Multi-Commodity Price Risk Hedging using Asymmetric Tail Dependence Modelling

Master Thesis Quantitative Finance

Sebastiaan Thomas Njio - 534274

July 14, 2021

Abstract

In this paper I investigate an application of Hierarchical Outer Power transformed Archimedean Copulas (HOPACs) recently proposed by Górecki et al. (2021). I use HOPACs for modelling the price-risk exposure of an oil refinery. I use Monte-Carlo simulations from said HOPACs to determine optimal hedging ratios for reducing the Value-at-Risk (VaR) and Expected-Shortfall (ES). This research extends the methods from Alexander et al. (2013), by targeting VaR and ES rather than variance, and extends the research by considering the HOPAC models. I find that some HOPAC implementations lead to a substantial reduction in ES. I also find that disregarding volatility clustering in favor of model parsimony, leads to increased performance in this context. However, a naive method reduces moderate losses more effectively than the HOPAC counterparts. A refinery concerned with reducing moderate losses should therefore consider targeting a different risk measure instead.

Supervised by

dr. J.W.N. Reuvers

Second assessor

dr. O. Kleen

Erasmus University Rotterdam, Erasmus School of Economics

Contents

1 Introduction

Energy from oil commodities enable essential services, from providing heat to providing energy for manufacturing and transportation. Downey (2009) states that refineries face high capital investments, and until recently, low profit margins. A refinery may increase their financial attractiveness to investors by mitigating the risk of substantial losses. Refinery profit margins are closely related to the *crack spread*, which is the price difference between crude oil and refined commodities (EIA, 2002). Haigh and Holt (2002) propose a two period framework to determine an appropriate hedging strategy. The hedging strategy mainly concerns the appropriate amount of hedging per commodity, expressed in hedging ratios. Ji and Fan (2011) use GARCH models to determine time varying minimum-variance hedging ratios, and find that the GARCH models significantly outperform a naive implementation. However, Alexander et al. (2013) find that GARCH models incur high hedging costs due to considerable parameter uncertainty. Moreover, Alexander et al. (2013) find that the hedging costs cause a naive strategy to outperform all GARCH implementations. Next, Sukcharoen and Leatham (2017) argue that refineries should only mitigate the risk of adverse price movements, rather than reduce variance, they consider the SV, LPM, VaR, and ES risk measures instead. Furthermore, Sukcharoen and Leatham (2017) use a Gaussian copula, Student's t copula, Clayton copula, C- and D-vine copulas to model the complete joint distribution of commodity spots, and futures prices. The dependence between spots and futures prices is important due to the risk eminating from the mismatch between commodity delivery and the futures expiry. Sukcharoen and Leatham (2017) find that the D-vine copula enables the highest risk reduction for each respective measure. Sukcharoen and Leatham (2017) state that the D-vine copula performance is likely linked to its versatility in modelling the asymetric (tail) dependence between the respective financial instruments. However, despite the findings of Alexander et al. (2013), Sukcharoen and Leatham (2017) do not consider trading costs. Furthermore, Sukcharoen and Leatham (2017) only compare the D-vine copula against alternatives that lack the ability to capture asymetric (tail) dependence structures.

I investigate an application of Hierarchical Outer Power transformed Archimedean Copulas (HOPACs) recently proposed by Górecki et al. (2021), and use them for modelling the

price-risk exposure of an oil refinery. Górecki et al. (2021) introduce HOPACs as a combination between Archimedean Copulas (ACs) with an outer power (OP) transformation, and Hierarchical ACs (HACs). The OP transformation allows for a better fit in both body- and tail-dependence compared to a regular HAC (Górecki et al., 2021). Like HACs, HOPACs also allow for modelling asymetric dependence between marginals. The extra flexibility in both body- and tail-dependence enables HOPACs to compete with the D-vine copulas. Furthermore, HOPACs benefit from a more intuitive and parsimonious dependence structure compared to vine copulas (Okhrin et al., 2013). Model parsimony might offer a distinct advantage with respect to parameter uncertainty in the volatile oil markets (Alexander et al., 2013). To my knowledge, no article yet exists that applies HOPACs to data outside of a simulation. Furthermore, I also consider hedging costs since Alexander et al. (2013) find them to be a practical limitation for most proposed models. Besides, I investigate the parsimony trade-off in the context of modelling volatility clustering, by using a GARCH$(1,1)$ model to filter each marginal.

I investigate HOPAC effectiveness in the oil refinery hedging framework from Haigh and Holt (2002), using historical data from Jun 1986 through Mar 2021. I estimate *two* HOPAC implementations from *four* families on historical observations from a *five*-year backward-looking moving window. The *two* HOPAC implementations are; with accounting for volatility clustering (HOPAC-GARCH), and without (HOPAC-i.i.d.). Disregarding the presence of volatility clustering leads to a more parsimonious model, which might increase hedging performance due to lower estimation uncertainty. I account for volatility clustering by using a GARCH$(1,1)$ model to filter observations from each marginal. I use *four* generating functions in each implementation; Ali–Mikhail–Haq, Clayton, Frank, and Joe. Next, I use Monte-Carlo simulations from said HOPACs to obtain hypothetical price changes for the estimation of optimal hedging ratios based on two risk measures; VaR, and ES. Next, I evaluate the efficacy of this hedging strategy on an ex post basis, and test for first order stochastic dominance using a one-sided Kolmogorov-Smirnov test.

The primary *crack spread* commodities are; crude oil, gasoline, and heating oil. Specifically, I consider WTI crude at Cushing Oklahoma, conventional gasoline at New York Harbor, and No.2 heating-oil at New York Harbor. For each commodity, I obtain weekly Wednesdays spots

and futures prices from the Energy Information Administration (EIA). The resulting sample spans from Jun 10, 1986 through Mar 10, 2021.

The HOPAC-i.i.d. implementations lead to a substantial reduction of the ES. However, regardless of the time dependence assumption, HOPAC implementations do not substantially reduce the VaR. A refinery concerned with ES can therefore definitely reduce its risk using HOPACs to estimate optimal hedging ratios. A reduction of risk might lower the capital requirement and increase the attractiveness to investors. However, HOPACs perform worse when considering typical losses when targeting the VaR- and ES-metric. Targeting a different risk-metric which measures the severity of typical losses might enable HOPACs to outperform instead.

Accounting for volatility-clustering adds 18 parameters to the existing 10, thus substantially increasing the amount of estimates. The extra parameters harm the model's performance in practice. Therefore, it is desirable to disregard volatility-clustering in favor of model parsimony in this particular setting, over this risk management horizon.

2 Data

2.1 Data Description

I obtain spots and futures prices for crude, gasoline, and heating-oil from the EIA (Energy Information Administration, 2021b, and Energy Information Administration, 2021a). I consider WTI crude at Cushing Oklahoma, conventional and regular RBOB gasoline at New York Harbor, and Number 2 heating oil at New York Harbor. I adjust the daily spot and future series to weekly Wednesdays closing prices. In the occasion that trading is closed on Wednesday, Tuesday or Monday prices are taken instead. Conventional gasoline futures are discontinued in Dec 2006, therefore I transition linearly into RBOB futures starting from Oct 2005. The linear transition prevents any artificial jumps caused by rolling over. Gasoline and heating-oil futures expire on the last business day of the calendar month, succeeding the trade date. The expiry of crude futures deviates slightly from this convention, which I disregard in my analysis. Hurricane Katrina caused many US oil refineries to suspend operations, thus leaving no price

risk to hedge. Therefore, I discard observations on 31 Aug 2005, and 7 Sept 2005. The resulting data span from Jun 4, 1986 until Mar 10, 2021, excluding 31 Aug 2005, and 7 Sept, leading to 1802 observations for each series.

2.2 Data Adjustments

For (tail) risk management purposes, I make all profits negative, and all losses positive. Next, I construct profit and loss (P&L) series for spots by taking the first order backward difference. The P&L spots series are denoted by ΔS_C, ΔS_G, and ΔS_{HO} for crude, gasoline, and heating-oil respectively.

Alexander et al. (2013) discuss two methods for creating P&L futures series. First being the standard *rollover* method, which uses fixed maturity future contracts. Second being the *constant-maturity* method, where we construct a dynamic portfolio consisting of two futures contracts with different maturities. Therefore, the *constant-maturity* method results in futures with a longer time-to-maturity compared to the *rollover* method. Alexander et al. (2013) argue in favor of the *constant-maturity* method, since the *rollover* method lead to jumps caused by the rolling over between contracts. Alexander et al. (2013) also state that the rollover jumps may harm the performance of certain models. I apply both methods, and find that the *rollover* method leads to slightly higher risk reductions. The higher risk reduction is likely linked to the stronger relation between the spots and corresponding futures, as a result of the shorter time-to-maturity. I continue with the *rollover* method, and provide results from the constant maturity method in appendix A. The P&L future series constructed with the first method, are denoted by ΔF_C, ΔF_G, and ΔF_{HO} for crude, gasoline, and heating-oil respectively. Both spot and future P&L series span from Jun 10, 1986 until Mar 10, 2021, leading to 1801 observations.

