a significant difference in HE. Concluding, we prefer to optimise hedge ratios over 100,000 draws for the three hedging strategies in Table 14, as either the mean HE is significantly higher than the hedging strategies based on 10,000 draws or we are not able to state that hedging performance differs significantly for each risk measure.

Table 15: T-statistics from two-tailed paired t-test, comparing the hedging performance of the single, fixed and flexible hedges based on 100,000 draws in Table 9 to the strategies based on 10,000 draws in Table 14. The positive (negative) t-statistic indicates a better performance for strategies based on 100,000 (10,000) draws. The values in brackets represent the p-values. The asterisks, *, ** and ***, represent the rejection of the equal HE at 10%, 5% and 1% significance, respectively.

	Single		Fixed		Flexible	
VaR 1%	0.481	(0.631)	5.383***	(0.000)	9.500***	(0.000)
VaR 5%	-0.187	(0.852)	2.851***	(0.004)	3.257***	(0.001)
VaR 10%	2.803***	(0.005)	0.852	(0.394)	1.349	(0.178)
ES 1%	-0.670	(0.503)	3.228***	(0.001)	2.867***	(0.004)
ES 5%	-1.404	(0.160)	2.938***	(0.003)	2.297**	(0.022)
ES 10%	0.069	(0.945)	2.525**	(0.012)	2.709***	(0.007)
LPM_2	-1.009	(0.313)	2.487**	(0.013)	1.150	(0.250)
LPM_3	-0.902	(0.367)	3.415***	(0.001)	-1.157	(0.248)

5.4.5 Sensitivity of the mixture vine copula with respect to inclusion criterion

In this section, we examine the impact of the inclusion criterion on in-sample fit and hedging performance. For the flexible hedge in Sections 5.2 and 5.3, we applied an AIC threshold of 99% of the copula type with lowest AIC as inclusion criterion. We find that this threshold increases the average HE within the flexible multi-commodity framework significantly for all risk measures except the VaR measures.

Table 16: Average log-likelihood, AIC, BIC and average number of parameters of the mixture R-vine copula models for AIC thresholds 95%, 98% and 99% and the R-vine model over the rolling window samples.

	Log-likelihood	AIC	BIC	# of parameters	# of multi-copula pairs
Mixt. R-vine 95%	1328.6	-2579.5	-2442.7	39	5.8
Mixt. R-vine 98%	1339.8	-2616.3	-2505.0	32	3.7
Mixt. R-vine 99%	1342.1	-2629.1	-2531.9	28	2.4
R-vine	1343.7	-2645.0	-2585.8	21	-

In Table 16, we observe the logical consequence of lowering the AIC threshold, namely a decreasing log-likelihood, AIC and BIC as well as an increase in the average number of copula types used to define the vine copula. We see that the mixture R-vine model with a 95% AIC threshold on average has $39 - 21 = 18$ more

copula parameters than the R-vine model. This increase in copula parameters is caused by the 5.8 pairs with multiple copula types assigned within the vine copula. This increase also directly decreases the AIC and BIC which penalise additional used parameters. In terms of log-likelihood, we find that the decrease of the 98% and 99% mixture R-vine score lower than the R-vine, but higher than the D-vine in Table 3.

Tables 17 and 18 show that an inclusion condition of 98% leads to either a significantly lower mean HE or a similar HE, for which the null hypothesis can not be rejected in both hedging frameworks with respect to the 99% R-vine mixture model (see Table 26 in Appendix E.4). We observe similar results for the R-vine model based on the 95% inclusion criterion, with the exception that this model results in a significantly higher HE for the VaR 5% risk measure.

Table 17: Average HE for the R-vine and mixture R-vine copula models in the flexible multi-commodity framework. 99%, 98% and 95% indicate the inclusion criterion based on AIC for the mixture R-vine models.

	R-vine		Mixt 99%		Mixt 98%		Mixt 95%	
	Mean	*Median*	Mean	*Median*	Mean	*Median*	Mean	*Median*
VaR 1%	21.803	*20.450*	**21.906**	***20.629***	21.221	*20.243*	19.748	*19.483*
VaR 5%	39.525	*39.066*	39.601	*39.101*	39.672	*39.152*	**39.841**	***39.677***
VaR 10%	42.876	***43.222***	**42.896**	*43.079*	42.853	*42.929*	42.942	*43.075*
ES 1%	19.257	*19.921*	**19.911**	***21.002***	18.926	*20.760*	17.580	*19.973*
ES 5%	30.933	*32.620*	**30.966**	***32.951***	30.829	*32.670*	30.905	*33.100*
ES 10%	35.995	*36.739*	**36.075**	*36.784*	36.015	***36.795***	36.053	*36.711*
LPM_2	55.406	*56.776*	**55.580**	*57.135*	55.488	***57.171***	55.476	*56.886*
LPM_3	52.213	*64.350*	**54.026**	***64.951***	53.203	*64.381*	52.454	*63.737*

Within the fixed proportion framework, we find that mean HE decreases strictly and significantly with each decrease in the inclusion criterion for the risk measures ES 1%, ES 5% and ES 10%, LPM_2 and LPM_3. Furthermore, for VaR 10%, we observe similar hedging performance from each (mixture) R-vine model. For VaR 1%, both the 99% mixture R-vine and R-vine model obtain significantly higher HE than the other mixture R-vine models.

Table 18: Average HE for the R-vine and mixture R-vine copula models in the fixed proportion framework. 99%, 98% and 95% indicate the inclusion criterion based on AIC for the mixture R-vine models.

	R-vine		Mixt 99%		Mixt 98%		Mixt 95%	
	Mean	*Median*	Mean	*Median*	Mean	*Median*	Mean	*Median*
VaR 1%	22.718	*22.567*	**22.734**	***22.631***	22.344	*22.246*	21.455	*20.119*
VaR 5%	40.251	*40.256*	40.186	*40.028*	40.185	*40.082*	**40.401**	***40.517***
VaR 10%	43.506	*44.308*	**43.534**	***44.437***	43.515	*44.348*	43.525	*44.262*
ES 1%	**23.590**	*24.246*	23.570	***24.652***	23.096	*23.654*	22.433	*20.832*
ES 5%	**32.432**	*33.566*	32.400	*33.670*	32.321	*33.545*	32.017	***33.682***
ES 10%	**36.984**	*37.661*	36.964	*37.660*	36.952	*37.723*	36.911	***37.771***
LPM_2	**57.054**	***58.819***	57.010	*58.798*	56.979	*58.789*	56.949	*58.446*
LPM_3	**64.549**	***70.460***	64.348	*70.365*	64.205	*70.350*	63.929	*69.396*

In conclusion, we find that the 99% mixture R-vine model achieves significantly higher HE than the other models within the flexible hedging framework (see Table 26). Moreover, in the fixed proportion hedging framework the R-vine model attains the highest HE for ES 5%, ES 10%, LPM_2 and LPM_3. Only for VaR 5% the 95% mixture outperforms the R-vine model. In general, we observe that the hedging performance decreases when the inclusion criterion is set to 98% or lower. This suggests that the setting the inclusion criterion to 99% is the sweet spot of the trade-off between adding additional copula types to define a variable pair and adding copula types with a strong fit.