Next, Table 1 shows summary statistics of the P&L series over the entire sample using conventional moment estimators. Here we see that the mean, denoted by $\hat{\mu}$, of each series is slightly negative, indicating that the commodity prices tend to rise. The weekly standard-deviation and kurtosis, denoted by $\hat{\sigma}$ and $\hat{\kappa}$ respectively, decrease in the order of gasoline, heating-oil, and crude. This means that gasoline facing market participants incur more price

Table 1: Summary statistics for the constructed P&L series over the entire sample period (Jun 1986 through Mar 2021), where $\hat{\mu}, \hat{\sigma}, \hat{\nu}, \hat{\kappa}$ denote the estimated mean, st.dev, skewness, and kurtosis respectively. Augmented Dickey-Fuller and Ljung-Box tests are performed on the P&L series, the test statistics are denoted by ADF and JB-Q respectively. I denote the Ljung-Box test statistic on the squared P&L series with LB-Q^2.

statistic	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\nu}$	$\hat{\kappa}$	ADF	LB-Q	LB-Q^2
ΔS_C	-0.028	2.432	0.293	5.140	-10.62*	64.12*	1124.34*
ΔS_G	-0.038	3.167	0.470	3.606	-14.20*	54.08*	691.37*
ΔS_{HO}	-0.032	2.863	-0.031	5.609	-12.83*	38.51*	925.77*
ΔF_C	-0.029	2.385	0.334	5.265	-10.55*	75.25*	1109.28*
ΔF_G	-0.039	3.147	0.441	4.123	-11.15*	44.38*	604.30*
ΔF_{HO}	-0.036	2.774	0.075	4.902	-12.49*	48.91*	1272.13*

Rejection of the null-hypothesis at 1% significance level is denoted with a *.

uncertainty. Also, Downey (2009) states that gasoline is the most profitable commodity for refineries, the combination between extra profitability and uncertainty reinforces the need for risk management. Table 1 shows that all futures contracts exhibit slightly lower skewness than their spots counterparts. Alexander et al. (2013) state that this might be the result of more unexpected shocks in short term consumption behavior. The skewness, denoted by $\hat{\nu}$, is positive for all spots, and futures, except heating-oil spots. The ADF test in Table 1 shows that all P&L series are stationary. A Ljung-Box test shows that all P&L loss series exhibit autocorrelation at a 1% significance level. Furthermore, a Ljung-Box test on the squared P&L series shows presence of volatility clustering at a 1% significance level.

Table 2 shows the estimated Kendall's τ matrix. The spot-future pairs for each commodity exhibit the highest dependence. The dependence decreases in the order of crude, heating-oil, and gasoline. Across the commodities, we also see high dependence in the range of $0.5 - 0.6$.

Table 2: Estimated Kendall's τ dependence matrix based on entire sample (Jun 1986 through Mar 2021).

	ΔS_C	ΔS_G	ΔS_{HO}	ΔF_C	ΔF_G	ΔF_{HO}
ΔS_C	1	0.50	0.59	0.89	0.53	0.62
ΔS_G	-	1	0.50	0.51	0.76	0.51
ΔS_{HO}	-	-	1	0.59	0.52	0.85
ΔF_C	-	-	-	1	0.54	0.62
ΔF_G	-	-	-	-	1	0.55
ΔF_{HO}	-	-	-	-	-	1

3 Multi-Commodity Hedging Problem

3.1 Refinery Portfolio Profit and Loss

Haigh and Holt (2002) propose an oil refinery hedging framework, they focus on gasoline and heating-oil as primary refined products from crude. A general oil refinery roughly cracks three units of crude oil into two units gasoline and one unit heating-oil (Energy Information Administration, 2002). This refining ratio is known as the $3:2:1$ *crack spread*, and can be used as proxy for refinery profit margins. Due to crack spread importance, the NYMEX introduced a $3:2:1$ *crack spread* contract (CME Group, 2017). Thus, A refinery might hedge their price risk by using said NYMEX contract, or any custom ratio of futures contracts. Consistent with Haigh and Holt (2002), I assume the refinery determines appropriate hedging ratios in week $t-1$. One week later, in week t, the refined products are sold to the market at the corresponding spot rates, where any loss might be offset by closing futures positions. The refinery's hedging problem stems from the required refining time, and the maturity mismatch between refining completion and the futures expiry. (1) Shows the per unit operational Profit and Loss (P&L) as function of time

$$\pi_t(\boldsymbol{\beta}_t) = \underbrace{-S_t^C + \frac{2}{3}S_t^G + \frac{1}{3}S_t^{HO}}_{\text{time } t \text{ unhedged profit}} + \underbrace{\boldsymbol{\beta}_t' \Delta \boldsymbol{F}_t^{CS}}_{\text{hedging effect}}, \tag{1}$$

where S_t^k with $k \in \{C, G, HO\}$ denotes crude, gasoline, and heating-oil spot rates respectively, and where $\boldsymbol{\beta}_t = (\beta_t^C, \beta_t^G, \beta_t^{HO})'$ denote the hedging ratios determined at time $t-1$. Furthermore, the vector $\Delta \boldsymbol{F}_t^{CS} = (\Delta F_t^C, -\frac{2}{3}\Delta F_t^G, -\frac{1}{3}\Delta F_t^{HO})'$ in (1), denotes the futures P&L (with $\Delta F_t^k = F_t^k - F_{t-1}^k$). Thus, the hedging effect in (1) is given by the futures positions $\boldsymbol{\beta}_t' \Delta \boldsymbol{F}_t^{CS}$. The, unhedged profit at time t is made by selling the gasoline and heating oil. To sustain operations, the refinery has to buy new crude oil at time t, resulting in a loss of S_t^C. Consistent with Sukcharoen and Leatham (2017) I rewrite (1) as a function of the spots P&L

$$\pi_t(\boldsymbol{\beta}_t) = -\Delta S_t^C + \frac{2}{3}\Delta S_t^G + \frac{1}{3}\Delta S_t^{HO} + \boldsymbol{\beta}_t' \Delta \boldsymbol{F}_t^{CS} - S_{t-1}^{CS}, \tag{2}$$

where $\Delta S_t^k = S_t^k - S_{t-1}^k$ is the marginal spot price change with $k \in \{C, G, HO\}$, for crude, gasoline, and heating oil spots respectively. Furthermore, the time $t-1$ crack spread profit $S_{t-1}^{CS} = -S_{t-1}^C + \frac{2}{3}S_{t-1}^G + \frac{1}{3}S_{t-1}^{HO}$ is known at $t-1$, and does therefore not impact optimal hedging ratios. Similar to Alexander et al. (2013) and Sukcharoen and Leatham (2017), I disregard S_{t-1}^{CS} from (2) to obtain the profit equation in (3)

$$\pi_t(\boldsymbol{\beta}_t) = -(\Delta S_t^C - \beta_t^C \Delta F_t^C) + \frac{2}{3}(\Delta S_t^G - \beta_t^G \Delta F_t^G) + \frac{1}{3}(\Delta S_t^{HO} - \beta_t^{HO} \Delta F_t^{HO}), \tag{3}$$

where ΔS_t^k, and ΔF_t^k with $k \in \{C, G, HO\}$, for crude, gasoline, and heating-oil spot and futures P&L respectively.

3.2 Hedging Cost

Costs of hedging are an important aspect of the hedging strategy. Alexander et al. (2013) find that the hedging costs can cause a naive implementation to outperform complex models. The hedging cost consists of three aspects; 1) being the transaction cost, 2) being the bid-ask spread, 3) being the margin requirement for positions futures contracts.[1] First, consistent with Alexander et al. (2013), I assume that transaction costs per futures contract amount to $C_{t-1}^{TC} = \$1.45$. Second, (4) formulates the time $t-1$ incurred cost due to the bid-ask spread

$$C_{t-1}^{BA}(\boldsymbol{\beta}_t) = \gamma_C \beta_t^C F_{t-1}^C + \frac{2}{3}\gamma_G \beta_t^G F_{t-1}^G + \frac{1}{3}\gamma_{HO} \beta_t^{HO} F_{t-1}^{HO}, \tag{4}$$

[1] A future contract concerns 1000 barrels or the volumetric equivalent, usually market quotes are per unit. For analysis, I consider fictive single barrel contracts that, when aggregated, are consistent with their real counterpart.

where γ_k for $k \in \{C, G, HO\}$, are the bid-ask spreads for each respective commodity. Dunis et al. (2008) find that the NYMEX bid-ask spreads are approximately 1 bps for crude, 10 bps for gasoline, and 12 bps for heating-oil. Third, the cost of the margin deposit is

$$C_{t-1}^{M}(\boldsymbol{\beta}_t) = \mathcal{F}_M r_d(\beta_t^C F_{t-1}^C + \frac{2}{3}\beta_t^G F_{t-1}^G + \frac{1}{3}\beta_t^{HO} F_{t-1}^{HO}), \tag{5}$$

where $\mathcal{F}_M$ is the required margin deposit as fraction of the contract value. Some simple analysis over the span of Jan 2020-Mar 2021, reveals that the fractional margin requirements are approximately equal for each of the three futures. I therefore take the margin requirement to be a constant fraction of $\mathcal{F}_M = 6\%$ for all contracts. Consistent with Alexander et al. (2013), I assume that the refinery finances the margin requirement using commercial debt. The three largest U.S. refineries (Energy Information Administration, 2020), without vertical integration, currently have BBB credit ratings from Standard & Poors. I therefore take the Standard & Poors Investment Grade Corporate Bond Index as proxy for the cost of debt, whose yield is approximately $r_d = 2.2\%$ as of April 2021. I disregard marking to market costs due to their relatively small impact, which can both be positive or negative. Moreover, the aforementioned refineries also do not report these costs.

The P&L equation for one business cycle of two weeks, including hedging costs, is therefore given by

$$\pi_t^*(\boldsymbol{\beta}_t) = \pi_t(\boldsymbol{\beta}_t) + C_{t-1}^{BA}(\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}) + C_{t-1}^{M}(\boldsymbol{\beta}_t) + C_{t-1}^{TC}, \tag{6}$$

where $\pi_t(\cdot)$ is the hedged P&L from (3), t_m denotes the start of the month, $C_{t-1}^{BA}(\cdot)$ is the bid-ask spread cost, $C_{t-1}^{M}(\cdot)$ is the margin requirement cost, and C_{t-1}^{TC} is the transaction cost.

3.3 Empirical Methodology

Hedging ratios $\boldsymbol{\beta}_t$ should minimize risk such that

$$\boldsymbol{\beta}_t^* = \arg\min_{\boldsymbol{\beta}_t} \left\{ \rho(\pi_t^*(\boldsymbol{\beta}_t)) \right\}, \tag{7}$$

where $\rho(\cdot)$ denotes a risk measure, and $\pi_t^*(\boldsymbol{\beta}_t)$ is the portfolio P&L from (6). I consider two risk measures: value-at-risk denoted by $\rho_{\text{VaR}}(\cdot)$, and expected shortfall denoted by $\rho_{\text{ES}}(\cdot)$.

Furthermore, I consider three probability levels $p \in \{0.10, 0.05, 0.01\}$. McNeil et al. (2015) provides empirical estimators

$$\hat{\rho}_{\mathrm{VaR}}(\pi_t^*) = F_n^{\leftarrow}(1-p) = L_{(\lceil n(1-p) \rceil)}, \tag{8a}$$

$$\hat{\rho}_{\mathrm{ES}}(\pi_t^*) = \frac{1}{np}\left(\left\{ \lceil n(1-p) \rceil - n(1-p) \right\} L_{(\lceil n(1-p) \rceil)} + \left(\sum_{i=\lceil n(1-p) \rceil + 1}^{n} L_{(i)} \right) \right), \tag{8b}$$

where $F_n^{\leftarrow}$ is generalized inverse of the empirical CDF of π_t^* , and $L_{(n)}$ denotes the n-th ordered statistic of observed losses.

Table 1 shows the presence of volatility clustering in each marginal P&L series, and might be crucial for risk management. However, modelling volatility clustering also adds parameter uncertainty. I investigate the potential parsimony benefit from assuming independence by comparing two HOPAC implementations. In the first implementation (HOPAC-i.i.d.), I assume i.i.d. observations from each marginal. In the second implementation (HOPAC-GARCH) I account for volatility clustering with a GARCH(1,1) model in each marginal. Therefore, this is in essence a multivariate-GARCH model where all cross-correlations are 0. In both cases I use a five-year backward looking window (n=260) to estimate the models.[2] I estimate the models spanning the entire sample leading to 1541 separate windows. Next, I simulate 10,000 values from each model, and transform them to hypothetical P&L observations to determine hedging ratios $\boldsymbol{\beta}_t$ from (7). Furthermore, I compare the HOPAC-i.i.d. and HOPAC-GARCH implementations to a naive hedging strategy which assumes complete delta hedging, such that $\boldsymbol{\beta}_t^{naive} = (1,1,1)'$ regardless of market conditions.

To fit the copula models, we need to construct pseudo observations. In accordance with Sklar's theorem (Sklar, 1959), I apply a quantile transformation to each marginal, in each window separately by using the empirical distribution function. I subsequently use the pseudo-observations from (9) to perform model estimation.

$$\boldsymbol{u}_j = \frac{n}{n+1}\hat{F}_{n,j}\left(\boldsymbol{x}_j\right) = \frac{\mathrm{Rank}\left(\boldsymbol{x}_j\right)}{n+1}, \tag{9}$$

[2]The five-year backward looking window is consistent with Ji and Fan (2011), Alexander et al. (2013), and Sukcharoen and Leatham (2017) to name a few.

where $\boldsymbol{x}_j$ denotes the vector of observations with corresponding variables $\boldsymbol{u}_j$ from marginal j, and where $\hat{F}_{n,j}(\cdot)$ denotes the respective empirical cdf. Let $j \in \{1, 2, \ldots, 6\}$ denote the corresponding margins of $\{S_C,\ S_G, \ldots, F_{HO}\}$. For the HOPAC-i.i.d. implementation, the original marginal P&L series $\boldsymbol{x}_j \in \{\Delta S_C,\ \Delta S_G, \ldots, \Delta F_{HO}\}$ are used. Let $\hat{\boldsymbol{z}}_t := (\hat{\boldsymbol{z}}_{1,t}, \hat{\boldsymbol{z}}_{2,t} \ldots, \hat{\boldsymbol{z}}_{6,t})$ denote a 260×6 matrix with GARCH residuals from window t. In the HOPAC-GARCH implementation, the marginal GARCH residuals are used instead $\boldsymbol{x}_j \in \{\hat{\boldsymbol{z}}_{1,t}, \hat{\boldsymbol{z}}_{2,t} \ldots, \hat{\boldsymbol{z}}_{6,t}\}$. Then, let the 260×1 vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2 \ldots, \boldsymbol{u}_6\}$ denote the set of marginal pseudo-observations obtained from (9). P&L observations can be reconstructed from $\{F^{\leftarrow}_{n,1}(\boldsymbol{u}_1), F^{\leftarrow}_{n,2}(\boldsymbol{u}_2) \ldots, F^{\leftarrow}_{n,6}(\boldsymbol{u}_6)\}$, after which applying the respective GARCH model if applicable. I evaluate the risk reduction using (10), which represents the risk reduction

$$\text{RR} = \left(1 - \frac{\rho\left(\pi^*_t(\boldsymbol{\beta}_t)\right)}{\rho\left(\pi^*_t(\mathbf{0})\right)}\right) \times 100, \tag{10}$$

where $\rho\left(\pi^*_t(\mathbf{0})\right)$ is the unhedged risk from (6), and RR measures the % risk reduction achieved by the hedging approach.

4 Hierarchical Outer Power transformed Archimedian Copulas

4.1 Background Theory

I briefly review the three main HOPAC ingredients; Archimedean copulas (AC), Outer Power transforms (OP), and hierarchical copulas. First, a regular bivariate AC can be used to model dependence between two random variables, say between the univariate P&L of the crude oil spot u_1, and future u_4 from Table 2, which gives

$$C_\psi\left(u_1, u_4\right) = \psi\left\{\psi^{-1}\left(u_1\right) + \psi^{-1}\left(u_4\right)\right\}, \tag{11}$$

where $C_\psi : [0,1]^2 \to [0,1]$ is a bivariate copula, with Archimedean generator $\psi(\cdot)$, and its inverse $\psi^{-1}(\cdot)$. Some restrictions apply to Archimedean generators such that the copula is a proper CDF. Specifically, $\psi(\cdot)$ is a strictly decreasing continuous convex function $\psi : [0, \infty) \to [0, 1]$ with $\psi(0) = 1$ and $\lim_{t\to\infty} \psi(t) = 0$. Additionally, if $(-1)^k \psi^{(k)}(t) \geq 0,\ \forall k \in \mathbb{N}_+,\ t \in \mathbb{R}$, then $\psi(\cdot)$ is called completely monotone (c.m.) (McNeil et al., 2015). Crucially, a Laplace-Stieltjes

transform can be applied to c.m. generators, which enable efficient sampling (McNeil, 2008). We investigate the sampling details at a later stage.

Next, Nelsen (2007) introduces the outer power (OP) transformation as modification to a regular AC. The OP transform enables a better fit in both body and tail dependence (Górecki et al., 2021). Let

$$\mathring{\psi}(u) = \psi(u^{1/\beta}), \text{ with } \beta \in [1, \infty), \tag{12}$$

where $\mathring{\psi}(\cdot)$ denotes the OP transformed generator with parameter β. The OP parameter β is restricted to $[1, \infty)$ such that the generator stays c.m. and creates a proper copula. If $\beta = 1$, the original AC generator is obtained. Then, let $\psi_{(a,\theta,\beta)}$ denote a generator from Archimedean family a, with parameter θ, and OP transform β. Also, let C_{ψ_k} denote a copula with generator $\psi_{(a,\theta_k,\beta_k)}$. I consider four families; Ali–Mikhail–Haq (A), Clayton (C), Frank (F), and Joe (J). Table 3 shows the original generators $\psi_{(a,\theta,1)}$, upper-tail dependence λ_u, and parameter space Θ_a. Clearly, the extra OP transform increases the ability to capture tail dependence.

Table 3: Archimedean family with generator $\psi_{(a,\theta,1)}$, upper tail dependence λ_u, and parameter space Θ_a.

Family a	$\psi_{(a,\theta,1)}(u)$	λ_u	Θ_a
Ali–Mikhail–Haq (A)	$(1-\theta)/\left(\exp(u)-\theta\right)$	$2-2^{1/\beta}$	$[0,1)$
Clayton (C)	$(1+\theta u)^{-1/\theta}$	$2-2^{1/\beta}$	$(0,\infty)$
Frank (F)	$-\theta^{-1}\log\left\{1-(1-\exp(-\theta))\exp(-u)\right\}$	$2-2^{1/\beta}$	$(0,\infty)$
Joe (J)	$1-\left(1-\exp(-u)\right)^{1/\theta}$	$2-2^{1/(\theta\beta)}$	$[1,\infty)$

Flexibility in (tail) dependence modelling is of particular importance to the oil refinery problem (Sukcharoen and Leatham, 2017).