6 Conclusion

The importance and relevance of hedging oil and oil related products for oil refineries have increased in recent years. In the last decade, with the rise of the vine copula, the modelling of such multi-product problems has improved. Sukcharoen & Leatham (2017) found that the D-vine performed best for hedging the downside risk of eight measures on a weekly basis within the flexible multi-commodity framework, while Liu et al. (2017) examined that for the daily hedging the flexible multi-commodity hedging framework outperformed the fixed proportion hedging framework based on a kernel copula model.

In our research, we evaluate three hedging strategies whether they can beat the naive hedge benchmark. These hedging strategies are based on three hedging frameworks: the single commodity, the fixed proportion and the flexible multi-commodity hedging framework. Within the latter two hedging frameworks, we examine whether the R-vine, which allows for additional vine copula structures besides the C- and D-vine, performed better than the C- and D-vine. Moreover, we propose a mixture R-vine copula model, which models bivariate dependence structures by multiple copula types under the condition that these copulae have an equivalent in-sample fit.

Our results show that the fixed hedge is the best hedging strategy, as it achieves the highest mean HE significantly for 7 of the 8 downside risk measures, beating the naive hedge. Behind the fixed hedge, the naive hedge achieves the highest average HE for the majority of the risk measures. Moreover, the single hedge scores the lowest risk reduction for 6 of the 8 risk measures, making it the least preferred hedging strategy for our problem.

In terms of vine copula models for the fixed proportion hedging framework, we find that the R-vine model is significantly preferred over the other vine copulae for four of the downside risk measures, while for the remaining four risk measures the hedging performance is indistinguishable for the D-, R- and mixture R-vine model. In the flexible multi-commodity hedging framework, we conclude that our proposed mixture R-vine model leads to significantly higher mean HE for all risk measures except the VaR risk measures.

Further analysis shows the benefit of modelling the dependency between all variables through vine copulae with respect to only modelling the spot-futures pair through three bivariate copulae. We observe this phenomenon in both the flexible hedging framework and the fixed proportion framework. Furthermore,

we find that the application of the fixed proportion constraint boosts the hedging performance more than drawing the samples from an R-vine instead of 3 paired copulae. Our study concludes that a single extreme lower tail observation, such as hurricane Katrina, deteriorates the hedging performance of the single and flexible hedge for the majority of the risk measures. Removing this observation from the estimation window increased the mean HE of the single and flexible hedge significantly. Moreover, we find that adjusting the rebalancing period from weekly to 4-weekly leads to similar results. By contrast, adjusting the draw sample from $100,000$ to $10,000$ draws for the fixed hedge leads to significantly lower mean HE for every risk measure except VaR 10%. Ultimately, we showed that lowering the AIC threshold below 99% generally does not improve hedging performance for most risk measures.

A limitation of our study are that we did not derive the drivers of the hedging performance of our proposed mixture R-vine model. As such, it would be interesting to study in further research why the mixture R-vine model increased performance in the flexible hedging framework but not in the fixed proportion framework. A related compelling starting point for future research would be the application and testing of mixture bivariate copulae in a risk management setting. Herein, we would also further explore more thoughtful methods to determine the inclusion criterion. Furthermore, we would like to investigate the hedging performance of our models in other multi-product settings and extend the oil refinery problem with additional petroleum products, such as kerosene and diesel which can be simultaneously produced from crude oil. Lastly, the hedging results of the model in the near future are interesting, given the unprecedented price movements in oil products since 2020.

Acknowledgements

I would like to express my sincere gratitude to each person that has helped me in this effort. Without their assistance this thesis would not have been written. My gratitude especially goes to Dr. C. Zhou for his extensive feedback and continuous guidance during this endeavour. I also extend my sincere gratitude to my parents and little sister: Roland Wessel, Monique van Ruitenburg and Luan Wessel, for their continuous support and patience during the pursuit of my academic career.

References

Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, *44*(2), 182–198.

Alexander, C., Prokopczuk, M., & Sumawong, A. (2013). The (de) merits of minimum-variance hedging: Application to the crack spread. *Energy Economics*, *36*, 698–707.

Awudu, I., Wilson, W., & Dahl, B. (2016). Hedging strategy for ethanol processing with copula distributions. *Energy Economics*, *57*, 59–65.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, *20*(4), 451–468.

Bedford, T., & Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, *32*(1), 245–268.

Bedford, T., & Cooke, R. M. (2002). Vines–a new graphical model for dependent random variables. *Annals of Statistics*, *30*(4), 1031–1068.

Brechmann, E. (2010). Truncated and simplified regular vines and their applications.

Dahlgran, R. A. (2009). Inventory and transformation hedging effectiveness in corn crushing. *Journal of Agricultural and Resource Economics*, 154–171.

Dissmann, J., Brechmann, E. C., Czado, C., & Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, *59*, 52–69.

Ederington, L. H. (1979). The hedging performance of the new futures markets. *The Journal of Finance*, *34*(1), 157–170.

Embrechts, P., Lindskog, F., & McNeil, A. (2001). Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*, *14*, 1–50.

Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, *82*(3), 543–552.

Haigh, M. S., & Holt, M. T. (2002). Crack spread hedging: accounting for time-varying volatility spillovers in the energy futures markets. *Journal of Applied Econometrics*, *17*(3), 269–289.

Hale, D. R., Lee, T., Zyren, J., Joosten, J., Kouserk, G., Rasmussen, J., & Hewlett, J. (2002). Derivatives and risk management in the petroleum, natural gas, and electricity industries. *Report, Energy Information Administration, US Department of Energy*.

Ji, Q., & Fan, Y. (2011). A dynamic hedging approach for refineries in multiproduct oil markets. *Energy*, *36*(2), 881–887.

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. *Lecture Notes-Monograph Series*, *28*, 120–141.

Kim, D., Kim, J.-M., Liao, S.-M., & Jung, Y.-S. (2013). Mixture of d-vine copulas for modeling dependence. *Computational Statistics & Data Analysis*, *64*, 1–19.

Kim, G., Silvapulle, M. J., & Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis*, *51*(6), 2836–2850.

Liu, P., Vedenov, D., & Power, G. J. (2017). Is hedging the crack spread no longer all it's cracked up to be? *Energy Economics*, *63*, 31–40.

Manner, H. (2007). Estimation and model selection of copulas with an application to exchange rates.

Mead, D., & Stiger, P. (2015). The 2014 plunge in import petroleum prices: What happened? *Beyond the Numbers: Global Economy*, *4*(9).

Morales-Napoles, O. (2010). Counting vines. In *Dependence modeling: Vine copula handbook* (pp. 189–218). World Scientific.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, *7*(4), 308–313.

Pindyck, R. S. (2001). The dynamics of commodity spot and futures markets: a primer. *The Energy Journal*, *22*(3).

Prim, R. C. (1957). Shortest connection networks and some generalizations. *The Bell System Technical Journal*, *36*(6), 1389–1401.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics*, *23*(3), 470–472.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de L'Institut de Statistique de L'Université de Paris*, *8*, 229–231.

Stoeber, J., Joe, H., & Czado, C. (2013). Simplified pair copula constructions—limitations and extensions. *Journal of Multivariate Analysis*, *119*, 101–118.

Sukcharoen, K., & Leatham, D. J. (2017). Hedging downside risk of oil refineries: A vine copula approach. *Energy Economics*, *66*, 493–507.