Next, A single multivariate AC restricts the dependence among all combinations of margins to be identical (Hofert, 2010). Asymmetric dependence structures can be modelled by constructing a hierarchy of ACs (HAC). One can construct a HAC by replacing some marginal arguments in (11) with other copulas, Figure 1 shows such hierarchy of copulas. Mathematically;

$$C(\boldsymbol{u}) = C_{\psi_5}\left(C_{\psi_4}\left(C_{\psi_2}\left(u_3, u_6\right), C_{\psi_1}\left(u_1, u_4\right)\right), C_{\psi_3}\left(u_2, u_5\right)\right) \tag{13}$$

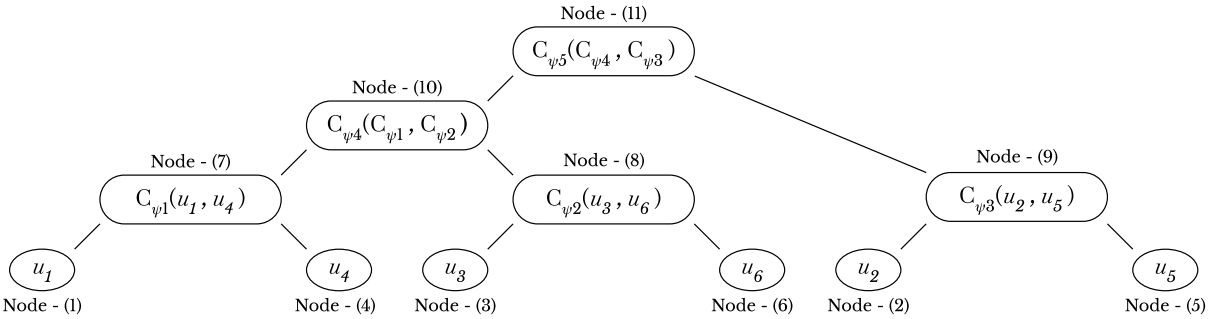

Figure 1: Hierarchical copula structure from (13). Descendant leaves of *node* k are denoted with $\downarrow(k)$. For example, *node* 7 corresponds to copula 1, such that $\downarrow(7) = \{1, 4\}$.

where $C_{\psi_k}(\cdot)$ denotes a copula C with generator $\psi_{(a,\theta_k,1)}$. Figure 1 illustrates the HAC example from (13). One can construct a HOPAC by using OP transformed generators in (13), instead of regular generators. Restrictions apply leading Górecki et al. (2021) to derive parameter bounds, using the sufficient nesting condition from Joe (1997). These parameter bounds are used in estimation, and the technical details are deemed beyond the scope of this research.

4.2 HOPAC Estimation

First is hierarchy estimation, for which I use *algorithm 1* from Górecki et al. (2021). I start with definitions; $\nu := \{1, \ldots, 2d-1\}$ denotes the set of *nodes*, and ϑ denotes the set of arcs. We have $d = 6$ price series called *leaves*, the $d-1$ remaining *nodes* are called *forks*. The *forks* form the hierarchy of copulas, which model the *leaves'* dependence structure. For example consider the hierarchy from Figure 1, here the copula margins are *nodes* $\{1, \ldots, 6\}$. The *forks* correspond to *nodes* $\{7, \ldots, 11\}$. Górecki et al. (2021) numerate *forks* uniquely by assigning the first *fork* (*node* $d+1$) with the highest dependence, and decreasing from there-on. The descendant *leaves* of *node* k are denoted with $\downarrow(k)$. In Figure 1 for example, $\downarrow(7) = \{1, 4\}$ and $\downarrow(10) = \{\{1, 4\}, \{3, 6\}\}$.

Algorithm 1 works by iterating until all *forks* are identified. We start with *node* $d+k$ (*fork* k), and search for its location in the hierarchy. In line 2, we calculate the average dependence between all descendant *leaves* of the remaining *nodes* in ξ. In the first iteration ξ contains all d *leaves*, such that all descendant *leaves* are the *leaves* themselves. Next, in line 3 let *nodes* i

Algorithm 1 Pseudo code for hierarchy estimation from Górecki et al. (2021)

Definitions

— $\hat{\boldsymbol{\tau}}(\boldsymbol{U})$: estimated Kendall's τ matrix using pseudo-observations $\boldsymbol{U}$ from (9).

— ξ: set of *nodes* to be coupled.

— $\boldsymbol{\eta}$: vector containing the children from *node* k

Initial	1. $\hat{\boldsymbol{\tau}}(\boldsymbol{U})$, $\xi := \{1, ..., 6\}$, $\vartheta := \emptyset$, $\boldsymbol{\eta} := \emptyset$
for $k = 1, ..., d-1$ **do**	
—	2. For *nodes* $(i,j) \in \xi$, let $\hat{\tau}_{(i,j)} := \text{mean}(\hat{\tau}_{\downarrow(i)\times\downarrow(j)})$, where $\downarrow(i)\times\downarrow(j)$ is the set of all cross combinations between $\downarrow(i)$ and $\downarrow(j)$.
—	3. Set $(i,j) := \arg\max_{i \leq j}\{\hat{\tau}_{(i,j)}\}$, with $i, j \in \xi$.
—	4. Set $\xi := \xi \cup \{d+k\} \setminus \{i,j\}$, where ξ is the set of *nodes* under consideration.
—	5. Set $\hat{\tau}_{d+k} := \hat{\tau}_{(i,j)}$.
—	6. Set $\vartheta := \vartheta \cup \{\{i, d+k\}, \{j, d+k\}\}$, where ϑ is the set of arcs.
—	7. Set $\eta_k := (i,j)$, which are the children of *node* k
end for	
return	$\{\ \vartheta,\ \boldsymbol{\eta}_{1:d-1},\ \hat{\boldsymbol{\tau}}_{d+1:2d-1}\ \}$

and j maximize $\hat{\tau}_{i,j}$. For example, consider *node* $d+k = 10$ with $\downarrow(10) = \{\downarrow(7), \downarrow(8)\}$ from Figure 1, which means that $\hat{\tau}_{7,8} = \frac{1}{4}(\hat{\tau}_{1,3} + \hat{\tau}_{1,6} + \hat{\tau}_{3,4} + \hat{\tau}_{4,6})$ is the combination which exhibits the strongest dependence. Therefore *algorithm 1* assigns *nodes* $\{7, 8\}$ to be the margins of *fork* 10. In line 4, we remove *nodes* i and j from the set of *nodes* under consideration in ξ, and add *node* $d+k$ to ξ. Line 4 ensures we only model dependence between *leaves* of which we have not modelled dependence yet. In line 5 through 7, we set the dependence $\hat{\tau}_{d+k}$, the arcs ϑ_{d+k}, and the children $\boldsymbol{\eta}_{d+k}$ for *node* $d+k$.

Next is HOPAC parameter estimation, for which Górecki et al. (2021) propose *Algorithm 2*. I start with an overview and provide an example afterwards. Consistent with Górecki et al. (2021), I let each copula be from the same family a for simplicity. The algorithm starts at the *root* (*node* $2d-1$) of the hierarchy, and traverses sequentially through all *forks*. Consider *fork*

Algorithm 2 Pseudo code for HOPAC estimation from Górecki et al. (2021)

Definitions

— $\psi_k(a,\, \theta, \beta)$: *node* k copula from family a with parameter θ and OP-transform β.

— Θ : the allowed parameter space for θ, initially determined by the copula family Θ_a.

— $\mathcal{B}$: the allowed parameter space for β

— $\downarrow\!(i) \times \downarrow\!(j)$: is the set of all combinations between $\downarrow(i)$ and $\downarrow(j)$.

Initial 1. $\boldsymbol{U},\ \boldsymbol{\eta},\ \Theta := \Theta_a,\ \mathcal{B} := [1, \infty)$

for *node* $k = (2d-1),\ ...,\ (d+1)$ **do**

— 2. Let $(i, j) := \eta_k$, be the children of *node* k (from algorithm 1).

— 3. Let $\tilde{i}$ and $\tilde{j}$ be *leaves*, where $\tilde{i} \in \downarrow(i)$ and $\tilde{j} \in \downarrow(j)$. Then obtain copula parameters $(\tilde{\theta}_{\tilde{i},\tilde{j}},\ \tilde{\beta}_{\tilde{i},\tilde{j}})$ by MLE, for all combinations between $\tilde{i}$ and $\tilde{j}$.

$$(\tilde{\theta}_{\tilde{i},\tilde{j}},\ \tilde{\beta}_{\tilde{i},\tilde{j}}) := \arg\max \left\{ \sum_{m=1}^{n} \log c_{\psi\left(a,\, \theta_{\tilde{i},\,\tilde{j}}, \beta_{\tilde{i},\tilde{j}}\right)} \left(u_{m,\tilde{i}}\,,\ u_{m,\tilde{j}} \right) \right\}$$

$$\text{Subject to: } \tilde{\theta}_{\tilde{i},\tilde{j}} \in \Theta \ \text{ and } \ \tilde{\beta}_{\tilde{i},\tilde{j}} \in \mathcal{B}$$

— 4. $\hat{\psi}_k\Big(a,\, \hat{\theta}, \hat{\beta}\Big) := (\text{mean}(\tilde{\theta}_{\downarrow(i)\times\downarrow(j)}), \text{mean}(\tilde{\beta}_{\downarrow(i)\times\downarrow(j)}))$

if $\hat{\beta} \in [1, \beta_{\mathcal{R}}]$ **do**

— 5. Set $\Theta := \Theta \cap [\hat{\theta}, \infty)$ and $\mathcal{B}$ remains unchanged.

else do

— 6. Set $\Theta := [\hat{\theta}, \hat{\theta}]$ and $\mathcal{B} := [\hat{\beta}, \infty)$.

end for

return $\{\ \hat{\boldsymbol{\Psi}}\Big(a,\, \hat{\theta}, \hat{\beta}\Big)\ \}$

k, and let i and j be its children. Then, the copula parameter $\hat{\theta}$, and OP parameter $\hat{\beta}$, a t *fork* k, are estimated by averaging the ML estimates $(\tilde{\theta}, \tilde{\beta})$ from all cross combinations of descendant *leaves* between $\downarrow(i)$ and $\downarrow(j)$. Two restrictions apply to each ML estimate, such that $\tilde{\theta} \in \Theta$ and $\tilde{\beta} \in \mathcal{B}$. The parameter bounds Θ and $\mathcal{B}$ follow from the sufficient nesting conditions derived by Górecki et al. (2021) for HOPACs. Initially Θ is defined by the generator family a such that $\Theta = \Theta_a$, whereas $\mathcal{B} = [1, \infty)$. Next, the parameters of the succeeding *fork* $k-1$ are subject to new bounds, determined by one of the two restrictions applied in line 5 or 6 of the pseudo

code. Whichever restriction applies depends on, if the estimated OP parameter $\hat{\beta}$, is within the interval $[1, \mathcal{B}_{\mathcal{R}}]$, where I let $\mathcal{B}_{\mathcal{R}} = 1.05$ consistent with Górecki et al. (2021).

For example, let *fork* k be the fourth copula (*node* 10) from Figure 1; such that its children $\{i, j\}$ are *nodes* $\{7, 8\}$, and its descendant *leaves* are $\downarrow(10) = \{\{1, 4\}, \{3, 6\}\}$. Furthermore, let $\left(\tilde{\theta}_{i,j}, \tilde{\beta}_{i,j}\right)$ maximize the likelihood for a copula with marginals i, j subject to the aforementioned constraints. Then, for the copula at *node* 10, the estimate $\hat{\theta} = \frac{1}{4}(\tilde{\theta}_{1,3} + \tilde{\theta}_{1,6} + \tilde{\theta}_{4,3} + \tilde{\theta}_{4,6})$, and the estimate $\hat{\beta} = \frac{1}{4}(\tilde{\beta}_{1,3} + \tilde{\beta}_{1,6} + \tilde{\beta}_{4,3} + \tilde{\beta}_{4,6})$.

4.3 Sampling

The HOPAC sampling algorithm can be inferred from the work of Górecki et al. (2021), and is closely related to the algorithm from McNeil (2008). Specifically, Górecki et al. (2021) prove two adjustments such that the algorithm from McNeil (2008) can be used for HOPACs. I discuss some important principles of the algorithm from McNeil (2008), and provide a simple example along with pseudo code intended to aid the reader. Besides, I provide a detailed account of the sampling method in algorithm 3.

The algorithm from McNeil (2008) is based on sampling from the equivalent Laplace-Stieltjes (LS) distribution of the generating function $\psi(\cdot)$ from (11). I closely follow the work of Joe (1997) to illustrate the equivalence. Specifically, let generator $\psi(\cdot)$ be the LS-transform of distribution function (df) $F(\cdot)$ on $[0, \infty)$ with $F(0) = 0$

$$\psi(t) = \int_0^\infty \mathrm{e}^{-t\alpha} \, \mathrm{d}F(\alpha), \quad t \geq 0, \tag{14}$$

where $\psi(0) = 1$. Next, Joe (1997) states that a unique cdf $G(\cdot)$ exists for a univariate cdf $M(\cdot)$ such that

$$M(x) = \int_0^\infty G^\alpha(x) dF(\alpha) = \psi(-\log G(x)). \tag{15}$$

Naturally, (15) implies that $G(\cdot) = \exp\{-\psi^{-1} \circ M(\cdot)\}$. Consider a simple bivariate case where $M_1(\cdot)$, $M_2(\cdot)$ are standard uniform (marginals), with the corresponding $G_1(\cdot)$, $G_2(\cdot)$ from (15). Then a bivariate copula can be constructed

$$C(u_1, u_2) = \int_0^\infty G_1^\alpha G_2^\alpha dM(\alpha) = \psi\left(-\log G_1 - \log G_2\right) = \psi\left(\psi^{-1} \circ M_1(u_1) + \psi^{-1} \circ M_2(u_2)\right), \tag{16}$$

where u_1, u_2 are rvs from the standard uniform marginals $M_1(\cdot)$, $M_2(\cdot)$. Clearly (16) is equivalent to the bivariate AC class in (11). Joe (1997) also provides the multivariate generalization, which I omit for brevity. Moreover, I stick to the bivariate copula case to illustrate the sampling algorithm from McNeil (2008) in its most intuitive form.

Bivariate AC Sampling Example

The algorithm from McNeil (2008) is essentially based on that of Marshall and Olkin (1988), who recognize that the LS representation provides a convenient basis for a sampling algorithm. I follow the formulation from McNeil (2008), which I restate for the bivariate example;

1. Generate a variate $V \sim F(\cdot)$ with $F = \mathcal{LS}^{-1}[\psi(\cdot)]$.
2. Generate two i.i.d. uniform variates (x_1, x_2).
3. Return $u_i = G^{-1}(x_i^{1/V})$ for $i = 1, 2$.

Marshall and Olkin (1988) show that a variate from marginal df M_i (step 3) can be generated from $G^{-1}(x_i^{1/V})$, which McNeil (2008) simply rewrite using the implication from (15) as $\psi(-\log(\mathrm{x_i})/V)$. I illustrate the algorithm with a simple bivariate AC. Figure 2 shows a simple

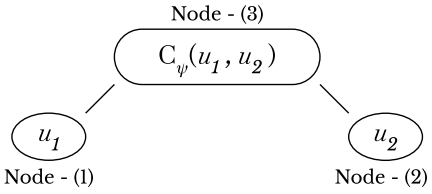

Figure 2: Example of a bivariate AC, with generator ψ and marginals $u1, u2$.

bivariate AC with generator $\psi(\cdot)$. *Step 1* in the algorithm is generating variate V from the LS df $F(\cdot)$ corresponding to the generator $\psi(\cdot)$. Next, in *step 2* we generate two i.i.d. uniformly distributed variates x_1, x_2. In *step 3* we use the result from Marshall and Olkin (1988) to obtain variates u_1, u_2 which follow the marginal distributions M_1, M_2. Clearly this algorithm provides a convenient way of simulating variates from the margins of the AC $C(u_1, u_2)$.

The more general algorithm from McNeil (2008) for sampling from hierarchical copula structures, can be obtained by generalizing the bivarate case through including a recursive argument

between *step 1* and *step 2*, say *step 1.5*. In step *step 1.5*, one samples from a LS distribution that corresponds to that of the nested generator family.[3] The OP-transform requires two intermediate, but straightforward modifications proposed by Górecki et al. (2021). Rather than repeating each adjustment separately, I provide the overview in *algorithm 3*.

Similar to *algorithm 2*, the HOPAC sampling algorithm starts at the *root* of the hierarchy, and traverses sequentially through all *forks*, introducing dependence on the way. The HOPAC

Algorithm 3 Sampling algorithm inferred from McNeil (2008) and Górecki et al. (2021)

Notation

— Let $\{\theta_k,\ \beta_k\}$ be parameters from the generator at *fork* k: $\psi_k(a,\ \theta, \beta)$.

— Let $\mathrm{S}\{\cdot, \cdot, \gamma, \delta\}$ denote the stable distribution with scale γ, and location δ

Initial 1. Sample $V_{d-1} \sim F$ with $F = \mathcal{LS}^{-1}[\psi_{d-1}]$ from Table 4

— 2. Independently sample $S \sim \mathrm{S}\left\{1/\beta_{d-1}, 1, \cos^{\beta_{d-1}}(\pi/(2\beta_{d-1})), \mathbb{1}_{\{\beta_{d-1}=1\}}\right\}$

— 3. Set $\mathring{V}_{d-1} := S\,(V_{d-1})^{\beta_{d-1}}$

for *fork* $k = (d-2),\ \ldots,\ 1$ **do**

— 4. Let p be the parent *fork* of k with generated random variate $\mathring{V}_p$

— 5. Sample $\mathring{V}_k \sim \mathrm{S}\left(\beta_p/\beta_k, 1, \left\{\cos\left(\frac{\beta_p}{\beta_k}\frac{\pi}{2}\right) V_p\right\}^{\beta_p/\beta_k}, \mathring{V}_p \mathbb{1}_{\{\beta_p/\beta_k=1\}}\right)$

if *fork* k has children, repeat for each child i **do**

— 6. Independently sample $X_i \sim \mathbf{U}\,(0, 1)$

— 7. $u_i := \psi_k\left(-\ln(X_i)/\mathring{V}_p\right)$

end for

return $\{u_1, u_2, \ldots, u_6\}$

sampling algorithm, inferred from the work of Górecki et al. (2021) and McNeil (2008), starts in line 1 at the *root* where variate V_{d-1} is sampled from $F(\cdot)$. An adjustment from Górecki et al. (2021) in line 2 and 3, makes sure that we sample from the OP-transformed generator. We continue on line 4, where we consider *node* k with parent *fork* p. In line 5 we introduce

[3] Hofert (2010) provides a comprehensive overview of the LS-transforms for some well known single parameter copula families. Joe (1997) provides a comprehensive overview of the mathematical derivations.

dependence of *fork* k by generating a new variate V_k, which depends on the parent variate V_p. A regular HAC is sampled using the LS transformed distribution, which depends on the nested generators. However, Górecki et al. (2021) prove that the LS transform of nested OP generators always correspond to a stable distribution. Last in line 6 and 7, if *fork* k has children, we apply the transformation from Marshall and Olkin (1988) to i.i.d. uniformly distributed variates to get the pseudo-observations from the corresponding marginals. These last two steps are *step 2* and *3* from the original algorithm proposed by Marshall and Olkin (1988). Naturally, we repeat line 4 through 7 until a pseudo-observation is generated for each dependent *leaf* in the structure.

Table 4: Archimedean family with corresponding Laplace-Stieltjes (LS) distribution $F(\cdot)$ from Hofert (2010). Where Geo$(\cdot)$ denotes the geometric distribution, $Gam(\cdot)$ denotes the gamma distribution, and Logser$(\cdot)$ denotes the log-series distribution. The Joe LS-df follows a non-standard discrete distribution (Hofert, 2010).