Tawn, J. A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika*, *75*(3), 397–415.

Wang, Y., Wu, C., & Yang, L. (2015). Hedging with futures: Does anything beat the naïve hedging strategy? *Management Science*, *61*(12), 2870–2889.

Weiß, G. N., & Scheffer, M. (2015). Mixture pair-copula-constructions. *Journal of Banking & Finance*, *54*, 175–191.

7 Appendix

A Copula types

We use the R package 'VineCopula' for the calculation of the copula types. Moreover, we visualise the copula types by simulation of observations in scatter plots in Figures 15 and 16. All copulae are defined for a bivariate instance, as we only apply bivariate copulae (pair-copulae) in our research. Hereafter, we list all copula types, the corresponding tail dependence and copula parameter bounds in Table 19 and 20. Ultimately, we provide the CDFs of the copula types

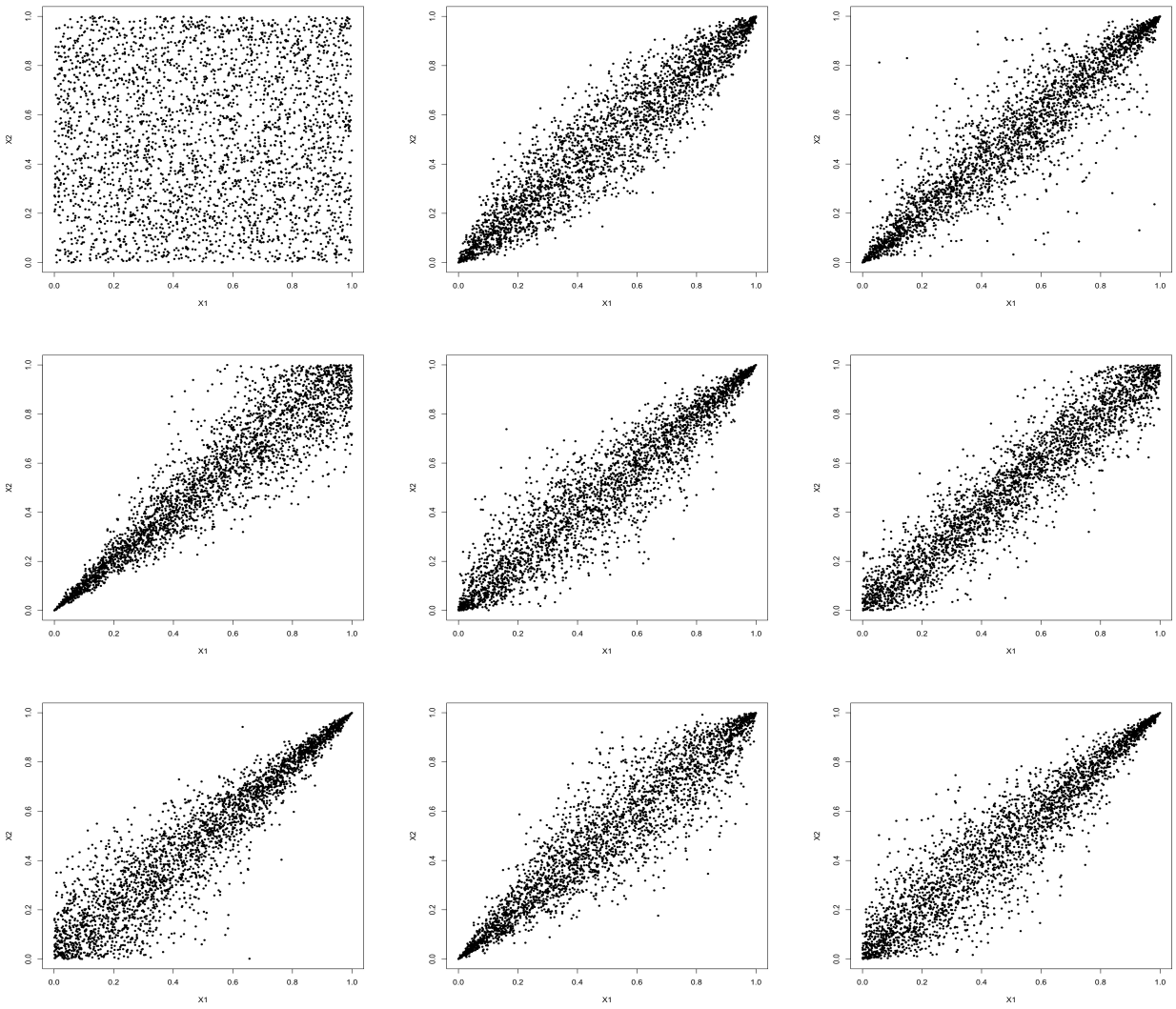

Figure 15: Scatter plots of 3,000 random observations on bivariate copulae. From left to right: Independence copula, Gaussian copula, Student-t copula, Clayton copula, Gumbel copula, Frank copula, Joe copula, $BB1$ copula and $BB6$ copula (all with a Kendall's $\hat{\tau}$ 0.8).

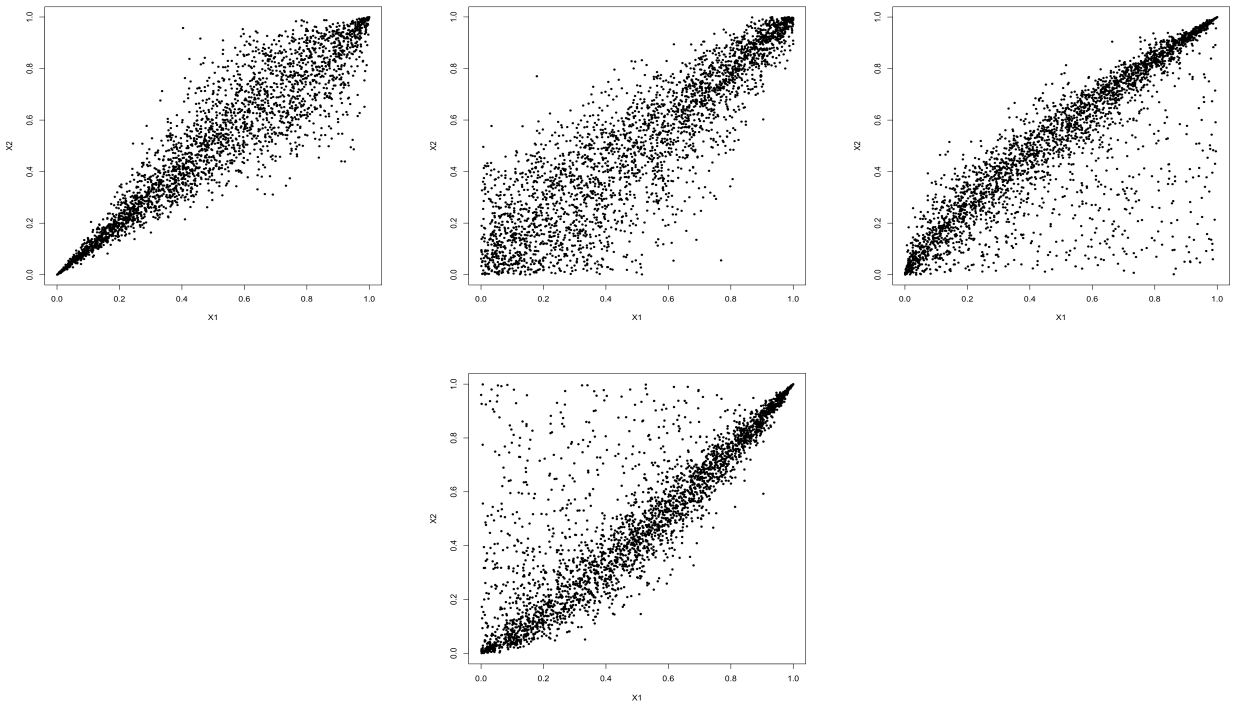

Figure 16: Scatter plots of 3,000 random observations on bivariate copulae. From left to right: $BB7$ copula with a Kendall's $\hat{\tau}$ 0.8, $BB8$ copula, Tawn type 1 copula and Tawn type 2 copula with $\hat{\tau}$ 0.7

Table 19: Part 1 of evaluated copula families, the corresponding tail dependence and copula parameter bounds.

	Copula type	Tail Dependence	LB_1	UB_1	LB_2	UB_2
1	Independence copula	None	-	-	-	-
2	Gaussian copula	None	-1	1	-	-
3	Student-t copula	Both	-1	1	2	Inf
4	Clayton copula	LTD	0	28	-	-
5	Gumbel copula	UTD	1	17	-	-
6	Frank copula	None	-35	35	-	-
7	Joe copula	UTD	1	30	-	-
8	BB1 copula	Both	0	7	1	7
9	BB6 copula	UTD	0	6	1	8
10	BB7 copula	Both	1	6	0	75
11	BB8 copula	UTD	1	8	0.0001	1
12	Rotated Clayton copula 180°	UTD	0	28	-	-
13	Rotated Gumbel copula 180°	LTD	1	17	-	-
14	Rotated Joe copula 180°	LTD	1	30	-	-
15	Rotated BB1 copula 180°	Both	0	7	1	7
16	Rotated BB6 copula 180°	LTD	0	6	1	8
17	Rotated BB7 copula 180°	Both	1	6	0	75
18	Rotated BB8 copula 180°	LTD	1	8	0.0001	1

Table 20: Part 2 of evaluated copula families, the corresponding tail dependence and copula parameter bounds.

	Copula type	Tail Dependence	LB_1	UB_1	LB_2	UB_2
19	Tawn type 1 copula	UTD	1	Inf	0	1
20	Rotated Tawn type 1 copula 180°	LTD	1	Inf	0	1
21	Tawn type 2 copula	UTD	1	Inf	0	1
22	Rotated Tawn type 2 copula 180°	LTD	1	Inf	0	1
23	Rotated Clayton copula 90°	None	-	-	-	-
24	Rotated Gumbel copula 90°	None	-	-	-	-
25	Rotated Joe copula 90°	None	-	-	-	-
26	Rotated BB1 copula 90°	None	-	-	-	-
27	Rotated BB6 copula 90°	None	-	-	-	-
28	Rotated BB7 copula 90°	None	-	-	-	-
29	Rotated BB8 copula 90°	None	-	-	-	-
30	Rotated Clayton copula 270°	None	-	-	-	-
31	Rotated Gumbel copula 270°	None	-	-	-	-
32	Rotated Joe copula 270°	None	-	-	-	-
33	Rotated BB1 copula 270°	None	-	-	-	-
34	Rotated BB6 copula 270°	None	-	-	-	-
35	Rotated BB7 copula 270°	None	-	-	-	-
36	Rotated BB8 copula 270°	None	-	-	-	-

A.1 Gaussian copula

The CDF of the Gaussian copula is defined as,

$$C(u_1, u_2; \theta) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\theta}} \exp\left\{ \frac{-s^2 - 2\theta st + t^2}{2(1-\theta^2)} \right\} \mathrm{d}s\mathrm{d}t,$$

where $\theta = \rho \in (-1, 1)$ and Φ^{-1} is the inverse of the normal distribution function.

A.2 Student-t copula

The CDF of the Student-t copula is defined as the inverse of,

$$T_v(u_1, u_2; \theta) = T_v(u_1, u_2; s) + \int_{s\frac{\pi}{2}}^{arcsin(\theta)} \left(1 + \frac{u_1^2 + u_2^2 - 2u_1u_2\sin(t)}{v\cos^2(t)} \right) \mathrm{d}t,$$

where $\theta \in (-1, 1)$, $s = sign(\theta)$ and $v > 0$ represents the degree of freedom. Furthermore, $T_v(u_1, u_2; s)$ is defined as,

$$T_v(u_1, u_2; s) = \begin{cases} T_v(\min\{u_1, u_2\}), & \text{if } s = 1 \\ \max\{0, T_v(b_1) - T_v(-b_2)\}, & \text{if } s = -1. \end{cases}$$

A.3 Clayton copula

The CDF of the Clayton copula is defined as,

$$C(u_1, u_2; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}},$$

if $u_1^{-\theta} + u_2^{-\theta} - 1 < largevalue$ to prevent overflow, otherwise $\min(u_1, u_2)$.

A.4 Gumbel copula

The CDF of the Gumbel copula is defined as,

$$C(u_1, u_2; \theta) = \exp\left\{-\left[(-\log\, u_1)^\theta + (\log\, u_2)^\theta\right]^{-\frac{1}{\theta}}\right\},$$

for $\delta > 1$.

A.5 Joe copula

The CDF of the Joe copula is defined as,

$$C(u_1, u_2; \theta) = 1 - \left(\bar{u}_1^\theta + \bar{u}_2^\theta - \bar{u}_1^\theta \bar{u}_2^\theta\right)^{\frac{1}{\theta}},$$

for $\delta \geq 1$ and $\bar{u} = 1 - \bar{u}$.

A.6 Frank copula

The CDF of the Frank copula is defined as,

$$C(u_1, u_2; \theta) = -\frac{1}{\delta}\log\left(1 - \frac{(1 - e^{-\theta u_1})((1 - e^{-\theta u_2})}{1 - e^{-\theta}}\right).$$

A.7 BB1, BB6, BB7, BB8 copulae

The BB1, BB6, BB7 and BB8 copulae are also known as the Clayton-Gumbel, Joe-Gumbel, Joe-Clayton and Joe-Frank copula, respectively. These copulae are instable for large parameters, therefore upper bounds are set at approximately 7. The CDF of the BB1 copula is defined as

$$C(u_1, u_2; \theta, \delta) = (1 + ((u_1^{-\theta} - 1)^\delta + ((u_2^{-\theta} - 1)^\delta)^{\frac{1}{\delta}})^{-\frac{1}{\theta}},$$

where it becomes a Clayton copula for $\delta = 1$ and a Gumbel copula for $\theta = 0$.