Family	A	C	F	J
$\mathcal{LS}^{-1}[\psi(\cdot)]$	Geo$(1-\theta)$	Gam$(1/\theta, 1)$	Logser $\left(1-e^{-\theta}\right)$	$\binom{1/\theta}{k}(-1)^{k-1}, k \in \mathbb{N}$

Table 4 shows the LS transforms corresponding to each copula family I use, and is adopted from Hofert (2010). Here, Geo(p) denotes the geometric distribution with success-probability $\pi \in (0,1]$, and density $p_k = \pi(1-\pi)^{k-1}$ on $k \in \mathbb{N}$. Gam(α, β) denotes the gamma distribution, with shape and scale $\alpha, \beta \in (0,\infty)$, and df $f(x) = \beta^\alpha x^{\alpha-1} \exp\{-\beta x\}/\Gamma(\alpha)$ on $x \in (0,\infty)$. Logser(π) denotes the log-series distribution with $\pi \in (0,1)$, and density $p_k = \pi^k/(-k\log(1-\pi))$ on $k \in \mathbb{N}$. For Joe $\theta \in [1,\infty)$ such that $1/\theta \in (0,1]$, I therefore generalize the density provided by Hofert (2010) using gamma functions, such that $p_k = \frac{\Gamma(1/\theta+1)}{\Gamma(1/\theta+1-k)\,k!}$ on $k \in \mathbb{N}$.

5 Results

I compare effectiveness of a naive strategy, and two HOPAC implementations. The naive hedging strategy corresponds to a complete delta hedge, where all hedging ratios are 1 regardless of market conditions. Instead, HOPACs can model the complete dependence structure of com-

modity spots and futures for determining risk minimizing hedging ratios. I implement HOPACs from four distinct Archimedean families; Ali–Mikhail–Haq (A), Clayton (C), Frank (F), and Joe (J). Furthermore, I also investigate whether GARCH$(1,1)$ models should be used to account for the presence of volatility clustering. Thus granting two distinct HOPAC implementations, HOPAC-i.i.d. wherein volatility clustering is disregarded, and HOPAC-GARCH wherein volatility clustering is accounted for. The four Archimedean families combined with two time dependence assumptions lead to eight implementations. Each model is estimated on a 5 year backward looking window ($n = 260$ weeks) in 1541 windows (12 Jun 1991 - 3 Mar 2021), and is used to determine optimal hedging ratios. I evaluate the realised risk across the 1541 windows in an ex-post analysis. Furthermore, I also use an adjusted Kolmolgorov-Smirnov (K-S) test, to test the stochastic properties of the realised loss distributions.

5.1 Performance

Table 5: Realised risk reduction (RR) in % for probabilities $p \in \{90\%, 95\%, 99\%\}$, using a naive and various HOPAC implementations, where A corresponds to Ali–Mikhail–Haq, C for Clayton, F for Frank, J for Joe. The risk reduction from (10) is the risk measure reduction with respect to no hedging.

Dependence	i.i.d.					GARCH$(1,1)$			
Model	Naive	A	C	F	J	AG	CG	FG	JG
$\text{VaR}_{90\%}$	**42.68**	37.43	38.11	39.12	37.7	34.17	**36.59**	35.48	35.31
$\text{VaR}_{95\%}$	37.24	38.47	37.25	38.31	**38.89**	**34.72**	34.62	33.91	31.80
$\text{VaR}_{99\%}$	23.95	22.95	**24.43**	23.63	24.34	16.79	20.85	**21.06**	17.72
$\text{ES}_{90\%}$	32.62	34.10	**34.21**	33.90	33.83	31.23	31.48	31.07	**31.94**
$\text{ES}_{95\%}$	28.54	30.99	30.63	31.47	**31.90**	**29.51**	28.73	29.31	29.29
$\text{ES}_{99\%}$	11.48	18.37	19.96	19.36	**23.37**	18.61	20.02	19.35	**20.53**

The highest risk reduction within each model category is denoted in bold.

Table 5 shows the ex post realised risk reduction (RR) including hedging cost for three probability levels $p \in \{90\%, 95\%, 99\%\}$, each risk measure, and each hedging approach. Over-

all, the HOPAC-i.i.d. methods attain the highest RR. The HOPAC-GARCH implementations underperform their HOPAC-i.i.d. counterparts in almost all instances. Besides, no copula family consistently outperforms, which is likely linked to the extra degrees of freedom from the OP-transform. Any differences are likely linked to the implied body dependence, which causes a different fit. The realised Value-at-Risk (VaR) for all HOPAC-i.i.d. models, is close to the naive method. The realised VaR is essentially based on one observation, which makes the metric more suseptible to noise. The realised Expected-Shortfall (ES) reduction for all HOPAC-i.i.d. models is higher than for the naive method. The HOPAC-GARCH models seem to outperform the naive method only when considering higher probability levels, however fail to outperform the HOPAC-i.i.d. models consistently. Unfortunately, expected shortfall is not elicitable such that there is no natural scoring function for comparing the realised risk (McNeil et al., 2015). Rather than testing the RR with one measure, I resort to testing the realised loss distribution with a Kolmogorov-Smirnov (K-S) test. An advantage of the K-S test is, that it tests the complete loss distribution rather than the (tail risk) measure exclusively. Furthermore, the K-S test can be used for both targetted risk measures.

Loss Distribution

Preferably, a model reduces the risk of high losses, whilst not increasing the severity of more common loss events. To establish which hedging implementation is preferred, I employ a one-sided two-sample Kolmolgorov-Smirnov (K-S) test (Smirnov (1939)) on the observed loss distributions. The one-sided K-S test can be used to establish the presence of first order stochastic dominance, which means that one model implementation is preferred regardless of the individuals risk aversion, provided some weak conditions (Massachusetts Institute of Technology [MIT], 2015). I formulate the null and alternative hypothesis as $H_0 : \hat{F}_{naive}(\text{P\&L}) \leq \hat{F}_{model}(\text{P\&L})$ and $H_1 : \hat{F}_{naive}(\text{P\&L}) > \hat{F}_{model}(\text{P\&L})$, where $\hat{F}_{model}(\cdot)$ denotes the empirical cdf from the model, given that a loss occurs in an unhedged scenario. A total of 746 unhedged losses occur, with which the cdfs are constructed. The K-S test is sensitive to autocorrelation in the observations.

Table 6: Kolmogorov–Smirnov one sided p-values with $H_0 : \hat{F}_{naive}(\text{P\&L}) \leq \hat{F}_{model}(\text{P\&L})$ and $H_1 : \hat{F}_{naive}(\text{P\&L}) > \hat{F}_{model}(\text{P\&L})$. I calculate p-values with the formula of Hodges (1958) using the effective sample-size adjustment from Xu (2013).

Dependence	i.i.d.				GARCH(1, 1)			
Model	A	C	F	J	AG	CG	FG	JG
$\text{VaR}_{90\%}$	0.02*	0.01*	0.01*	0.01*	0.04*	0.04*	0.06	0.02*
$\text{VaR}_{95\%}$	0.06	0.03*	0.02*	0.01*	0.01*	0.01*	0.01*	0.00*
$\text{VaR}_{99\%}$	0.08	0.06	0.03*	0.01*	0.01*	0.01*	0.00*	0.00*
$\text{ES}_{90\%}$	0.02*	0.01*	0.01*	0.01*	0.04*	0.04*	0.06	0.02*
$\text{ES}_{95\%}$	0.06	0.03*	0.02*	0.01*	0.01*	0.01*	0.01*	0.00*
$\text{ES}_{99\%}$	0.08	0.06	0.03*	0.01*	0.01*	0.01*	0.00*	0.00*

Rejection of the null-hypothesis is denoted with an asterisk *.

Xu (2013) proposes a straight forward, and intuitive, effective sample size adjustment.[4] I use the adjustment from Xu (2013) and subsequently apply the formula from Hodges (1958) to obtain p-values. Table 6 shows the resulting p-values. The p-values from Table 6 show that the naive P&L cdf tends to lie above almost all model implementations at a 5% significance level (and sometimes at 1%). A higher model cdf means that, at a fixed probability level, the P&L is higher than of the naive cdf. For example, if $\hat{F}_{naive}(\text{P\&L}) > \hat{F}_{model}(\text{P\&L})$, consider a probability level α. Then, $\hat{F}^{\leftarrow}_{naive}(\alpha) < \hat{F}^{\leftarrow}_{model}(\alpha)$ such that an α probability loss is lower for the naive approach. The naive implementation thus, at a fixed probability level, tends to incur lower losses than most HOPAC implementations. The naive implementation therefore exhibits first order stochastic dominance over most HOPAC models. The HOPAC-i.i.d. ES-performance is therefore linked to some low probability losses, which are successfully mitigated at the cost of more commonly occurring loss events.

[4]The adjustment from Xu (2013) can only be used in the case of AR(1) auto-correlation. In this case, an AR(1) model captures all auto-correlation.

5.2 Hedging Ratios

Table 7: Average hedging ratios with standard deviation in parentheses, before a 10% loss event is observed without hedging.