The CDF of the BB6 copula is defined as

$$C(u_1, u_2; \theta, \delta) = 1 - (1 - \exp(-[(-\log(1 - \bar{u}_1^\theta))^\delta + (-\log(1 - \bar{u}_2^\theta))^\delta]^{\frac{1}{\delta}}))^{\frac{1}{\theta}},$$

where $\theta, \delta \geq 1$ and $\bar{u} = 1 - \bar{u}$. The CDF of the BB7 copula is defined as

$$C(u_1, u_2; \lambda_L, \lambda_U) = 1 - (1 - [[1 - u_1)^{-\kappa}]^{-\gamma} + [1 - (1 - u_2)^\kappa]^{-\gamma} - 1]^{-\frac{1}{\gamma}})^{\frac{1}{\kappa}},$$

where γ is a function of the lower tail dependence coefficient λ_L and κ is a function of the upper tail dependence coefficient λ_U. The parameters are defined as follows $\gamma = -\frac{1}{log_2(\lambda_L)}$ and $\kappa = -\frac{1}{log_2(2-\lambda_U)}$. Both tail dependence coefficients are within the range $(0, 1)$.

The CDF of the BB8 copula is defined as

$$C(u_1, u_2; \theta, \delta) = \delta^{-1}\left(1 - [1 - (1 - (1 - \delta)^{\theta})^{-1}(1 - (1 - \delta u_1)^{\theta})(1 - (1 - \delta u_2)^{\theta}]^{\frac{1}{\theta}}\right),$$

where $\theta \geq 1$ and $\delta \in (-1, 1)$. The BB8 copula converges to the Frank copula for $\theta -> \infty$ and a Joe copula for $\delta = 1$.

A.8 Tawn type 1 & 2 copulae

The Tawn copulae are an extension of the Gumbel copula and known as extreme value copulae with 3 copula parameters (Tawn (1988)). However, our copulae are reduced to 2 parameter copulae.

$$\begin{aligned} C(u_1, u_2; \theta_1, \theta_2) &= C(u_1, u_2)^{A(w)}, \\ w &= \frac{\log(u_1)}{\log(u_1 u_2)}, \\ A(t) &= (1 - \theta_2)(1 - t) + (1 - \theta_1)t + [(\theta_1(1 - t))^{\delta} + (\theta_2 t)^{\delta}]^{\frac{1}{\delta}}, \end{aligned}$$

where $\theta_1, \theta_2 \in [0, 1]$ and $\delta \geq 1$. The Tawn type 1 (2) refers to the CDF where $\theta_1 = 1$ ($\theta_2 = 1$).

A.9 Rotated copulae

We apply the 90°, 180° and 270° rotated copula types of the earlier mentioned copula types f, as follows

$$C_{\text{rotated-180°},f}(u_1, u_2; \theta) = u_1 + u_2 - 1 + C_f(1 - u_1, 1 - u_2; \theta),$$

$$C_{\text{rotated-90°},f}(u_1, u_2; \theta) = u_2 - C_f(1 - u_1, u_2; -\theta),$$

$$C_{\text{rotated-270°},f}(u_1, u_2; \theta) = u_1 - C_f(u_1, 1 - u_2; -\theta).$$

B Minimum Spanning Tree algorithm

To find the tree with the maximum sum of dependence measured with Kendall's τ, Dissmann et al. (2013) propose the minimum spanning tree (MST) algorithm of Prim (1957). For each tree we assess the possible edges and estimate the corresponding empirical Kendall's $\hat{\tau}$ again as described in Section 4.3.2. To align our objective of finding the graph with the highest cumulative Kendall's τ with MST, we compute and assign weight $w_{i,j} = 1 - \hat{\tau}_{i,j}$ to edge $\{i, j\}$.

The MST algorithm for tree $T = (N, E)$, where N is the node set consisting of 6 nodes representing the 6 variables in T_1 and E represents the edge set follows:

- Initialise by randomly picking a node n_1 in node set N;
- Identify all neighboring nodes to n_1 and add edge connected to node n_2 with the smallest weight to current tree T;
- Repeat by identifying all neighbors of nodes already in T and add edge with the smallest weight connected to an isolated node. An isolated node is a node not yet connected with another edge;
- Iterate until all nodes are covered by T.

Once a lower tree T_r is specified in vine V, the algorithm for T_{r+1} becomes simpler, as the available nodes for T_{r+1} are the edges in T_r. This limits the number of possible trees T_{r+1}.

C Hedge ratios of hedging strategies

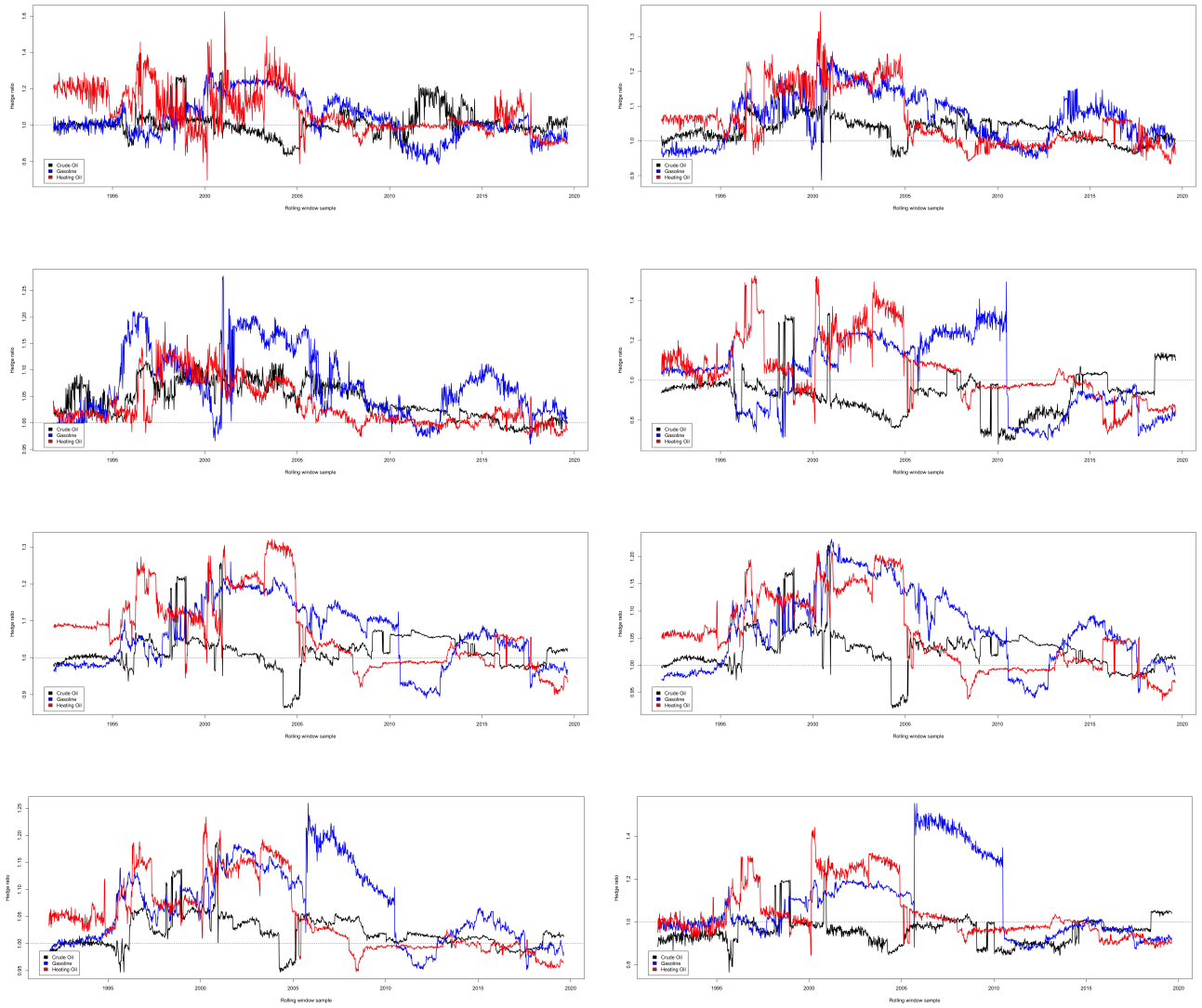

Figure 17: Hedge ratios optimised within the single commodity hedging framework for the downside risk measures: VaR 1%, VaR 5%, VaR 10%, ES 1%, ES 5%, ES 10%, LPM_2, LPM_3.