	A	C	F	J	AG	CG	FG	JG
Panel A:	Crude hedging ratios $\bar{\beta}^{C}_{10\%}$ (S.D.)							
$VaR_{90\%}$	0.93 (0.04)	0.93 (0.04)	0.93 (0.05)	0.93 (0.05)	0.90 (0.09)	0.90 (0.09)	0.90 (0.09)	0.90 (0.08)
$VaR_{95\%}$	0.94 (0.05)	0.94 (0.05)	0.95 (0.06)	0.94 (0.06)	0.91 (0.09)	0.91 (0.10)	0.90 (0.09)	0.92 (0.09)
$VaR_{99\%}$	0.98 (0.09)	0.98 (0.08)	0.98 (0.09)	0.98 (0.09)	0.94 (0.12)	0.94 (0.11)	0.94 (0.12)	0.95 (0.13)
$ES_{90\%}$	0.93 (0.04)	0.93 (0.04)	0.93 (0.05)	0.93 (0.05)	0.90 (0.09)	0.90 (0.09)	0.90 (0.09)	0.90 (0.08)
$ES_{95\%}$	0.94 (0.05)	0.94 (0.05)	0.95 (0.06)	0.94 (0.06)	0.91 (0.09)	0.91 (0.10)	0.90 (0.09)	0.92 (0.09)
$ES_{99\%}$	0.98 (0.09)	0.98 (0.08)	0.98 (0.09)	0.98 (0.09)	0.94 (0.12)	0.94 (0.11)	0.94 (0.12)	0.95 (0.13)
Panel B:	Gasoline hedging ratios $\bar{\beta}^{G}_{10\%}$ (S.D.)							
$VaR_{90\%}$	0.84 (0.08)	0.83 (0.07)	0.84 (0.08)	0.81 (0.08)	0.84 (0.13)	0.82 (0.13)	0.83 (0.13)	0.81 (0.14)
$VaR_{95\%}$	0.84 (0.08)	0.84 (0.08)	0.84 (0.08)	0.79 (0.08)	0.82 (0.13)	0.81 (0.13)	0.81 (0.13)	0.77 (0.13)
$VaR_{99\%}$	0.84 (0.08)	0.85 (0.08)	0.83 (0.08)	0.77 (0.08)	0.80 (0.13)	0.81 (0.14)	0.80 (0.13)	0.74 (0.14)
$ES_{90\%}$	0.84 (0.08)	0.83 (0.07)	0.84 (0.08)	0.81 (0.08)	0.84 (0.13)	0.82 (0.13)	0.83 (0.13)	0.81 (0.14)
$ES_{95\%}$	0.84 (0.08)	0.84 (0.08)	0.84 (0.08)	0.79 (0.08)	0.82 (0.13)	0.81 (0.13)	0.81 (0.13)	0.77 (0.13)
$ES_{99\%}$	0.84 (0.08)	0.85 (0.08)	0.83 (0.08)	0.77 (0.08)	0.80 (0.13)	0.81 (0.14)	0.80 (0.13)	0.74 (0.14)
Panel C:	Heating-oil hedging ratios $\bar{\beta}^{HO}_{10\%}$ (S.D.)							
$VaR_{90\%}$	1.11 (0.08)	1.10 (0.08)	1.09 (0.09)	1.11 (0.08)	1.11 (0.21)	1.13 (0.21)	1.12 (0.20)	1.13 (0.19)
$VaR_{95\%}$	1.10 (0.08)	1.10 (0.09)	1.10 (0.08)	1.11 (0.09)	1.11 (0.19)	1.13 (0.17)	1.11 (0.21)	1.13 (0.21)
$VaR_{99\%}$	1.09 (0.10)	1.09 (0.08)	1.10 (0.10)	1.14 (0.11)	1.11 (0.20)	1.10 (0.20)	1.12 (0.19)	1.11 (0.17)
$ES_{90\%}$	1.11 (0.08)	1.10 (0.08)	1.09 (0.09)	1.11 (0.08)	1.11 (0.21)	1.13 (0.21)	1.12 (0.20)	1.13 (0.19)
$ES_{95\%}$	1.10 (0.08)	1.10 (0.09)	1.10 (0.08)	1.11 (0.09)	1.11 (0.19)	1.13 (0.17)	1.11 (0.21)	1.13 (0.21)
$ES_{99\%}$	1.09 (0.10)	1.09 (0.08)	1.10 (0.10)	1.14 (0.11)	1.11 (0.20)	1.10 (0.20)	1.12 (0.19)	1.11 (0.17)

Table 7 shows hedging ratios, before at least a 10% unhedged loss event occurs. Hedging ratios before a significant loss occurs, grant some insight in model characteristics. The naive method always assumes full delta hedging, thus a hedging ratio of 1 in all futures. Notably the HOPAC implied mean hedging ratios are similar across risk measures, and across different probability levels. Furthermore, hedging ratios are nearly identical across the various HOPAC implementations. However, the HOPAC-GARCH ratios exhibit a higher standard deviation.

This higher standard deviation can have two sources; 1) parameter uncertainty, or 2) a deliberate adjustment due to increased risk as implied by the GARCH models. However, a two-sided K-S fails to reject difference between the loss distributions of the HOPAC-i.i.d. and HOPAC-GARCH implementations. Furthermore, Table 5 shows that the HOPAC-GARCH models fail to reduce risk more effectively than the more parsimonious models. Therefore, the higher standard deviation likely originates disproportionately from parameter uncertainty.

All models take hedging ratios close to 1 for crude ($\bar{\beta}^{C}_{10\%}$), lower than 1 for gasoline ($\bar{\beta}^{G}_{10\%}$), and higher than 1 for heating-oil ($\bar{\beta}^{HO}_{10\%}$). Therefore, it seems that the some of the gasoline price risk is hedged using heating-oil futures. The smaller gas (higher heating-oil) positions are likely linked to the higher (lower) volatility, and higher (lower) skewness as seen in Table 1. Lower volatility is desirable, because it represents more price certainty in the future. Lower (positive) skewness means that the probability of prices rising (causing losses for the refinery who is short) is slightly lower. The higher gasoline volatility is partially driven by consumer behaviour, which causes distortions like the summer driving season for instance (Downey, 2009). This additional price uncertainty means that gasoline exposure can, to some extent, be hedged more efficiently by using heating-oil futures, despite the lower cross dependence between gasoline spots and heating-oil futures (see Table 2).

5.3 Model Fit

Table 8: AIC and BIC model selection criteria, with parameter count k. The HOPAC-i.i.d. models {A, C, F, J} are nested in the HOPAC-GARCH {AG, CG, FG, JG} models. The reported statistics are averaged results across the 1541 estimation windows.

Model	A	C	F	J	AG	CG	FG	JG
k	10	10	10	5	28	28	28	23
AIC	-2337	-2329	-2334	-2258	-2229	-2210	-2241	-2155
BIC	-2301	-2294	-2298	-2241	-2130	-2110	-2142	-2073

The J and JG specifications collapse to Gumbel family models which have 5 parameters less.

I use the Akaike Information Critereon (AIC) and Bayesian Information Critereon (BIC) to evaluate the model fit. The HOPAC-i.i.d. implementations are nested in the HOPAC-GARCH models, which is easily verified through applying straight forward restrictions on the GARCH parameters. Table 8 shows the averaged AIC and BIC values across the 1541 windows. Both selection criterea always prefer the i.i.d. implementation within each family, indicating that the additional GARCH specification is unfavourable. The Joe family seems to lag behind the other three families in terms of model fit. Notably, the Joe parameter $\hat{\theta}$ consistently attains its lower bound of 1 across the entire sample for both i.i.d. and GARCH specifications. The OP-transformed Joe generator then collapses to a regular Gumbel generator. The parameter count k is therefore 5 for the HOPAC-i.i.d. or 23 for HOPAC-GARCH respectively. The Gumbel family apparently has a favorable body dependence compared to the Joe family, since the tail dependence characteristics are identical. Notably, the Gumbel family is the only Archimedean extreme-value copula (Hofert, 2010). Extreme-value copulas should theoretically be well suited to this hedging problem due to the positive dependence structure amongst the marginals, and interest in rare events. The extreme-value property might explain the attractive hedging performance shown in Table 5, however it comes at the cost more moderate losses shown in Table 6, and leads to a slightly worse fit compared to the other copula models as shown in Table 8.

6 Discussion and Conclusion

My findings show that HOPAC implementations can lead to a substantial reduction of the Expected-Shortfall (ES). However, the HOPAC implementations do not substantially reduce the Value-at-Risk (VaR). I also find that the assumption of independent observations, in favor of model parsimony ultimately enhance model performance in this setting. I also find that the HOPAC models imply stable hedging ratios, which many implementations fail to attain (Alexander et al., 2013).

I perform an ex-post evaluation on eight model-based hedging approaches using two different risk measures; VaR and ES. I find that the realised ES is superior to the naive approach for all HOPAC-i.i.d. implementations. However, both HOPAC-i.i.d. and HOPAC-GARCH imple-

mentations fail to consistently outperform the naive approach with respect to the VaR-measure. Besides, I find that the additional GARCH specification increases parameter uncertainty and harms out-of-sample performance. Furthermore, I show that the naive approach exhibits first order stochastic dominance, such that a refinery might still prefer the naive method for mitigating losses.

HOPAC-i.i.d. models outperform under the ES-metric, whilst they do not consistently outperform under the VaR-metric. This diverging finding is likely linked to the fact that the VaR-metric only considers one quantile observation of the (loss)distribution, whereas the ES-metric measures a wider segment of the tail. The HOPAC-i.i.d. models effectively mitigate some extreme losses, which substantially reduce the ES. Therefore, A refinery concerned with ES can reduce its risk using HOPACs, which might increase the attractiveness to investors. I conclude that HOPACs can be used effectively for modelling the complete joint distribution, and used it to mitigate extreme losses.

However, the naive implementation dominates both HOPAC implementations in the first-order stochastic sense. The stochastic domination is a result of the naive-method's ability to reduce moderate losses more effectively than the HOPAC counterparts. The HOPACs therefore might perform more favorably when targeting a different risk measure that captures moderate losses instead. Besides, I only consider HOPACs for estimating ideal next week hedging ratios. McNeil et al. (2015) states that the risk-management horizon should reflect the market conditions in which its core business activities lie. A horizon of one week is not necessarily appropriate, and could therefore be a subject for further research. Besides, an alternative risk management horizon can also benefit the i.i.d. assumption, since the conditional (significant autocorrelation) and unconditional (stationarity) distributions for all prices series differ.