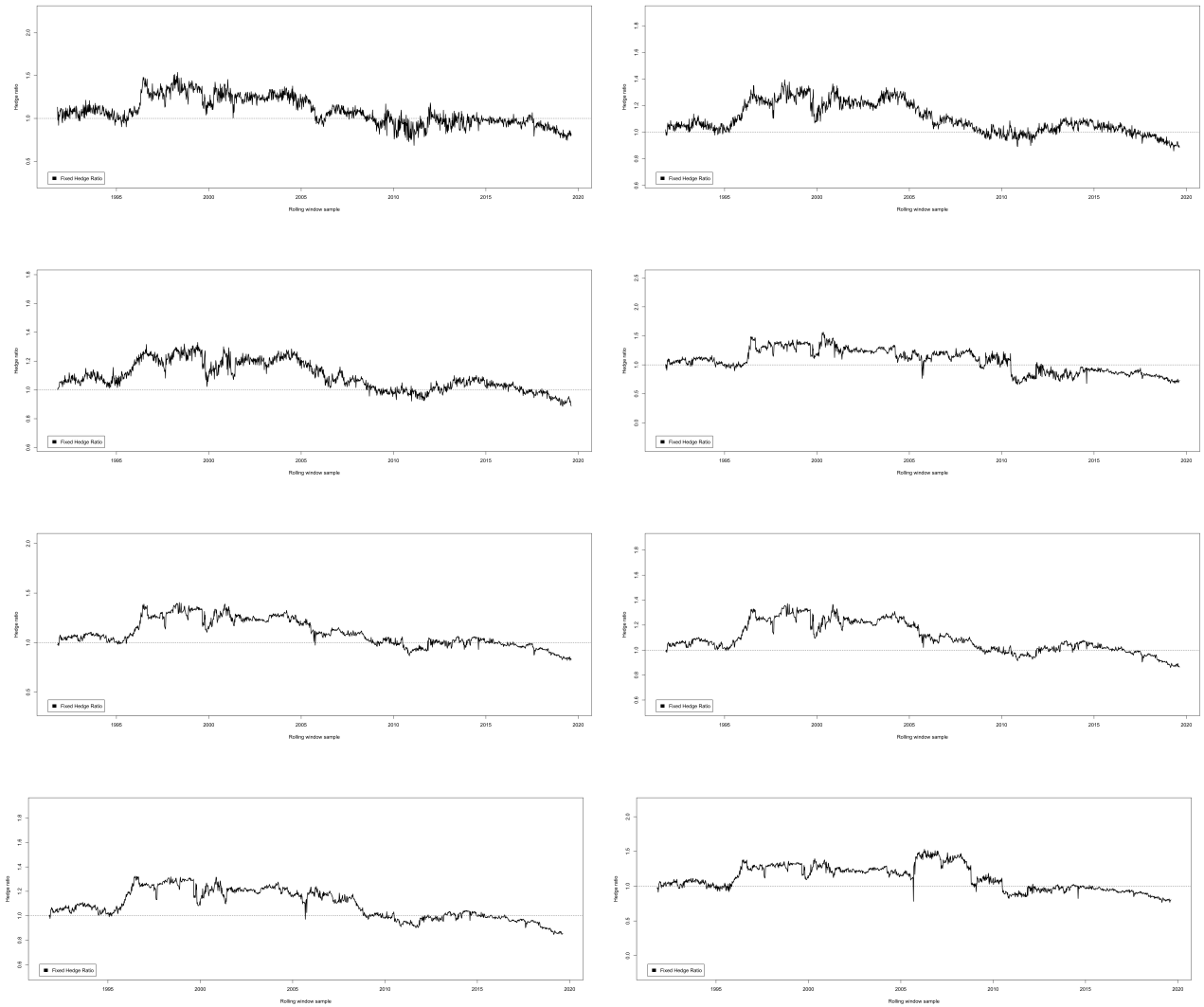

Figure 18: Hedge ratios optimised within the fixed proportion hedging framework based on the R-vine model for the downside risk measures: VaR 1%, VaR 5%, VaR 10%, ES 1%, ES 5%, ES 10%, LPM_2, LPM_3.

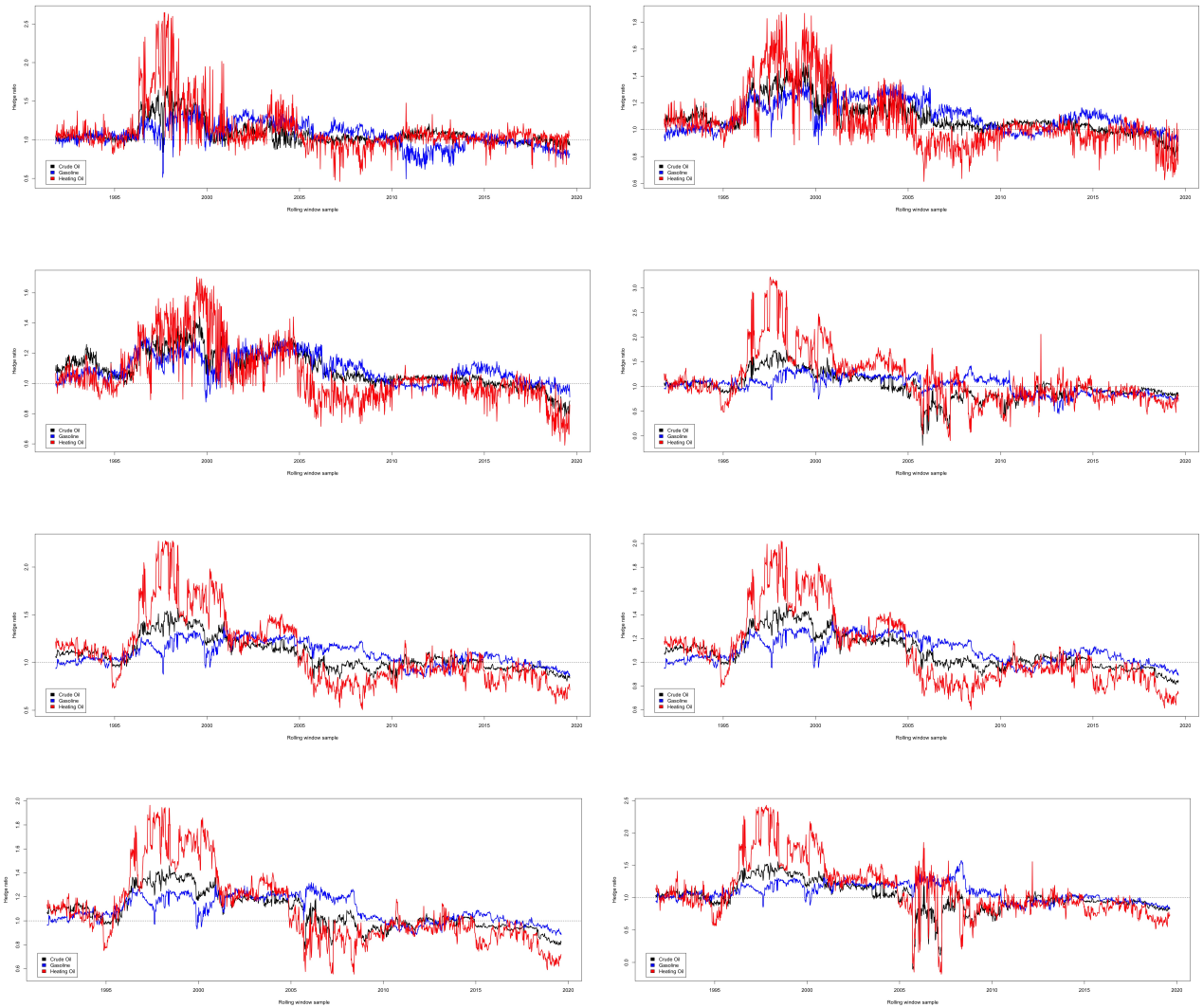

Figure 19: Hedge ratios optimised within the flexible multi-commodity hedging framework based on the mixture R-vine model for the downside risk measures: VaR 1%, VaR 5%, VaR 10%, ES 1%, ES 5%, ES 10%, LPM_2, LPM_3

D Hedge effectiveness statistics

Table 21: Additional HE statistics for Table 9. Left hand side shows the number of times (in %) a hedge strategy is in rank 4 (lowest HE) to 1 (highest HE) over the whole dataset for each risk measure. Right hand side presents the average HE difference to the strategy with the highest HE for the samples with the corresponding rank.

VaR 1%	Naive	Single	Fixed	Flexible	VaR 1%	Naive	Single	Fixed	Flexible
4	29	34	18	19	4	5.467	9.490	4.656	5.896
3	32	20	22	26	3	4.875	4.404	3.424	4.213
2	24	21	23	32	2	3.142	1.735	1.902	2.434
1	15	24	37	23	1	-	-	-	-
VaR 5%	Naive	Single	Fixed	Flexible	VaR 5%	Naive	Single	Fixed	Flexible
4	34	28	14	24	4	5.337	5.970	4.209	4.741
3	22	31	18	29	3	3.618	3.323	2.971	3.246
2	24	21	31	25	2	1.146	1.559	2.228	1.924
1	20	20	37	23	1	-	-	-	-
VaR 10%	Naive	Single	Fixed	Flexible	VaR 10%	Naive	Single	Fixed	Flexible
4	19	25	19	36	4	4.429	4.156	4.254	4.760
3	21	27	27	25	3	2.837	2.624	2.757	2.729
2	29	27	28	15	2	1.258	1.348	1.368	1.559
1	31	21	25	24	1	-	-	-	-
ES 1%	Naive	Single	Fixed	Flexible	ES 1%	Naive	Single	Fixed	Flexible
4	36	25	12	27	4	9.507	22.070	6.327	16.292
3	21	30	27	22	3	8.404	8.624	4.674	5.221
2	20	22	31	27	2	3.191	2.094	3.277	3.649
1	24	23	30	23	1	-	-	-	-
ES 5%	Naive	Single	Fixed	Flexible	ES 5%	Naive	Single	Fixed	Flexible
4	27	39	05	29	4	3.494	4.142	1.995	3.280
3	24	35	12	29	3	3.542	2.226	1.410	1.942
2	22	10	42	26	2	0.478	0.940	0.639	1.614
1	28	16	41	15	1	-	-	-	-
ES 10%	Naive	Single	Fixed	Flexible	ES 10%	Naive	Single	Fixed	Flexible
4	26	28	06	40	4	2.681	2.626	1.255	1.983
3	25	37	12	26	3	1.778	1.529	0.947	1.326
2	18	21	40	22	2	0.311	0.716	0.527	0.720
1	31	13	43	13	1	-	-	-	-
LPM_2	Naive	Single	Fixed	Flexible	LPM_2	Naive	Single	Fixed	Flexible
4	27	27	06	40	4	2.778	3.921	0.811	3.283
3	26	35	15	25	3	2.016	2.550	1.292	1.728
2	17	23	36	24	2	0.389	0.574	0.687	0.777
1	31	15	43	11	1				
LPM_3	Naive	Single	Fixed	Flexible	LPM_3	Naive	Single	Fixed	Flexible
4	23	35	08	34	4	6.535	14.447	1.983	28.708
3	27	28	19	25	3	4.695	16.764	4.023	5.885
2	22	18	30	29	2	1.805	1.048	2.470	2.473
1	28	19	42	11	1	-	-	-	-

E Additional information sensitivity analysis

E.1 Impact of vine copulae and the fixed proportion constraint

Table 22: T-statistics from two-tailed paired t-test, comparing the HE based on 3 paired copulae without and with fixed proportion constraint (3-Flex and 3-Fixed, respectively) and the HE based on R-vine copula model without and with fixed proportion constraint (R-Flex and R-Fixed, respectively). The positive (negative) t-statistic indicates a better performance for model left (top) of the statistic. The asterisks, *, ** and ***, represent the rejection of the equal hedging effectiveness at 10%, 5% and 1% significance, respectively.

VaR 1%	3-Fixed	R-Flex	R-Fixed
3-Flex	-9.298***	-10.860***	-12.229***
3-Fixed	-	-2.145**	-9.601***
R-Flex	-	-	-6.178***
VaR 5%	3-Fixed	R-Flex	R-Fixed
3-Flex	-10.385***	-4.866***	-12.802***
3-Fixed	-	4.092***	-3.278***
R-Flex	-	-	-9.335***
VaR 10%	3-Fixed	R-Flex	R-Fixed
3-Flex	-4.908***	5.145***	-1.632
3-Fixed	-	8.183***	3.210***
R-Flex	-	-	-7.216***
ES 1%	3-Fixed	R-Flex	R-Fixed
3-Flex	-12.511***	-2.986***	-13.042***
3-Fixed	-	9.012***	-8.376***
R-Flex	-	-	-12.968***

ES 5%	3-Fixed	R-Flex	R-Fixed
3-Flex	-20.047***	-5.506***	-20.554***
3-Fixed	-	10.385***	-13.274***
R-Flex	-	-	-24.547***
ES 10%	3-Fixed	R-Flex	R-Fixed
3-Flex	-22.831***	-1.643	-26.602***
3-Fixed	-	11.987***	-11.602***
R-Flex	-	-	-25.943***
LPM_2	3-Fixed	R-Flex	R-Fixed
3-Flex	-18.892***	0.698	-24.589***
3-Fixed	-	13.209***	-14.157***
R-Flex	-	-	-21.759***
LPM_3	3-Fixed	R-Flex	R-Fixed
3-Flex	-17.480***	6.197***	-19.692***
3-Fixed	-	12.214***	-12.970***
R-Flex	-	-	-13.509***

E.2 Additional information excluding the hurricane Katrina observation in estimation

Figure 20: Out-of-sample hedging effectiveness (HE) in percentages for 4 hedging strategies where the observation after hurricane Katrina is removed from the estimation windows. The graph displays the reduction in ES 1% (top) and LPM_2 (bottom).