The assumption of i.i.d. observations is invalid due to the presence of volatility clustering. However, despite the presence of volatility-clustering, the HOPAC-GARCH implementations perform unfavorably compared to the naive and HOPAC-i.i.d. approaches. Thus, assuming i.i.d. observations in favor of parsimony leads to a better performing model in this particular context.

Poor performance of the HOPAC-GARCH implementations might be exaggerated by the

two-stage estimation, which results in an efficiency loss. A unified estimation technique might reduce parameter uncertainty, and thus increase model performance. Besides, one might disregard GARCH models in favor of directly incorporating the time-dependence modelling into the HOPAC structure. Therefore, research into a different estimation technique, or an extension for HOPACs to capture time dependence, might lead to improved results.

References

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

CME Group. (2017). *Introduction to Crack Spreads*. Retrieved April 21, 2021, from https://www.cmegroup.com/education/articles-and-reports/introduction-to-crack-spreads.html

Downey, M. (2009). *Oil 101*. Wooden Table Press LLC.

Dunis, C. L., Laws, J., & Evans, B. (2008). Trading futures spread portfolios: Applications of higher order and recurrent networks. *The European Journal of Finance*, *14*(6), 503–521. https://doi.org/10.1080/13518470801890834

Energy Information Administration. (2002). Derivatives and risk management in the petroleum, natural gas, and electricity industries. *U.S. Department of Energy, Washington DC*.

Energy Information Administration. (2020). *Refinery Capacity Report 2020*. Retrieved April 21, 2021, from https://www.eia.gov/petroleum/refinerycapacity/table5.pdf

Energy Information Administration. (2021a). *NYMEX Future Prices*. Retrieved June 29, 2021, from https://www.eia.gov/dnav/pet/PET_PRI_FUT_S1_D.htm

Energy Information Administration. (2021b). *Spot Prices*. Retrieved June 29, 2021, from https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm

Górecki, J., Hofert, M., & Okhrin, O. (2021). Outer power transformations of hierarchical archimedean copulas: Construction, sampling and estimation. *Computational Statistics and Data Analysis*, *155*.

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.

Hodges, J. L. (1958). The significance probability of the smirnov two-sample test. *Arkiv för Matematik*, *3*(5), 469–486.

Hofert, M. (2010). *Sampling nested archimedean copulas with applications to cdo pricing* (Doctoral dissertation). Universität Ulm. https://doi.org/10.18725/OPARU-1787

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

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.

Marshall, A. W., & Olkin, I. (1988). Families of multivariate distributions. *Journal of the American statistical association*, *83*(403), 834–841.

Massachusetts Institute of Technology. (2015). *Microeconomic theory iii*. Retrieved April 21, 2021, from https://ocw.mit.edu/courses/economics/14-123-microeconomic-theory-iii-spring-2015/lecture-notes-and-slides/MIT14_123S15_Chap4.pdf

McNeil, A. J. (2008). Sampling nested archimedean copulas. *Journal of Statistical Computation and Simulation*, *78*(6), 567–581. https://doi.org/10.1080/00949650701255834

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools-revised edition*. Princeton university press.

Nelsen, R. B. (2007). *An introduction to copulas*. Springer Science & Business Media.

Okhrin, O., Okhrin, Y., & Schmid, W. (2013). On the structure and estimation of hierarchical archimedean copulas. *Journal of Econometrics*, *173*(2), 189–204.

Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Publications de l'Institut Statistique de l'Université de Paris.

Smirnov, N. V. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bull. Math. Univ. Moscou*, *2*(2), 3–14.

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

Xu, X. (2013). *Methods in hypothesis testing, markov chain monte carlo and neuroimaging data analysis* (Doctoral dissertation). Harvard University. Cambridge, Massachusetts.

Appendix A Constant Maturity Futures Results

Here, I briefly discuss the hedging results obtained when using the *constant-maturity* futures construction method suggested by Alexander et al. (2013). Furthermore, Alexander et al. (2013) provide a comprehensive review of the method. I use *contract 3* futures from the Energy Information Administration (2021a) due to the limited time-span availability for *contract 2* futures. The *contract 3* futures expire two months after *contract 1* futures. Therefore, I use a slightly different expiry time than Alexander et al. (2013); $T = 60$ instead of $T = 44$. Next, I apply the same data adjustments as with the *rollover* method.

Table 9: Realised risk reduction (RR) in % for probabilities $p \in \{90\%, 95\%, 99\%\}$, using a naive and various HOPAC implementations using *constant-maturity* futures, where A corresponds to Ali–Mikhail–Haq, C for Clayton, F for Frank, J for Joe. The risk reduction from (10) is the risk measure reduction with respect to no hedging.

Dependence	i.i.d.					GARCH(1,1)			
Model	Naive	A	C	F	J	AG	CG	FG	JG
$\text{VaR}_{90\%}$	**25.20**	21.72	23.51	21.54	24.41	**26.66**	24.55	25.57	23.66
$\text{VaR}_{95\%}$	24.76	24.43	23.71	23.35	**26.78**	26.17	**27.57**	25.09	26.63
$\text{VaR}_{99\%}$	**28.91**	24.88	28.71	24.91	25.20	**25.56**	24.41	23.49	20.80
$\text{ES}_{90\%}$	26.91	26.94	27.27	27.17	**27.48**	27.32	27.16	**27.44**	27.09
$\text{ES}_{95\%}$	27.62	28.11	27.53	27.35	**28.19**	26.67	**28.15**	27.51	27.99
$\text{ES}_{99\%}$	23.86	23.55	**23.97**	22.79	22.51	**23.01**	21.52	21.76	22.92

The highest risk reduction within each model category is denoted in bold.

Table 9 shows the resulting RR in % compared to no hedging. The RR using the *constant-maturity* futures is lower, but appears more consistent across probability levels compared to Table 5. The HOPAC-GARCH implementations seem to perform slightly better with the *constant-maturity* futures, although no single implementation consistently outperforms the naive or HOPAC-i.i.d. approaches across measures and probability levels. However, instead of outperforming, the HOPAC-i.i.d. models realize comparable RR compared to the naive

method. The HOPAC-GARCH models seem to perform slightly better, but fail to significantly outperform the naive and HOPAC-i.i.d. models.

A.1 Stochastic Dominance

Table 10: Kolmogorov–Smirnov one sided p-values with H_0 : $\hat{F}_{naive}(\text{P\&L}) \leq \hat{F}_{model}(\text{P\&L})$ and H_1 : $\hat{F}_{naive}(\text{P\&L}) > \hat{F}_{model}(\text{P\&L})$. I calculate p-values with the formula of Hodges (1958) using the effective sample-size adjustment from Xu (2013).

Dependence	i.i.d.				GARCH(1, 1)			
Model	A	C	F	J	AG	CG	FG	JG
$\text{VaR}_{90\%}$	0.395	0.395	0.316	0.14	0.423	0.292	0.292	0.395
$\text{VaR}_{95\%}$	0.367	0.395	0.316	0.114	0.395	0.423	0.452	0.341
$\text{VaR}_{99\%}$	0.697	0.636	0.667	0.226	0.207	0.395	0.636	0.316
$\text{ES}_{90\%}$	0.395	0.395	0.316	0.14	0.423	0.292	0.292	0.395
$\text{ES}_{95\%}$	0.367	0.395	0.316	0.114	0.395	0.423	0.452	0.341
$\text{ES}_{99\%}$	0.697	0.636	0.667	0.226	0.207	0.395	0.636	0.316

Table 10 shows that the Kolmogorov-Smirnov (K-S) test fails to reject the null-hypothesis, under which the naive method is preferred by virtue of stochastic dominance. Failure to reject the K-S null-hypothesis means that all HOPAC models exhibit better behavior with regards to moderate losses compared to the *rollover* futures approach. Besides testing for the stochastic dominance of the naive method, I also test for dominance of the HOPAC models. The K-S test with the alternative hypothesis of stochastic dominance for HOPAC implementations, leads to p-values even closer to 1 which is why I omit the results for brevity. Therefore, the HOPACs do not perform significantly worse or better in terms of stochastic dominance.

A.2 Model Fit

Table 11 shows the model fit across the 1541 estimation windows. Both AIC and BIC favor the restricted HOPAC-i.i.d. models in favor of the additional GARCH specification. The

Table 11: AIC and BIC model selection criteria, with parameter count k. The HOPAC-i.i.d. models {A, C, F, J} are nested in the HOPAC-GARCH {AG, CG, FG, JG} models. The reported statistics are averaged results across the 1541 estimation windows.

Model	A	C	F	J	AG	CG	FG	JG
k	10	10	10	5	28	28	28	23
AIC	-2200	-2197	-2191	-2112	-2101	-2086	-2107	-2019
BIC	-2164	-2161	-2155	-2095	-2001	-1986	-2008	-1937

The J and JG specifications collapse to Gumbel family models, which have 5 parameters less.

alternative futures specification does therefore not substantially impact conclusions based on model specification criteria.

A.3 Conclusion Constant Maturity Futures

The HOPAC performance using the *constant-maturity* futures is lower in terms of absolute risk reduction. Also, the HOPAC-GARCH models seem to benefit moderately from the construction of *constant-maturity* futures. However, the performance of all HOPAC implementations is arguably better in terms of stochastic dominance, compared to the implementations when using the *rollover* futures approach. Both futures approaches therefore entail different potential. For instance, the *rollover* method has the potential of greater risk reduction, likely linked to stronger dependence between futures and spots. Whereas the *constant-maturity* futures have greater potential of attaining stochastic dominance over the naive method, for example by use of an alternative risk measure or time management horizon.