Table 23: T-statistics from two-tailed paired t-test, comparing performance of Table 9 including hurrican Katrina in estimation to the performance in Table 12 of the single, fixed and flexible hedge. The positive (negative) t-statistic indicates a better performance for estimation excluding (including) the observation of hurricane Katrina. The values in brackets represent the p-values. The asterisks, *, ** and ***, represent the rejection of the equal hedging effectiveness at 10%, 5% and 1% significance, respectively.

	Single		Fixed		Flexible	
VaR 1%	2.211**	(0.027)	0.555	(0.579)	6.018***	(0.000)
VaR 5%	0.829	(0.407)	-0.837	(0.403)	-1.048	(0.295)
VaR 10%	0.286	(0.775)	-0.035	(0.972)	2.390**	(0.017)
ES 1%	12.986***	(0.000)	-3.445***	(0.001)	14.007***	(0.000)
ES 5%	14.002***	(0.000)	-1.327	(0.185)	13.328***	(0.000)
ES 10%	10.676***	(0.000)	1.993**	(0.046)	11.401***	(0.000)
LPM_2	15.187***	(0.000)	4.456***	(0.000)	14.348***	(0.000)
LPM_3	15.658***	(0.000)	-0.262	(0.793)	13.036***	(0.000)

Table 24: T-statistics from two-tailed paired t-test, comparing 4 hedging strategies in pairs over the dataset without the Katrina observation in the estimation windows, which consist of the strategies to the left and directly above the statistic. The positive (negative) t-statistic indicates a better performance for the model left (top) of the statistic. The asterisks, *, ** and ***, represent the rejection of the equal hedging effectiveness at 1%, 5% and 10% significance, respectively.

VaR 1%	Single	Fixed	Flexible
Naive	2.825***	-13.127***	-10.032***
Single	-	-11.648***	-10.444***
Fixed	-	-	3.912***
VaR 5%	Single	Fixed	Flexible
Naive	0.996	-8.737***	-3.178***
Single	-	-12.156***	-4.768***
Fixed	-	-	7.554***
VaR 10%	Single	Fixed	Flexible
Naive	3.235***	1.886*	7.000***
Single	-	-1.556	4.086***
Fixed	-	-	5.754***
ES 1%	Single	Fixed	Flexible
Naive	3.967***	-12.666***	-8.973***
Single	-	-9.308***	-9.335***
Fixed	-	-	2.848***

ES 5%	Single	Fixed	Flexible
Naive	3.541***	-18.232***	-3.508***
Single	-	-16.916***	-5.783***
Fixed	-	-	16.752***
ES 10%	Single	Fixed	Flexible
Naive	3.553***	-16.212***	-0.776
Single	-	-24.572***	-3.799***
Fixed	-	-	18.403***
LPM2	Single	Fixed	Flexible
Naive	-0.699	-16.959***	-5.089***
Single	-	-15.61***	-3.558***
Fixed	-	-	11.357***
LPM3	Single	Fixed	Flexible
Naive	4.261***	-13.956***	-3.566***
Single	-	-10.163***	-6.367***
Fixed	-	-	8.771***

E.3 Monthly rebalancing

Table 25: T-statistics from two-tailed paired t-test, comparing the weekly rebalancing performance of Table 9 to the 4-weekly rebalancing performance in Table 13 of the single, fixed and flexible hedge. The positive (negative) t-statistic indicates a better performance for weekly (monthly) rebalancing. The values in brackets represent the p-values. The asterisks, *, ** and ***, represent the rejection of the equal hedging effectiveness at 10%, 5% and 1% significance, respectively.

	Single		Fixed		Flexible	
VaR 1%	-0.367	(0.713)	-0.582	(0.561)	-1.037	(0.300)
VaR 5%	-0.805	(0.421)	1.035	(0.301)	-1.269	(0.205)
VaR 10%	1.197	(0.231)	0.278	(0.781)	-0.721	(0.471)
ES 1%	-2.004**	(0.045)	1.323	(0.186)	0.063	(0.950)
ES 5%	-1.629	(0.103)	0.612	(0.541)	-1.916*	(0.056)
ES 10%	0.230	(0.818)	0.030	(0.976)	-1.395	(0.163)
LPM_2	-1.250	(0.212)	-0.182	(0.856)	0.854	(0.393)
LPM_3	-1.608	(0.108)	-1.579	(0.115)	2.654***	(0.008)

E.4 Inclusion criterion

Table 26: T-statistics from two-tailed paired t-test, where the positive (negative) t-statistic indicates a better performance for the model left (top) of the statistic. Left hand side provides the t-statistic based on the HE in flexible multi-commodity hedging framework, whereas the right hand side shows the HE in the fixed proportion hedging framework. The asterisks, *, ** and ***, represent the rejection of the equal HE at 10%, 5% and 1% significance, respectively.

VaR 1%	98%	99%	R-vine	98%	99%	R-vine
95%	-9.812***	-13.32***	-12.157***	-9.086***	-11.476***	-11.484***
98%	-	-5.732***	-4.410***	-	-4.961***	-4.434***
99%	-	-	0.920	-	-	0.202
VaR 5%	98%	99%	R-vine	98%	99%	R-vine
95%	2.213**	3.082***	3.593***	4.190***	4.107***	2.626***
98%	-	1.001	2.008**	-	-0.020	-1.235
99%	-	-	1.107	-	-	-1.390
VaR 10%	98%	99%	R-vine	98%	99%	R-vine
95%	1.265	0.611	0.866	0.240	-0.220	0.438
98%	-	-0.633	-0.352	-	-0.492	0.244
99%	-	-	0.281	-	-	0.693
ES 1%	98%	99%	R-vine	98%	99%	R-vine
95%	-8.281***	-12.897***	-8.041***	-6.074***	-9.815***	-9.526***
98%	-	-7.449***	-1.979**	-	-6.457***	-5.638***
99%	-	-	4.851***	-	-	-0.310
ES 5%	98%	99%	R-vine	98%	99%	R-vine
95%	1.689*	-1.195	-0.499	-9.629***	-11.008***	-11.651***
98%	-	-4.363***	-2.761***	-	-4.850***	-6.089***
99%	-	-	1.166	-	-	-2.402**
ES 10%	98%	99%	R-vine	98%	99%	R-vine
95%	1.479	-0.809	1.706*	-3.429***	-4.286***	-5.764***
98%	-	-3.133***	0.847	-	-1.493	-3.681***
99%	-	-	4.173***	-	-	-2.691***
LPM_2	98%	99%	R-vine	98%	99%	R-vine
95%	-0.416	-3.147***	1.743*	-2.193**	-4.534***	-7.538***
98%	-	-3.825***	2.633***	-	-3.228***	-6.733***
99%	-	-	6.262***	-	-	-4.600***
LPM_3	98%	99%	R-vine	98%	99%	R-vine
95%	-5.122***	-7.884***	0.803	-5.592***	-8.119***	-11.769***
98%	-	-4.089***	3.463***	-	-4.360***	-8.793***
99%	-	-	5.922***	-	-	-5.902***